

# Online Public Access Catalog Research in the United Kingdom: An Overview

---

JANET KINSELLA  
PHILIP BRYANT

## Introduction

DESPITE THE MANY STATEMENTS that have been made about the rapid pace of technological developments in libraries—particularly advances in automation—it is surprising to reflect on how slow some of the developments have been.

The Centre for Catalogue Research (CCR) evolved from work that commenced at the end of the 1960s in the library at the University of Bath. Maurice Line, then the university librarian, had been offered the use of a terminal by the university's computer unit, and he wanted to experiment with the development of an online catalog which would use very brief entries. The catalog record as originally conceived by Line was extremely minimal—surnames only, short title, date, class number, and book number. The aim was for "direct access in the simplest possible way from all names or titles to the entry and its location, giving editors, etc. more or less equal status to authors so far as use of the catalogue is concerned."<sup>1</sup> As it happened, the terminal failed to materialize and the result was the development of the offline Bath Mini-Catalogue which used variable field records considerably shorter than those customarily provided in university library catalogs. Although the library was not, at the time, provided with its own terminal, Gillian Venner, who worked with the Computer Unit at Bath University from 1969-74 and who was recently a member of the Online Keyword Access

---

Janet Kinsella is Research Officer, Centre for Catalogue Research, The Library, University of Bath, Bath, England; and Philip Bryant is Director, Centre for Catalogue Research, The Library, University of Bath, Bath, England.

to Public Information (OKAPI) team at the Polytechnic of Central London, set up a small experimental online catalog in the unit. The system permitted retrieval just by book number or by personal author name. The appendix of the Bath Mini-Catalogue report that described this experiment concluded by saying: "A large online file must earn its keep!"

The subsequent development of online catalogs in the United Kingdom has indeed been slow. As recently as April 1986, Juliet Leeves, speaking at the CCR's second National Conference on "Online Public Access to Library Files"<sup>2</sup> held at the University of Bath, showed that commercially produced OPACs were only currently available in nineteen libraries (twelve from one supplier) and a number of these are public query facilities to circulation files. Very few of the systems yet offer all the features belonging to a true OPAC.

As Stephen Walker has pointed out,<sup>3</sup> OPACs are information retrieval systems not unlike online retrieval systems such as DIALOG. Their important feature is that they have to be designed *not* to require a human intermediary. The research and development work required to achieve this goal is considerable and, until recently, the resources and moral support have not been available in the United Kingdom to assure significant progress. CCR was pressing for research in this area at the end of the 1970s but little progress was made until 1982. That year the centre undertook a series of surveys on the use of public inquiry facilities to circulation files and, in 1983, organized a number of one-day seminars entitled "Introducing the Online Catalogue."<sup>4</sup> More significantly, in that year a start was made by the Polytechnic of Central London on the OKAPI project and the design of an OPAC to run on a local area network.<sup>5</sup>

### Recent Research

In the three years since 1983, there has been no shortage of research into OPACs. Work has fallen into two categories: (1) the study of existing systems; and (2) the design of experimental and prototype systems.

Many projects (and the majority of the work to be considered in this overview) have been sponsored by the British Library Research & Development Department; however, there has also been considerable research activity by individual institutions. Examples include the University of Hull, where papers have been produced by a special research group established within the library, and the Cranfield Institute of Technol-

## *Online Public Access Catalog Research*

ogy which has been engaged in a major replanning of its total library service and where preparing for online public access has involved evaluation of some commercial systems. In addition to institutions, individuals such as Keith Renwick at the University of Manchester Institute of Science and Technology (UMIST), who has conducted a study comparing the use of various forms of catalog, have made their particular contribution to this area of research.

How relevant has recent research been to the needs of libraries and users? Has it changed the design of commercial online systems? An apparent gap between research and its application was just one of a number of factors that led in 1985 to the formation by the British Library Research & Development Department of a working group of researchers already active in the field to discuss current research and to identify future priorities. Five main areas for research were identified:

1. Systems design—e.g., search sequencing and intelligent browsing.
2. Impact studies relating to user behavior—e.g., monitoring of system use and changes in user behavior.
3. Impact studies relating to organizational change—e.g., monitoring impact of OPACs on resource allocation, staffing, and library structure.
4. Visual browsing/ergonomics—e.g., screen content, layout, typography, use of color.
5. Bibliographic factors—e.g., database enhancement, bibliographic standards, full text searching, multiple database searching.

