DOCUMENT RESUME

ED 231 364 IR 050 238

AUTHOR Friend, Linda; Bonta, Bruce
TITLE Reference Use of Online Databases: An Analysis.
PUB DATE Mar 81
NOTE 10p.; Paper presented at the National Online Meeting (New York, NY, March 24-26, 1981).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Libraries; Costs; Databases; Higher Education; *Information Retrieval; Library Expenditures; *Online Systems; *Recordkeeping; Reference Materials; *Reference Services; Relevance (Information Retrieval); Tables (Data); Time
IDENTIFIERS *Library Statistics; *Pennsylvania State University

ABSTRACT

Since July 1978 the general reference section of the Pennsylvania State University library has compiled statistical information related to the use of online searching for reference purposes. An evaluation form completed by the librarian for each online reference search recorded such information as search date; status of requestor; printed sources consulted and time involved; online databases used with time and cost figures; and a statement as to whether the search was successful. This paper summarizes the information collected on the evaluation forms and identifies at least five major categories of questions that are feasible for online reference, particularly in a large academic library: (1) directory-type information; (2) literature retrieval on a limited scale; (3) bibliographic verification of specific works; (4) definitions or explanations of concepts or acronyms; and (5) location of book reviews. It is noted that a detailed statistical study showed that the actual nature of a question was what seemed to determine search success or failure. Five tables and seven references are provided. (Author/ESR)

REFERENCE USE OF ONLINE DATABASES: AN ANALYSIS

Linda Friend

and

Bruce Bonta

REFERENCE USE OF ONLINE DATABASES: AN ANALYSIS

Linda Friend and Bruce Bonta, The Pennsylvania State University
University Park, Pennsylvania, U.S.A.

This paper examines statistical information compiled from reference uses of the computer since 1978 in the General Reference Section at Penn State. An evaluation form was completed by the librarian involved each time the computer was used to help answer a reference request, and we collected such information as: date; status of requestor (faculty, undergraduate, etc.); printed sources consulted and time involved; online databases used with time and cost figures; and a statement of whether the search had been successful. Both the questions asked and the ways the terminal was used to answer them varied widely, but upon analysis some patterns have appeared. We have identified at least five major categories of questions that are feasible for online reference, at least in a large academic library: 1). directory-type information; 2). literature retrieval on a limited scale; 3). bibliographic verification of specific works; 4). definitions or explanations of concepts or acronyms; 5). book reviews. A detailed statistical study concluded that the actual nature of the question seemed to determine search success or failure.

1. INTRODUCTION

Computerized literature searching has been available on a fee basis to our clientele at Penn State since 1975. As the librarians' online expertise developed, and as more databases in more subject fields became available from the vendors, we discovered another attractive use for our computer capabilities—to answer various reference questions that could not be readily found in the printed sources available, or that we felt could be answered more economically online. This dimension of reference service resulted in a steady trickle of terminal uses for ready reference, and all five of the current Penn State searching locations (Life Sciences, Government Documents, General Reference, Physical Sciences, and Engineering) have used the terminal to some extent to answer reference questions.

In July 1978 the General Reference Section began to keep a fairly detailed record of each attempt to use one or more online databases in answering questions for patrons. Our purpose, aside from professional curiosity, was to amass enough data to justify our request for a terminal at the Reference Desk (we presently use the literature searching terminal, located in a separate room at least 50 feet away from the desk). Since General Reference is primarily responsible for reference service to the social sciences and humanities, this study will reflect that emphasis, which is particularly apparent when we detail the databases used and printed sources consulted.

