Subject Searching in an Online Catalog with Authority Control

Noelle Van Pulis and Lorene E. Ludy

Results of previous catalog research indicate that users want improvements in subject searching features, noting display of terms related to their topics as especially desirable. This study reports the findings of an examination of subject searchers' use of such a feature provided through online display of authority information from Library of Congress Subject Headings for terms used in the catalog. It was found that, of the terms entered by users, 78 percent were an exact or close match to catalog terms, and another 14 percent matched "see" references. Although additional information was available for nearly half of the searches performed, it was used in only 28 percent of these cases. It is suggested that improvements in subject searching features are needed to encourage use of available syndetic structures.

This study examined subject searching in a large research library's online public access catalog with a controlled vocabulary. The main focus was on the system's response to the user's search. It was hoped that the findings would shed light on how to improve subject searching within the existing online system and standard catalog records.

In particular, the authors sought answers to a number of related questions, raised by their experience with the online system and by earlier studies of catalog use:

What kind of terms do users enter in subject searching (e.g., broad, narrow, subdivided)?

How often does a user's search term match the terminology of the catalog?

How does the catalog respond (term found, not found, cross-referenced)?

How do references affect this response?

How do users interpret the response and how do they proceed based on it?

How often is a syndetic structure available?

How often is it used?

What are the users' perceptions of the value of syndetic structure?

Do online subject searchers view the card catalog as an alternative or supplemental source for subject searching?

What is the relationship of subject search activity in the online catalog and use of a closed card catalog?

BACKGROUND

In 1980, Neal Kaske and Nancy Sanders reported on an OCLC-sponsored research project on subject searching in traditional catalogs. Users indicated that "[t]hey wanted an easy, transparent translation from the terms they entered to the terms used by the catalog." The results of the national study of online public access catalogs in 1982 sponsored by the Council on Library Resources revealed that subject searching is very much wanted and needed but can be difficult or problematic.

Noelle Van Pulis is Systems Librarian, and Lorene E. Ludy is former Catalog Planning Librarian at the Ohio State University Libraries, Columbus, Ohio 43210.
for the user. Many of those surveyed, including users at the Ohio State University Libraries (OSUL), reported difficulty finding subject terms. When asked to identify features that they would want added to the online catalog, respondents ranked an online display of related subject headings as their first choice. Traditional card catalogs display related subject headings in two ways. First, the alphabetic arrangement of the card file brings together related terms. Secondly, most card catalogs include syndetic structures, defined as "the array of devices used to bring to the user’s attention relations among terms that are not located near each other in the index." In a subject catalog these devices are typically see references, see also references, and, less commonly, scope notes.

The availability of syndetic structures in online catalogs has been limited, although more systems are beginning to add these features. The addition of syndetic structure to Ohio State’s online catalog in 1984 provided an opportunity to examine how users view and use this feature.

THE ONLINE CATALOG

The online catalog at the Ohio State University Libraries evolved from the online circulation system, Library Control System (LCS), which was implemented in 1970 and used brief records converted from the shelflist. During subsequent years, many revisions and enhancements were made to allow LCS to function as the OSU libraries’ online catalog. Two of the most important enhancements were the implementation in 1978 of the capability to store and display a full catalog record and, in 1983, the creation of a headings file for authority control. With the addition of these features, LCS met the traditional definition of a library catalog.

The establishment of the headings file allowed for inclusion of cross-references for names, subjects, series, and uniform titles. The process of adding references for subject headings was facilitated by the availability of Library of Congress Subject Headings (LCSH) in machine-readable form. In September 1984, the OSUL tape-loaded LCSH information for those subject headings used by the libraries.

LCS now serves as the online catalog for the OSUL and the State Library of Ohio and also includes records for portions of the collections of other libraries (e.g., OSU law library and the Center for Research Libraries). However, full catalog records are available in LCS for fewer than half of all titles held by the OSUL, because almost all titles cataloged before 1972 have only brief circulation records. Thus, the subject search capability in LCS applies only to State Library of Ohio titles, a portion of OSUL titles, and some titles from other libraries.

SUBJECT SEARCH AND DISPLAY IN LCS

The LCS subject search requires the user to enter the command "SUB/" followed by any word or phrase, preferably a valid LCSH. The resulting display, shown in figure 1, is an alphabetic list of subject headings, called the LCS subject index display. The LCS response shows the user’s entry, as keyed, in an alphabetic sequence of subject headings.