As a result of the group's meetings, a *Programme of Research* was formulated and published by the British Library Research & Development Department in January 1986.<sup>6</sup> In addition to the funding of current OPAC work, this report recommended the allocation of an additional £300,000 to OPAC research over the next three years. The program especially encouraged cooperative projects that involve working libraries cooperating with one or more research teams in the application of information retrieval techniques, evaluation, etc. Many of these projects will draw heavily on the techniques used in previous research especially in terms of monitoring and assessment. These techniques include: (1) feature analysis, (2) transaction log analysis, (3) questionnaire surveys, and (4) comparative studies. Such work has contributed greatly to our knowledge of OPACs, but the techniques do contain a number of inherent weaknesses.

*Feature Analysis*

This defines the nature of OPACs and aids the dissemination of information about systems to the library profession. For example, the recent survey of systems (including OPAC modules) in the United Kingdom by Juliet Leeves<sup>7</sup> fulfills the need for factual information about "the marketplace."

Last year CCR concentrated on a particular aspect of OPACs in a "browse-screen" questionnaire that looked specifically at the display and format of brief bibliographic records and index entries by OPACs. Analysis of responses received from forty-one selected United Kingdom, U.S. and European suppliers and libraries with in-house systems demonstrated the great diversity in current OPAC design. For example, the length of a brief entry varies between systems as does the number of entries displayed. Not only is there a lack of consensus between the use of single-line and multiline, but also between displaying fixed length or variable length brief records.<sup>8</sup>

However, feature analysis of this type is mainly a counting exercise; it rarely gives critical evaluation of the features enumerated. For example, the CCR study shows that just over half of United Kingdom systems surveyed (ten out of nineteen) would respond to all or some types of inquiry resulting in zero matches with a message (e.g., "no books found"). How helpful is this to the user? Surely a display of brief bibliographic records, or index entries, including the closest match would be of greater use.

*Computer Logging and Transactional Tape Analysis*

Considerable use has been made of this technique in the United States. Well-known examples of such research include the work by John Tolle of the OCLC Office of Research for the Council on Library Resources<sup>9</sup> and by Christine Borgman of the University of California at Los Angeles whose research was supported by OCLC.<sup>10</sup> In the United Kingdom examples of its use include work at Sussex University<sup>11</sup> and by the OKAPI team at the Polytechnic of Central London.<sup>12</sup>

Transaction logs enable researchers to quantify readers' exploitation of the facilities provided by a particular OPAC. Analysis of transaction logs at Sussex University Library demonstrated the changing pattern of use over two years, with the gradual increase in popularity of the "quick" search (i.e., 4,4 derived search key). Readers' ability to use systems should not be overestimated. Logs have indicated that users have difficulty with such tasks as reading instructions, responding to prompts, and keyboarding. Analysis of ninety-six OKAPI sessions identified spelling errors in about 10 percent of user input. One major

failing of transaction log analysis is that it is difficult to examine the patterns of use by individual users. Methods of delineating between sessions are not accurate and methods of obtaining information on the identity of the users (e.g., by asking for a user number), raise the question of confidentiality. A university community is not a homogeneous body in terms of ability. Even though users bring to the terminal differing amounts of experience in library, catalog, and computer use, transaction log analysis cannot identify characteristics of individuals. Users are restricted by the facilities provided by the particular system in use. They may require features, such as the facility to limit or broaden a search, that are not provided or are inadequately presented on the screen. Thus the results of a transaction log analysis may not reveal the real needs of users.

### *Surveys*

Researchers in the United States have conducted large-scale surveys.<sup>13</sup> On a modest scale, the Centre for Catalogue Research has carried out surveys at the libraries of Hull University and the Polytechnic of the South Bank where users had access to Geac online inquiry modules and at the Polytechnic of Central London and Bath University libraries where access was provided to certain elements of the SWALCAP circulation system. Over 800 questionnaires were collected giving information on users, their inquiries, and their views. However, surveys leave many questions unanswered. Would users have found items more quickly through another form of catalog? What were users' true understanding of the system? What errors did they make?