2. RESEARCH METHODOLOGY

Although much has been written on comparisons of manual vs. computerized literature searches, we could find very little in the literature to use as a guide for developing the reference aspect of computer searches (1, 2, 3). To begin the study, we listed all the types of information likely to be useful in evaluating such a service. One of our main purposes was to discover the various ways the computer could be creatively used to aid us in our work. At the National Online Meeting in 1980 Jan Egeland of BRS discussed a number of those uses (4):

- -when a term or concept is difficult to locate;
- -citation verification;
- -searching a file you do not have in hard copy;
- -use of CROS [or Lockheed's DIALINDEX] to direct patrons to the most appropriate reference source;
- -to find subject descriptors or terminology most appropriate to use in a manual index;
- -a few good "hits" to get a patron started, and
- -directory information.

In the 232 uses of the computer that we have evaluated so far, all of the above possibilities were explored at one time or another, and will be examined using the five categories we ourselves identified.

3. THE EVALUATION FORM

The evaluation form that the Section began using in July 1978 asked for the date; searcher's name; status of requestor; the question; specific printed sources used, if any, and time involved; online databases used; online time and cost; and an evaluation by the searcher of whether or not the online search could be considered successful. The completed evaluation sheet was stapled to the search printout (patrons could photocopy the printout if they wanted to) and filed for future reference. After 30 months, July 1978—December 1980, and 232 documented reference uses of the terminal, some patterns have begun to appear.

3.1 Status of Requestor

At Penn State's main campus the Library serves about 38,200 students (Fall 1980 figures), the majority of whom are undergraduates. This group is traditionally the least likely to make use of our fee-based literature search service, so we were curious to see whether we were reaching some of them through our quick reference searching. Fifty-four of the searches, or 23%, were for undergraduates, and, not surprisingly, the vast majority of their requests were answered by supplying literature retrieval on a limited scale. Of the remaining requests, 34 (15%) were for teaching faculty; 58 (25%) for library faculty or staff, or other university staff; 52 (22%) were for graduate students; 13 (6%) for "others" (mostly people from the surrounding community or local businesses); and 21 (9%) were unknown. Since there has never been, within memory of the present library staff, an analysis of the users of the reference room, it is tempting to suspect that similar proportions of patrons may apply to the harder questions received at the desk; this postulation would need further research to confirm or deny it.

4. CATEGORIES OF QUESTIONS

Certainly the most interesting part of our study turned out to be an analysis of the questions themselves, and the variety of ways the terminal was used to answer them. For example:

1). A faculty member wanted to find references to any dissertations that related to athletic departments and their structures within colleges or universities. The search through the printed dissertation sources had been futile, but when a librarian tried Comprehensive Dissertation Index online, one excellent dissertation was found in just three minutes.

2). A faculty member needed to find the correct bibliographic information for an article by Benjamin DeMott that had appeared in American Psychologist at an unknown date. After giving up in frustration with the printed indexes, he asked a librarian for help; in three minutes online, the correct reference to Psychology Today was found.

3). A graduate student needed information on the elderhostel program and was having very little luck. Five minutes online produced five excellent citations.

4). An undergraduate was looking for information on the motor skills needed in dart throwing. After spending over an hour with the printed Psychological Abstracts searching under motor performance, she came to a librarian who, at a cost of less than $4.00, found exactly the citations needed.

5). A librarian needed to find a government document (for a class on the brewing industry) that she thought was called "Beer Is a Four-Letter Word." After spending ten minutes with the printed Monthly Catalog, she tried the corresponding online database and found that the document was called "Is Beer a Four-Letter Word?"

In examining our terminal uses (See Table 1), we found that they could be categorized like any other reference questions. The five categories we set up reflect the kinds of printed reference sources one might use to find the answers:

I. Directory Type: institutions, names, addresses, biographical information

II. Literature retrieval on a limited scale*

III. Bibliographic verification of specific works: books, dissertations, newspaper articles, documents, reports, conference proceedings

IV. Dictionary or encyclopedia type: definition of a term or acronym; explanation of a concept

V. Book reviews

---

*Literature retrieval on a limited scale is fairly distinct from fee-based literature searching. The former is initiated by the reference librarian when online access seems to be the logical way to handle a question, and the result is usually up to ten references, with perhaps five minutes spent online. The full-scale literature search is requested by the patron and implies a thorough coverage of a topic.

