International Symposium
on

NEW TECHNIQUES AND APPLICATIONS IN LIBRARIES

图书馆新技术应用国际学术讨论会

Xi'an, P. R.C. Sept. 8-11, 1988

Sponsored by
Xi'an Jiaotong University, P. R. C.
Ohio University, U.S.A.

Cooperated by
Kanazawa Institute of Technology, Japan

Donated by

B. H. Blackwell, Ltd.
Blackwell North America, Inc.
Lotus Tours Ltd.
National Library and Information Committee for
    Universities and Colleges, P. R. C.
National Natural Science Foundation of China
Online Computer Library Center

NEW DIRECTIONS FOR RESEARCH
IN ONLINE PUBLIC ACCESS CATALOGS

Raya Fidel
Graduate School of Library and Information Science
University of Washington, FM-30
Seattle, Washington 98195, USA
(206) 543-1888

ABSTRACT

The introduction of online public access catalogs (OPACs) brought a qualitative change in information science research: large-scale analyses of user requirements for catalogs were performed for the first time, and the distinction between catalogs and bibliographic databases began to diminish. Further OPAC investigations should be integrated with the study of databases to create a multi-faceted body of research, simultaneously studying areas such as human-computer interface, subject access, and user training.

1. INTRODUCTION

Most online public access catalogs (OPACs) are actually computerized versions of a manual catalog: the records that were used to reproduce catalog cards have been made available for information retrieval through direct, online interaction. Neither cataloging nor online searching are new phenomena in libraries, yet the seemingly simple technological transition that made catalog records available for online searching introduced both quantitative and major qualitative changes in library and information science research.

The quantitative change is manifested by the steady growth in the number of catalog studies and in their comprehensiveness. OPACs have also generated renewed interest in areas such as catalog use, Library of Congress Subject Headings, and bibliographic instruction as topic for research.

Most important, however, is the effect of the introduction of OPACs on the direction of research: (1) for the first time, information in online systems is actually considered from the user point of view; (2) the traditional distinction between online bibliographic databases and library catalogs is diminishing so that "indexing" and "subject cataloging," for example, are one and the same, as are "thesauri" and "subject-heading lists"; and (3) the study of catalogs has become multi-faceted and integrated--it draws from and contributes to variety of areas that previously were isolated from one another.

The purpose of this paper is to briefly examine these qualitative changes and to indicate a new direction for OPAC research.

2. PREVIOUS RESEARCH

OPACs are a relatively new phenomenon in libraries. During 1981-83, in the United States, the Council on Library Resources (CLR) funded a nationwide study of 17 OPACs (both in-house and commercial) in nearly 30 libraries.

While many articles about experience with online catalogs in specific libraries have been published (e.g., Ref. 1, Ref. 2, Ref. 3), this first large-scale study provides most of our knowledge about the OPAC. It was conducted by various agencies, and although attempts to summarize and synthesize its results have been made (e.g., Ref. 4), further interpretation and integration of the vast amount of data are required before specific conclusions can be drawn.

Typical of a large-scale study--although novel in research about online searching behavior--the CLR study applied a variety of methods. A survey of 8,094 users and 3,981 non-users of OPACs collected data through a questionnaire administered by various agencies (Ref. 5). Focused-group interviews with library staff and patrons--both users and non-users of online catalogs--were carried out in six libraries; they were conducted through open, in-depth discussions led by a moderator (Ref. 6). The results from the focused-group interviews were complemented by the results of individual and group interviews conducted among library staff at three research libraries (Ref. 7). In addition, transaction log analysis, in which the protocols of individual searches are analyzed, were performed in seven libraries (Ref. 8, Ref. 9). The CLR study also stimulated further analyses of the data, additional explorations, and considerations of possible implications.