If the user’s term does not exactly or partially match a term in the catalog, LCS displays the message "Nothing was found under:" in front of the user’s entry. The user sees the entered term in the alphabetical list and can check for spelling or keying errors, or select an alternate heading from the list. The index display includes see references, labeled on LCS with "Search under:" (see figure 2).

Prompts at the bottom of the display suggest choices for the next step in the user’s search. The command "TBL/" displays the list of titles assigned the subject heading displayed on the line. The number of titles available is indicated by the posting in the "items" column. Titles are displayed in reverse chronological order by date of publication.

An alternative second step is the "SAL/" command, which has the meaning "search also, by line." With this command, LCS displays the syndetic structure of a subject heading, as taken from LCSH: classification numbers(s), notes, and see also references, labeled "search also under" (see figure 3). Excluded from
this related headings display are see from and see also from (x and xx) references found in the printed LCSH. (A staff version of the same display includes the latter information, as well as headings control numbers and some additional authority control information. The headings file was designed to mask this information in the public display.)

Another alternative for the user is to page through the subject index. Function keys on library terminals enable easy access to preceding and following screens. This allows users to browse through the complete alphabetical list of subject terms used in LCS.

**STUDIES OF SUBJECT SEARCHING**

Several studies have touched upon subject searching at Ohio State, all predating the addition of cross-references to the online catalog. Carol Weiss Moore reported that the overall level of subject searching (versus searches by author and title) was 30 percent among four libraries surveyed. One of these was LCS at Ohio State, which lacked a true subject search capability at the time of the study. User-reported success rates for subject searches ranged from 70 to 75 percent. In a study of the terminal command usage at OSUL during the late 1970s, David Norden and Gail Herndon Lawrence found that subject search "apparently has met an important need of LCS users despite the present lack of online cross-references and authority control." The study did not attempt to measure success rates. The results of the 1982 survey of LCS searching by Sammy Alzofon and Noelle Van Pulis reconfirmed...
METHODOLOGY

The study consisted of a questionnaire distributed at OSUL’s main building between 9 a.m. and 5 p.m. during a period of four days in May 1985. Previous studies based on use of LCS indicated that users of the main library are representative of the libraries’ user population. The building contains primarily a humanities and social sciences research collection, is centrally located, has a study lounge, is one of the few libraries on campus with large reading areas, and has the largest number of LCS public user terminals in one location. Thus, in addition to serving graduate students and faculty, the main library is a popular location among undergraduate students despite the presence of an undergraduate library elsewhere on campus. Graduate students and faculty in the sciences are served primarily by department libraries outside the main library.

Users approaching an LCS terminal were asked if they would be searching by subject and, if so, would be willing to complete a brief questionnaire as they performed their search(es). Most users expressed willingness, and 248 forms were distributed. The questionnaire focused on the use of the subject search (“SUB/”) and its related headings search (“SAL/”). Other searches that have been intended as a subject approach (e.g., title searches on topical terms) were not considered. Some questions were asked about use of the LCSH volumes prior to beginning an online search and use of the card catalog (closed in 1982) in relation to the subject search.

The searches reported by users were later replicated in LCS by the authors to verify reported system responses. Of the questionnaires, 203 were judged usable for analysis. In addition, data available from the subject searches analyzed in the 1982 Alzofon and Van Pulis study were examined to determine if the success of searches performed prior to the addition of cross-references would have been improved by the addition of references.

RESULTS AND DISCUSSION

User Characteristics

Figure 4 summarizes the characteristics of the participants, most of whom were undergraduate students (77 percent) and male (55 percent). There were five faculty members, and five participants did not indicate their standing. These ten participants together with four persons who did not indicate their sex, are the “other” category in figures 6-8.

With respect to users’ previous experience with LCS, nearly half of the participants (49 percent) in this study reported using LCS a few times each week or month, and another 5 percent said they used LCS every day. Only 5 percent said they had never used LCS before. Most participants (93 percent) said they had done at least one subject search prior to the search recorded in this study, and more than half (54 percent) reported ten or more previous subject searches. Clearly,
Subject Searching in an Online Catalog

many participants were frequent users of LCS and most had experience using it for subject searches.