Despite the limitations of surveys, they are still important tools for the study of users' reactions and user characteristics. Surveys are the main data collection instruments for CCR's impact studies which are currently attempting to monitor changes in library use after the installation of an OPAC. Such "before and after" studies are taking place in four libraries (Coventry Polytechnic, Leicester Polytechnic, Devon Public Library System, and the Lancashire Library System) and cover a range of commercial integrated systems offering online public access (CLSI, BLCMP, DS, and Geac respectively). A questionnaire survey of library users (including nonusers of the catalog) is to be supplemented with brief semistructured interviews and observation of catalog use at selected periods over several days at each of the participating libraries.

### *Comparative Studies*

Comparative studies have been, and are being, conducted using experimental methods. For example, in a series of small controlled

experiments organized by CCR and carried out by Linda Reynolds at the Polytechnic of the South Bank and by Hans-Ove Frid<sup>14</sup> at Bath University, the use of the public inquiry facility provided by Geac and the SWALCAP Library Cooperative respectively was compared with the use of parallel files on COM microfiche. Users' performance was measured in terms of speed of, and success in, searching common and uncommon personal names correctly or incorrectly cited. When users were given correct citations they were able to retrieve the relevant titles faster in the online file; however, these studies identified possible sources of difficulty in online browsing. The user was less successful at retrieving titles online for which an incorrect personal name was cited.

### Current Research

Studies based on existing systems have a number of inherent limitations. The control of variables is particularly difficult, and therefore results are not usually generalizable. Working with a commercial system within the constraints of a "live" library, it is often not possible to vary one feature of the system while keeping others constant. By designing experimental and prototype systems, researchers have attempted to solve such problems.

CCR for example has developed an experimental system for the study of interface design based on the information retrieval facilities of BRS/Search. The system, currently mounted on an Onyx 16-bit microcomputer, retrieves records from a specially constructed 3000 record database on education-related topics taken from Bath University Library's SWALCAP file. The user interface has been developed using the high-level programming language MENACE. The basic screen design consists of a three-part display showing system information, bibliographic information (in the form of full or brief bibliographic entries), and commands. "Help" information can be displayed by partially "wiping" the screen, allowing the user to view records, "help," and relevant commands simultaneously. The flexibility of the system permits the design and testing of a variety of experimental interfaces.<sup>15</sup>

Research at the Polytechnic of Central London has been based on the design of a prototype online catalog (OKAPI), developed for a microcomputer network allowing simultaneous access for up to twenty users. Great emphasis has been placed on making the catalog as user friendly as possible both in terms of system design and ergonomics. Users' reactions to existing OPACs formed the basis for the design of the OKAPI user interface. The traditional distinction between "known item" and "subject" searches has been maintained, but these searches

## *Online Public Access Catalog Research*

are expressed by options to look for "a specific book" or "a book about something." Command input has been minimized to a single keystroke and brightly colored function keys have been incorporated into the keyboard for common commands (e.g., green to "continue" and yellow for "help").<sup>16</sup>

Unlike earlier work which concentrated on the passive task of monitoring existing systems, experimental research has begun to be recognized as having an active role to play in the improvement of commercially produced OPACs. The success of such OPACs depends on three factors: (1) technical performance (both of systems and communications technology), (2) interface design, and (3) quality of the database. This multifaceted approach to OPAC design is clearly demonstrated in current research activity concerned with subject searching.

### *Technical Performance*

One approach to improving subject access is by incorporating advanced information retrieval techniques into systems—e.g., automatic word stemming, spelling correction, and synonym generation. Analysis of OKAPI search sessions proved useful in identifying the potential of automatic word stemming. An analysis of 119 consecutive search statements, forming around seventy-two discrete OKAPI searches, indicated that a simple stemming procedure, conflating singular and plural noun forms and the verbal endings "ing," "ed," and "s," would have improved the performance of 23 percent of the statements. The work, being carried out by Stephen Walker, is based on Porter's stemming algorithm and involves the construction of a synonym dictionary and incorporation of stemming/synonym procedures into the OKAPI search program to enable all subject search words and personal names within specific item searches to be processed. The effectiveness of such techniques is to be evaluated through the analysis of user searches to be carried out through both the control and improved OPAC.<sup>17</sup>

Other basic information retrieval research relevant to OPACs is being carried out in the field of relevance feedback. Examples include the work by Niall Teskey at the Department of Computing and Cybernetics, Brighton Polytechnic on the evaluation of various methods of ranking and by Stephen Robertson of the Department of Information Science, City University on the development of a front-end to Medline, providing weighting, ranking, and relevance feedback.