The CLR study was done at the beginning of the OPAC era, when neither users nor librarians were experienced with the new tool. Though its findings cannot provide much insight into searching behavior, they do point to the areas in which research is needed. In a summary of the study, Matthews and Lawrence (Ref. 4) outlined the principal findings:

--Experience with the library and its catalog is the most important factor in determining success in using the catalog and satisfaction with it.
--OPACs should have a variety of interfaces available, depending on the type of search and the level of user experience.
--Users adapt to the capabilities and limitations of the OPAC being used.
--The form and nature of user training and user assistance are important.

These findings substantiate common knowledge among librarians; they do not provide new information about how OPACs should be designed.

The spread of OPACs in Great Britain has been slower than in the United States. Nevertheless, much research has been conducted there, most of it sponsored by the British Library Research & Development Department. While only a few large-scale studies have been completed, researchers take a multi-faceted approach: they study retrieval techniques and interface design, as well as the quality of the database. Controlled experiments are common, and most are based on theoretical developments in the area of bibliographic databases. In addition to direct use-studies, researchers have conducted comparative studies--either before-and-after surveys or examinations of the performance of different systems--and created experimental systems to study user-catalog interface (Ref. 10). Findings from these various projects (some of which are still incomplete) have yet to be integrated.

In summary, OPAC studies have focused on user requirements. Most studies have been set up to discover users' attitudes to and acceptance of the new catalogs by examining characteristics of both users and catalogs. The majority of these studies, however, have concentrated on the human-computer interface--that is, on how easy it is to "converse" with the computer. Only a few have addressed the retrieval problem: whether users are satisfied with the results of their searches and what could be done to improve the results. The problem of retrieval, though somewhat neglected in OPAC research, is central to research about searching bibliographic databases.

3. ONLINE BIBLIOGRAPHIC DATABASES

The study of information retrieval from bibliographic databases (e.g., Chemical Abstracts or ERIC) has been considered to be an independent, self-contained subject area that is different from the study of information retrieval from catalogs. However, information retrieval from bibliographic databases and OPACs is one and the same process: searching for bibliographic information. There is a diminishing distinction between the database as a store of citations to journal articles which can be manipulated with sophisticated techniques, and the OPAC as a store of information about monographs which can be manipulated with simple techniques. Consequently, before examining OPAC research, it is important to review briefly the approach taken by studies about searching bibliographic databases.

The first search systems for bibliographic databases were designed for professional searchers, primarily librarians. Research in online searching began in the 1970s with two large-scale studies that focused on user attitudes, satisfaction, and success (Ref. 11, Ref. 12).

Most of the research that followed, however, concentrated on the attributes of a "good searcher": the personal characteristics a librarian should possess to become a successful online searcher. Experience in online searching (Ref. 13), type of training (Ref. 14), personal traits such as creativity (Ref. 15) and cognitive abilities (Ref. 16) are among the characteristics examined. No conclusive results have been found, and most investigators have observed that their studies are impeded by the large individual variability among searchers.

There may be a number of reasons why these studies have failed to produce conclusive results, but the failure should not be of concern. While studies that identify the characteristics of a good searcher may at times help administrators to decide which employees should perform online searching in their libraries, the issue of personal characteristics is of very little significance. This is particularly so when both OPACs and bibliographic databases are searched by library users themselves.

The important issue is to discover the characteristics of a "good search." Identifying strategies and moves that can enhance the success of online searches is beneficial to users of all online retrieval systems.

Online searchers have long known the importance of the search process, but research in this area is sparse. Based on the experience of librarians, Bates has proposed a number of information search tactics that could be employed in online searches (Ref. 17). Fidel, on the other hand, analyzed search protocols and verbal protocols of seven experienced searchers performing approximately 80 searches as part of their regular workload. As a result, she listed the moves--changes in search formulations--that searchers made to increase the size of the retrieval, to decrease it, or to improve it altogether (Ref. 18).

Studying the search process is a complex task requiring the probing of phenomena that are not easy to observe or analyze. Nevertheless, user training and the design of useful online retrieval systems will not be successful until this process is thoroughly understood.