All entering freshmen and transfer undergraduate students at Ohio State are required to take a University College course (UVC) that includes a library instruction component. Later, students may also receive specific course-related library instruction. However, only 35 percent of the participants in this study reported learning to use LCS as a result of such instruction. Many participants (47 percent) reported that the brief LCS instructional brochure was their source of knowledge about online catalog. A small number of participants (3 percent) indicated that they had learned to use LCS in a workshop offered by the libraries. These figures are not mutually exclusive, as a number of participants checked multiple responses. These findings are similar to those reported in the CLR study, where users reported that printed aids were consulted most often as type of assistance during their most recent online search.

Search Terms

The first question on the survey asked participants to write down the actual word or words they used to search the catalog, rather than an expression of their topic. Most (80 percent) of the 203 responses were single-concept terms or names. A description of these responses is given in figure 5. The tendency to use brief terms or names was noted by Karen Markey in her study of subject searching in card catalogs.

Only fifteen participants (7 percent) entered a subdivided term, such as Women—Language, German Language—Dictionaries, Scotland—History, which suggests an understanding of the construction of subject headings. One of the combinations reported, however, was not constructed according to LCSH practice (Ethics—Abortion). Few user access points were constructed as expressions of a topic rather than single word or brief-phrase subject headings (e.g., Wine Regions of France, Socialization of American Children). No relationship was observed between the user's previous experience conducting subject searches on LCS and the appropriateness of entry vocabulary. Only two of the fifteen first-time subject searches had correctly formulated headings (Taxation and Welfare and Health Diseases (Homosexuality)).

Search Results

Searchers were asked to indicate the initial LCS response by selecting an option: Showed the topic I was looking for, Nothing was found under what I typed.
Table:

<table>
<thead>
<tr>
<th>Users’ Access Points</th>
<th>N = 203</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-word noun (e.g., weather, philosophy, linguistics)</td>
<td>74</td>
<td>36</td>
</tr>
<tr>
<td>Phrase (single concept) (e.g., Zen Buddhism, human anatomy, women athletes)</td>
<td>54</td>
<td>27</td>
</tr>
<tr>
<td>Personal name (e.g., Lenin, Mario Cuomo)</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>Corporate name (e.g., OSU, LEAA)</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Unsubdivided geographical name</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Single adjective (e.g., German, Mayan)</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Heading formulated with subdivision (e.g., Scotland-History, Women-Language, Art-Roman)</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Multiple concept (e.g., Wine regions in France, Television and children, Taxation and welfare)</td>
<td>18</td>
<td>9</td>
</tr>
</tbody>
</table>

Figure 5

Suggested other words to use
Other: ________. Figure 6 summarizes the responses to this question.

Most participants (74 percent) reported that LCS showed the topic for which they were searching. Since the LCS display after a subject search is an alphabetic index of subject headings, this response could mean that the entered term matched a subject heading or so approximated a subject heading that a useful heading appeared on the screen. Nine percent of the participants reported that LCS suggested other words to use, that is, provided a cross reference. Sixteen percent of the participants said that the first LCS response was “nothing was found” under the words they typed. Most (61 percent) of these searchers tried another subject search; 23 percent investigated other headings on the screen by doing a “TBL/” or “SAL/” on a displayed heading; 13 percent paged through the index. Only one searcher (3 percent) quit at this point.

The authors replicated these 203 initial searches. The results are shown in figure 7. Fifty-three percent of the entered terms showed no match on the LCS display; 12 percent were found and resulted in a successful search; 7 percent were found and provided a cross reference; and 2 percent were found but resulted in an unsuccessful search. Some respondents checked more than one answer. Percentages have been rounded to the nearest whole number.

Table:

<table>
<thead>
<tr>
<th>User</th>
<th>N = 203</th>
<th>Showed Topic</th>
<th>Nothing Found</th>
<th>Other Words</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male Undergraduate</td>
<td>96</td>
<td>73%</td>
<td>13%</td>
<td>7%</td>
<td>3%</td>
</tr>
<tr>
<td>Female Undergraduate</td>
<td>60</td>
<td>82%</td>
<td>12%</td>
<td>7%</td>
<td>0</td>
</tr>
<tr>
<td>Undergraduate</td>
<td>156</td>
<td>76%</td>
<td>15%</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>Male Graduate</td>
<td>16</td>
<td>69%</td>
<td>19%</td>
<td>13%</td>
<td>0</td>
</tr>
<tr>
<td>Female Graduate</td>
<td>17</td>
<td>59%</td>
<td>24%</td>
<td>24%</td>
<td>0</td>
</tr>
<tr>
<td>Graduate</td>
<td>33</td>
<td>64%</td>
<td>21%</td>
<td>15%</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>14</td>
<td>79%</td>
<td>14%</td>
<td>7%</td>
<td>0</td>
</tr>
<tr>
<td>All</td>
<td>203</td>
<td>74%</td>
<td>16%</td>
<td>9%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Some respondents checked more than one answer. Percentages have been rounded to the nearest whole number.