One approach to the application of relevance feedback to OPACs is being investigated by Stephen Walker at the Polytechnic of Central London. The research aims to improve subject searching by combining the use of subject information embedded within a classification scheme

(e.g., Dewey Decimal Classification) with relevance feedback. Having asked the user to specify the relevance of records retrieved from a free language search request, the proposed system (based on OKAPI) will assign relevance weights to the user's search terms and to the classification numbers of relevant records. The system will then automatically perform another search, presenting the user with related records.

The traditional approach to user-initiated subject access by information retrieval systems has been through the use of Boolean operators. However, many OPACs restrict their explicit use, offering only the implicit use of "AND" between terms. The OKAPI team rejected this simple post-coordinate approach to subject searching because of its lack of precision and have incorporated a complex hyper-Boolean facility based on an algorithm developed by D.J. Harper. Operating on the keywords found in the user's subject search string, the system implicitly puts terms in an OR relationship, calculating a weight for each posting based on the frequency of the terms within the index.<sup>18</sup>

### *Interface Design*

Research into the use of Boolean operators can also be approached from an ergonomic viewpoint. The CCR is studying alternative methods of presenting operators. For example, the experimental system includes a number of interfaces in which natural language explanations have replaced traditional terms (e.g., SELECT, INCLUDE, EXCLUDE, for AND, OR, NOT).

While the importance of research into "intelligent browsing" (e.g., relevance feedback, weighting, and Boolean search modification) should not be underestimated, "visual browsing" (i.e., the examination of data by the eye) also has a significant role in enabling greater ease of use. The CCR is undertaking a series of user tests, studying the effect of brief bibliographic entry length on "visual browsing." Tests take the form of controlled experiments using the centre's experimental system. For example, students are asked to identify particular titles within alternative "browse screen" displays. Initial indications from a pilot test show that single line entries are preferred by users and are quicker "to browse."

### *Quality of the Database*

The introduction of OPACs and other forms of online access to bibliographic data has renewed users' interest in the subject approach to information. Many of the comments gathered from the CCR questionnaire surveys suggested that users would like to see the expansion of individual entries to include notes, abstracts, and contents pages.

## *Online Public Access Catalog Research*

Enhancement of the database to improve subject access may sound attractive but first data currently included in MARC records should be fully utilized. United Kingdom MARC records contain PRECIS (Preserved Context Indexing System) indexing strings. Their possible use in providing structured subject access to an OPAC is being investigated by Juliet Congreve at Middlesex Polytechnic. As with the Bath experimental catalog, work at Middlesex Polytechnic is based upon the information retrieval capabilities of an existing software package, in this case STATUS. The system has been designed to allow users subject access without their needing to know the structure of the indexing language. Terms are matched against a thesaurus which is held as two indexed sequential files, one arranged in alphabetical order of index terms, the other sequenced by the PRECIS Reference Index Number (RIN). The RIN file, by acting as an automated subject authority file, facilitates cross referencing to related records. The system translates the verbal form of a match into an RIN number which, when compared against the RIN file, identifies related records.<sup>19</sup>

### **Future Research**

The OPAC Research Programme is already attracting a number of proposals for the British Library's consideration. Scarcity of resources in the United Kingdom means that it is essential that there is coordination of research effort and also that there is proper awareness of projects overseas—especially in North America and Europe. There has been an effort within the OPAC Research Programme to establish closer relations between OPAC designers and basic researchers in the field of information retrieval. There is, however, a need to widen the range of disciplines involved—psychology, ergonomics, graphic design, and so on. Unfortunately much of the research is limited by existing technology (e.g., interface design research is conducted with the standard eighty character by twenty-four line VDU screen) and becomes outdated as the technology changes. It is hoped that a greater emphasis on long-term, basic research could lead to results of general application. It is interesting to note that long-term work has been encouraged in Sweden. LIB-LAB, established at the University of Linköping and directed by Roland Hjerpe, draws on the experience and the talents of a range of disciplines (as recommended earlier) to develop a research laboratory in library and information science.<sup>20</sup>

### **Conclusion**

Frederick Kilgour has stated that OPACs will “profoundly change the way people go about the business of living.”<sup>21</sup> This is arguable.