4. USER SEARCHING

OPAC studies are guided by an approach that is essentially different from the approach taken by researchers in the area of online searching in bibliographic databases. In studying OPACs, researchers assume that the user population is a given, and that features of the catalogs themselves should be

examined in order eventually to design an online catalog that is most useful. In contrast, recent research on retrieval from online databases assumes that it is the databases and the search systems (e.g., DIALOG, BRS) which are a given--thus research focuses instead on the characteristics of users (i.e., professional searchers) that supposedly will guarantee successful online searching.

Surely, the difference between the two research approaches stems from economic reality. Library administrators do not elect which users to serve but they are free to select the system that provides the best catalog for their users. In contrast, the same library administrators--operating in the world of competitive economy--cannot change commercial databases and search systems but must instead select among a number of candidates the persons who will perform best on these online systems.

The difference in research approaches is an impediment to both areas of research: OPAC studies cannot rely on research that already has been performed in online searching, and vice versa. There is one exception: studies of users searching their own requests in bibliographic databases (rather than studies of search intermediaries only) may have direct implication to the design of OPACs.

The idea that users can search bibliographic databases for their own requests is rapidly gaining popularity. While several in-house studies have been performed on gateway systems that provide a simple interface with search systems, studies of users searching their requests directly would have far more impact. The only example of such a study is a research project that examined 11 years of searching by pathologists and pharmacists in the databases of the United States' National Library of Medicine (Ref. 19). The results of this study are many, but as yet not all have been published. Of particular interest are the findings that users did not perceive that they had many problems with the techniques of searching, but rather with the vocabulary and content of the system, and that most performed simple searches. Another finding revealed that they encountered major problems with the more sophisticated capabilities of the database, problems that sometimes caused a substantive loss of citations.

Although this study examined a specific population, its results substantiate the findings of OPAC studies: the weakest point in users' searching their own requests is their inability to formulate successful strategies. Existing gateway systems do not provide help because they do not employ any of the sophisticated capabilities of databases, and some even eliminate the simple capabilities.

5. NEW DIRECTION: INTEGRATED RESEARCH

There are two major reasons why integrated and multi-faceted research is needed.

First, the widespread and ever-increasing use of automated systems for information retrieval enables researchers to investigate phenomena under actual, rather than experimental, conditions. The initial large-scale studies of OPACs and bibliographic databases provided general and descriptive data about use, satisfaction, and problems. While important in pointing to research needs and in providing evidence to substantiate common knowledge, most of these data cannot be used directly for systems use or in systems design. At present, however, librarians and users are routinely searching online systems, and researchers can observe and analyze their searching behavior. Such analyses, based on actual searching, provide a deeper understanding of the search process because they can answer the "why"

questions. They also facilitate investigations into individual variability in online searching, and they will produce findings that are relevant to systems use and design.

Second, recent developments in libraries are eliminating the traditional barriers among specialties in librarianship. For instance, searches of OPACs and bibliographic databases are essentially the same. In fact, in the future, libraries probably will provide access to both OPACs and bibliographic databases through a single interface. Realizing this, one sees immediately that OPAC studies would be more useful if they examined the quality of retrieved sets rather than defining success as a non-zero hit; and that research on bibliographic databases should look at system features. Similarly, investigating subject access through OPACs is no longer limited to the study of Library of Congress Subject Headings, but must consider other controlled vocabularies, indexing (manual and, possibly, automated), full-text retrieval, and expert systems.

Thus, the future calls for integrated, rather than fragmented, research. Studies should be designed to make their results relevant to the searching of both OPACs and bibliographic databases. Moreover, with the descriptive data already collected, research in online retrieval systems can now focus on its central issue, the search process. Examination of the search strategies that are available, and the conditions under which each of them should be used to achieve satisfactory retrieval, is necessary for the research to be relevant to systems use and design.

This integrated research should simultaneously address a number of areas. Human-machine interface, subject access, the cataloging code, the librarian's role, and user training are examples of areas that are briefly discussed here.

(a) Human-Machine Interface. At present, there is some evidence that users have technical problems when communicating with the current automated systems in libraries, but the nature of these problems is not yet clear. Studies of users performing actual searches can identify the nature of these problems and guide systems designers in improving the human-computer interface. Such studies should examine problems in using commands and in understanding the computer's responses and displays.