CATEGORIES OF QUESTIONS								
	A	B	C	D	E	F	G	H
Category of Question	Number of Questions in that Category	Questions per Category as % of Total	Number of Uses of Individual Databases	Average Number of Databases Used per Search (C ÷ A)	On-line Time in Minutes	Average On-line Time per Search in Minutes (E ÷ A)	Number of Successful Searches	Percent Successful (G ÷ A)
I. Directory	24	10%	37	1.6	106	4.4	8	33%
II. Literature	109	48%	182	1.7	541	5.0	95	87%
III. Bibliographic Verification	89	39%	126	1.4	324	3.6	53	60%
IV. Dictionary or Encyclopedia	8	7%	13	1.6	32	4.0	6	79%
V. Book Review	2	1%	2	1.0	5	2.5	1	50%
TOTALS	232	100%	360	1.5	1008	4.3	163	70%

Table 1: Categories of Questions

Referring to Column H in Table 1, the low success rate of directory-type questions (33%) relates to the use of printed directory sources. Questions at the reference desk requiring directory information are usually answered pretty readily by a standard group of printed sources kept a few feet behind the desk. The more difficult questions require us to range farther afield, and at this point an online database used as a reference source is probably just as likely to provide an answer as a printed source.

5. BIBLIOGRAPHIC VERIFICATION

Bibliographic verification questions, Category III, were further subdivided by type (proceedings, articles, etc), as shown in Table 2. The single advertisement verification is another good example of successful reference use of the terminal. During the week after John Lennon was killed, a local minister planning a memorial service came in to try to locate a full page ad on peace that Lennon had placed in the New York Times sometime in 1979. Using the New York Times database on BRS, a librarian found the correct citation in less than three minutes; a manual search could have taken hours.

Type	Number	Percentage
Dissertations	27	30
Journal articles	22	25
Books	16	19
Reports	5	5
Journal titles	5	5
Conference proceedings	4	4
Documents	2	3
Newspaper articles	2	3
Advertisements	1	1
Not clearly specified	5	5
	89	100

Table 2: Question Types Within Category III

6. DATABASES AND PRINTED SOURCES USED

Fifty-three different online databases were consulted at least once, though only fourteen were used at least six different times and seem worth noting. (Table 3)

Database	Number of Uses
ERIC	89
Psychological Abstracts	55
Comprehensive Dissertation Index	49
Social Sciences Citation Index	25
Magazine Index	21
National Newspaper Index	15
ABI/Inform	9
Management Contents	8
Exceptional Child Education Resources	7
Language and Language Behavior Abstracts	7
Newsearch	7
Books Info	6
CROS	6
NTIS	6

Table 3: Online Databases Used

One hundred thirty-six different printed sources were used at least once, but of these only 17 were used six or more times. (Table 4)

Printed Source	Number of Uses
ERIC (RIE or CIJE)	45
Psychological Abstracts	39
Comprehensive Dissertation Index or Dissertation Abstracts International	22
Card Catalog	22
Social Sciences Citation Index	20
Reader's Guide	18
Education Index	18
Books in Print	13
New York Times Index	12
Cumulative Book Index	8
National Union Catalog	7
Encyclopedia of Associations	7
Business Periodicals Index	6
Forthcoming Books	6
Ulrich's International Periodicals Directory	6

Table 4: Printed Sources Used

We were not surprised to find that four of the top five sources in each list directly correspond. Judging by the number of printed book and journal title verification sources used, once our reference librarians have access to RLIN it will undoubtedly place high in the database ranking, certainly among the first five. We will have our opportunity to judge when the RLIN database becomes available to us in the first half of 1981.

7. TIME AND COST

All 232 evaluation sheets indicated the amount of online time involved in the search, and most estimated the amount of time a librarian had spent on manual searching as well. Table 5 indicates the results of this aspect of our study.