Certainly there is a communications revolution which will effect a change, but OPACs are strictly concerned with providing access to bibliographic data in libraries and library-related systems. They are being developed in some quarters as public inquiry systems for the exploitation of a whole range of sources of data—e.g., community information, viewdata services, external databases, etc. It would appear that at present there is more than enough research to be undertaken in connection with the more conventional functions of an OPAC if the return on the investment in library collections is to be realized.

### References

1. Bryant, Philip, et al. *The Bath Mini-Catalogue: A Progress Report*. Bath, England: Bath University Library, 1972.
2. Leeves, Juliet. "OPACS Off the Shelf." In *Online Public Access to Library Files: Second National Conference*, edited by Janet Kinsella, pp. 35-48. Oxford, England: Elsevier International Bulletins, 1986.
3. Walker, Stephen. "The Free Language Approach to Online Catalogues: The System." In *Keyword Catalogues and the Free Language Approach*, edited by Philip Bryant, pp. 31-40. Bath, England: Bath University Library, 1985.
4. Seal, Alan, ed. *Introducing the Online Catalogue*. Bath, England: Bath University Library, 1984.
5. Mitev, Nathalie N., et al. *Designing an Online Public Access Catalogue* (Library and Information Research Report 39). London: British Library, 1985, pp. 208-10.
6. British Library. *Online Public Access Catalogues: Programme of Research*. London: British Library, 1986.
7. Leeves, Juliet. *Library Systems: A Buyer's Guide*. Aldershot, England: Gower Publishing Group Ltd., 1987.
8. Bath University Library. *Online Public Access to Library Files: A Progress Report*. Bath, England: Bath University Library, to be published in 1987.
9. Tolle, John. *Current Utilization of Online Catalogs: Transaction Log Analysis* (final report to the Council on Library Resources), vol. 1, series no. OCLC/OPR/RR-82/2. Dublin, Ohio: OCLC, 1983.
10. Borgman, Christine. *End User Behavior on the Ohio State University Libraries' Online Catalog: A Computer Monitoring Study* (OCLC Research Report Series no. OCLC/OPR/RR-83/7). Dublin, Ohio: OCLC, 1983.
11. Young, Robert, et al. "Geac with Local Enhancements: The Integrated Real-time System at the University of Sussex Library." *Program* 20(Jan. 1986):1-26.
12. Mitev, et al., *Designing an Online Public Access Catalogue*.
13. Matthews, Joseph, et al., eds. *Using Online Catalogs: A Nationwide Survey*. New York: Neal-Schuman, 1983.
14. Frid, Hans-Ove. *Speed and Accuracy of Searching a COM Catalogue Compared to Using the Public Access Facility of an Online Circulation File*. Stockholm, Sweden: Stockholm University Library, 1985.
15. Prowse, Stephen G. "Use of BRS/Search in OPAC Experiments." *Program* 2(April 1986):178-95.
16. Mitev, et al., *Designing an Online Public Access Catalogue*.
17. Jones, Richard. "Improving OKAPI: Transaction Log Analysis of Failed Searches in an Online Catalogue." *Vine* 62(1986):3-13.

## *Online Public Access Catalog Research*

18. Mitev, et al., *Designing an Online Public Access Catalogue*, pp. 90-96.
19. Congreve, Juliet. "Browsing through PRECIS: Structured Subject Access in an Online Catalogue." In *Online Public Access to Library Files*.
20. Hjerppe, Roland. "The HyperCAT project at Linköping University." In *Online Public Access to Library Files*.
21. Kilgour, Frederick G. "The Online Catalogue Revolution." *Library Journal* 109(15 Feb. 1984):320.

This Page Intentionally Left Blank