Special attention should be given to factors that are important to the search process, such as user's experience or level of motivation. It is likely, for instance, that inexperienced users and experienced ones would encounter different problems, and consequently might require different sets of commands or computer displays.

(b) Subject Access. Research on subject access through automated systems should cover two main themes: user aids and models of the search process. A variety of existing schemes (e.g., Dewey Decimal Classification) and arrangements (e.g., alphabetic display of the Library of Congress Subject Headings) can be made available to users to enhance subject searching. These aids should be tested with actual users, and their effectiveness and suitability for specific user needs and characteristics should be compared.

Most of the aids to subject access suggested in the literature are automated versions of aids that already exist in print form. Effort should be put into discovering new aids that are applicable in the online environment, regardless of their suitability to printed tools.

Models of tactics and moves used to improve results of subject searches, as well as a model of the selection of search keys--whether free-text terms or descriptors--already exist (Ref. 17, Ref. 18, Ref. 20). They have been developed for bibliographic database searching, and their applicability to

OPACs should be tested. Further, other processes in the search for bibliographic information need to be described by such models: deciding when to use automated tools and when to use printed tools; the selection of databases to search; and deciding when to stop searching.

(c) The Cataloging Code. Unlike subject cataloging, descriptive cataloging seems to be untouched by OPAC research. The Anglo-American Cataloguing Rules (and any other cataloging code), however, are actually a set of rules for a database--a bibliographic database. The Rules have evolved through a century and a half-long experience in cataloging, but their adequacy to computerized databases should be examined.

A study using both the Entity-Relationship Model and a generalized database approach to analyze the Rules showed that the current arrangement of the Rules and their individual compositions disagrees with general principles of database design. These principles, in turn, suggest an outline for an alternate organization of the Rules (Ref. 21). Further studies should be conducted to create a cataloging code that is structured for use in modern databases.

(d) The Librarian's Role. OPACs and gateway systems are likely to increase the use of catalogs and bibliographic databases by users. Moreover, as indicated by the CLR study, users expect more from the new online catalogs than from card catalogs. While we would like to think that online systems of the future will be friendly and helpful enough to be self-explanatory, much research and development is still needed before this ideal can become a material fact. Although current online systems provide increasingly friendly interfaces, users still do not know enough about the coverage of these systems and about successful search strategies.

For this reason, the role of the librarian in assisting users has increased: in addition to serving as mediators or consultants, the training of users to search online systems is now a major responsibility for librarians. These new roles, and the modifications introduced into existing roles, should be reviewed and examined in order to determine the performance of these roles and their effectiveness.

(e) User Training. Theoretically, user training can be performed either by computers or by librarians. In reality, most user training is performed by librarians. Although studies in individual libraries (and the CLR study) showed that some user training is needed, two basic questions require rigorous testing: (1) When is training needed? (2) Which methods are most effective for each user group?

These questions are not simple to answer, and ordinary comparative studies do not provide the necessary results, as has already been shown (Ref. 22). Additional exploratory research is needed before studies to answer these questions can be designed. For example, observation of users' searching behavior and an in-depth analysis of their thought processes (e.g., Ref. 23) will point to problems that might be remedied with training. Once these problems are identified, the contribution of training to their solution can be tested. Further, the effectiveness of each training method is likely to depend on the characteristics of individual users. While experience in online searching immediately comes to mind as an example of such a characteristic, observations of actual training sessions and the resultant searching behavior will suggest other factors that should be considered in testing instructional methods.

In conclusion, OPAC research has only began to uncover the issues that are involved and the approaches needed for fruitful research. The ever-growing use of OPACs generates an urgent need for intensive research that will lead to improvements in systems use and design. Further, the new direction of multi-faceted and integrated research will be beneficial to the retrieval of information from all bibliographic online systems.