	Average Time of Librarian's Search in Minutes	Cost of Librarian Time at $10/hour*	Average Cost Per Search for Online Time	Average Total Cost Per Search
Manual Search	24.0	$4.00	---	$4.00
Online Search	4.3	$0.71	$3.08	$3.79

*$10.00 per hour was chosen as a reasonable compromise with Saffady's $9-12 per hour figure in 1979. (5)

Table 5: Time and Cost for Online and Manual Searches

Aside from librarian time and online time, the many other costs of online searching have been consciously omitted since the terminal was used primarily for literature searches throughout the study period, and for reference retrieval only if the terminal was not being used for a scheduled search. (See (6) and (7) for fuller treatments of the general question of costs).

The most interesting result of the calculations in Table 5 was that there appears to be an insignificant difference in the average cost of searching them manually. We would therefore not expect the librarians to exhaust all printed sources before turning to an online database, and as with any reference question, a trained professional should turn to whatever sources are most likely to answer a query, in the most efficient way possible.

Sometimes the patron's time was also noted on the form, and allows speculation on the amount of time our readers may waste in fruitless searches for information. Their time was recorded on many of the evaluation sheets, at an average of 57 minutes per search. These statements of patrons' time were based on chance comments ("I've been searching for an hour now, and haven't found anything."), and are probably not precise.

8. THE FUTURE

Once a terminal is provided at the Reference Desk, its use for answering reference questions will increase steadily. Every month new online databases are announced that would have obvious applicability for ready reference use. Some correspond to printed reference books, where the online use would offer new forms of access or more up-to-date information, while others have no printed counter part. As use of the databases for ready reference work continues to develop nationwide, it will reach the point where the first reference librarian on duty in the morning will dial-up the local telecommunication node (Telenet or Tymnet) and the reference desk terminal will stay connected all day. The librarians will thus be able to quickly choose different vendors and databases that they need to use on a moment by moment basis. Needless to say, lengthy and time-consuming evaluation sheets will be replaced by statistics gathering by sampling methods.

Other than the growth of this service, which seems so obvious and inevitable, the nature of the reference questions and the manner in which librarians handle them is likely to change also. The above discussion of five different categories of reference questions which were handled online is based unavoidably on the printed-reference-book mode of thinking. While librarians commonly categorize questions into directory type, dictionary type, encyclopedia type and so forth, it is really the category of the reference book used to answer the question that is applied to the question itself. The terms commonly used to categorize online databases-- bibliographic and non-bibliographic--hardly fit the need for evaluating and understanding the online ready reference questions.

Thus, a clear description of the categories of questions which may be answerable online is certainly not possible at this time. New types of reference desk services--such as interactive manipulation of statistical databases--could represent categories of reference questions which are not at all part of present reference desk work. The only conclusions possible are that online ready reference service is certain to grow and to change significantly the nature of reference service as we know it today.

REFERENCES

1. Sweetland, James H. "Using Online Systems in Reference Work." Online 3 (3) July 1979, pp. 10-19.
2. Raedeke, Amy, "Machine Assisted Reference Service in a Public Library: a One Month Test Period." Online 2 (4) October 1978, pp. 56-59.
3. Frazier, Ken. "Subsidized Online Search Service in an Agricultural Library." Agricultural Libraries Information Notes 6 (9/10) September-October 1980, pp. 1-4.
4. Egeland, Janet. "Reference Database Use." Paper presented at National Online Information Meeting; March 25-27, 1980.
5. Saffady, William. "The Economics of Online Bibliographic Searching: Costs and Cost Justification." Library Technology Reports, September-October 1979, p. 597.
6. Saffady, pp. 567-653.
7. Elchesen, Dennis R. "Cost-Effectiveness Comparison of Manual and On-line Retrospective Bibliographic Searching." Journal of the American Society for Information Science 29, March 1978, pp. 54-66.