6. REFERENCES

1. Ashoor, M.S.; Khurshid, Z. User Reactions to the Online Catalog at the University of Petroleum and Minerals Library. J. Acad. Libr., 1987, 13(4).
2. deBruijn, D.; Matheson, A. University of Calgary: Remote Access to On-Line Catalogues. Can. Libr. J., 1987, 44(4).
3. Lipetz, B.; Paulson, P.J. A Study of the Impact of Introducing an Online Subject Catalog at the New York State Library. Libr. Trends, 1987, 35(4).
4. Matthews, J.R.; Lawrence, G.S. Further Analysis of the CLR Online Catalog Project. Info. Technol. Libr., 1984, 3(4).
5. Matthews, J.R.; Lawrence, G.S.; Ferguson, D.K., eds. Using Online Catalogs: A Nationwide Survey. New York: Neal-Schuman, 1983.
6. Markey, K. Thus Spake the OPAC User. Info. Technol. Libr., 1983, 2(4).
7. Ferguson, D. Public Online Catalogs and Research Libraries. Stanford, CA: Research Libraries Group, 1982.
8. Tolle, J.E. Understanding Patron's Use of Online Catalogs: Transaction Log Analysis of the Search Method. In: Productivity in the Information Age, Proceedings of the 46th ASIS Annual Meeting, ed. R.F. Vondran et al. White Plains, NY: Knowledge Industry Publications, 1983.
9. Larson, R.R. Users Look at Online Catalogs: Results of a National Survey of Users and Non-Users of Online Public Access Catalogs. In: Interacting with Online Catalogs, Part 2. Berkeley, CA: University of California, Division of Library Automation and Library Research and Analysis Group, 1983.
10. Kinsella, J.; Bryant, P. Online Public Access Catalog Research in the United Kingdom: An Overview. Libr. Trends, 1987, 35(4).
11. Lancaster, F.W. Evaluation of Online Searching in MEDLARS (AIM-TWX) by Biomedical Practitioners. Urbana, IL: University of Illinois, Graduate School of Library Science, Occasional Paper No. 101, 1983.
12. Wanger, J.; Cuadra, C.A.; Fishburn, M. Impact of Online Retrieval Services: A Survey of Users, 1974-1975. Santa Monica, CA: Systems Development Corp., 1976.
13. Fenichel, C.H. Online Searching: Measures that Discriminate Among Users with Different Types of Experience. J. Amer. Soc. Info. Sci., 1981, 32(1).
14. Wanger, J.; McDonald, D.; Berger, M.C. Evaluation of the Online Process. Santa Monica, CA: Cuadra Associates, 1980.
15. Bellardo, T. An Investigation of Online Searcher Traits and their Relationship to Search Outcome. J. Amer. Soc. Info. Sci., 1985, 36(4).
16. Woelfl, N.N. Individual Differences in Online Search Behavior: The Effect of Learning Styles and Cognitive Abilities on Process and Outcome. Ph.D. Dissertation, Case Western University, Cleveland, OH, 1984.
17. Bates, M.J. How to Use Information Search Tactics Online. Online, 1987, 11(3).
18. Fidel, R. Moves in Online Searching. Online Rev., 1985, 9(1).
19. Sewell, W.; Teitelbaum, S. Observations of End-User Online Searching

Behavior Over Eleven Years. <u>J. Amer. Soc. Info. Sci.</u>, 1986, 37(4).

20. Fidel, R. Towards Expert Systems for the Selection of Search Keys. <u>J.</u>
<u>Amer. Soc. Info. Sci.</u>, 1986, 37(1).

21. Fidel, R.; Crandall, M. The AACR2 as a Design Schema for Bibliographic
Databases. <u>Libr. Quart.</u>, 1988, 58(2).

22. Nielsen, B.; Baker, B. Educating the Online Catalog User: A Model
Evaluation Study. <u>Libr. Trends</u>, 1987, 35(4).

23. Janosky, B. Online Library Catalog Systems: An Analysis of User Errors
<u>Internat. J. Man-Machine Stud.</u>, 1986, 25.