Figure 6

Reported Search Results
matched subject headings used in LCS. Another 15 percent were close to existing subject headings, so that the display included the heading sought (e.g., Deviance entered, Deviant Behavior is the correct heading). This confirms the utility of the index display. The authors found that 14 percent of the entered terms matched see references on LCS, and 18 percent of the terms were not found on LCS. Thus, user terms searched in LCS matched similar catalog terms for 81 percent of the searches. This result reaffirms Markey’s finding that users’ search terms frequently match standardized vocabulary in large catalogs.13

There are obvious discrepancies between the authors’ results and those reported by the participants. The participants’ higher percentage of “showed topic I was looking for” means they may have checked that category when they encountered references, because the first line of the reference is, in fact, the entered term. Seven of the forty-one persons who found see references answered “showed the topic I was looking for.” The discrepancy between the “nothing found” results reported by the users and those of the authors probably reflects a difference in interpretation (e.g., the authors counted Film as not found because there is no heading Film, although headings begin with that word; the user in this case marked “showed topic I was looking for”). In a few instances, there was no obvious explanation why the users thought their topics were found.

The fact that 14 percent of catalog users’ search terms matched cross-references shows how well cross-references direct users to appropriate terminology. This result is close to Markey’s findings at Syracuse University, where 16 percent of the searchers studied entered a valid cross-reference.14 A similar finding resulted from a reexamination of data from the 1982 Alzofon-Van Pulis study. In that study, the data were collected before subject heading references had been added to LCS. In seventy-six subject searches the respondents indicated they did not find something on the topic. When these searches were repeated, 18 percent (fourteen search terms) matched references now in the catalog. For these searches (6 percent of the total of 244 subject searches studied), the probability of success has been greatly increased.

Even though 74 percent of the participants in the current study reported a match on their searched term, fewer than half (48 percent) of these looked at titles at this point. Nearly a quarter (24 percent) paged though the subject index. These users had entered unsubdivided terms and, on seeing the index display the entered term with subdivisions, paged through to browse the more specific (subdivided) headings. For some 90 percent of the matched headings there were additional headings with further subdivisions on LCS.

Availability and Use of the Syndetic Structure

Users reported 230 access points (203 initial subject searches and 27 listed in response to questions three and five). Of these headings, 109 (47 percent of the total number of headings and 58 percent of the matched headings) had syndetic features associated with them (see also references, scope notes, class numbers). Questions three and four revealed how often this information was retrieved by the user.
Of the 30 displays of syndetic structures, nineteen provided a classification number for potential browsing; twenty-four provided see also ("Search also under") references; and two included a scope note.

Thirty retrievals (uses of the "SAL/" command) were recorded. Four users repeated the command, so twenty-six users (13 percent) retrieved the additional information. Thus, additional information was available for 47 percent of the searches, but it was viewed in only 28 percent of these cases.

Of the thirty displays of syndetic structures, nineteen provided a classification number for potential browsing; twenty-four provided see also ("Search also under") references; and two included a scope note. Asked what information they used, 18 percent of the users responding to this question reported using the call number in a shelf-position search ("browsing by call number"); 35.3 percent reported that they read the information, and the same percentage indicated that they searched the related terms. Users provided multiple responses to this question.

The syndetic structure was not used in the majority of the cases where it was available. This behavior may reflect searching habits developed earlier. If users are unfamiliar with online aids, they may be less likely to recognize their potential value. Also, the LCS subject index display did not indicate clearly the nature of the "SAL/" information. The display and instructional prompts were revised in September 1986 to facilitate use and encourage searching of "SAL/" information.

Question ten asked users' opinion of the "SAL/" search. Slightly more than half (55 percent) had never used it. Of the participants who reported having seen this display, 81 percent marked that the "SAL/" search usually provides helpful information, and 19 percent thought it was not useful. Although only 28 percent of the participants used the "SAL/," 42 percent had used it at some earlier time.

Failure to pursue related headings information has been seen in previous studies of card catalog users. In one case reported by Markey, none of 189 subject searchers consulted the Library of Congress Subject Headings in the course of a subject search.6 In this study, of the participants who responded to question five concerning use of LCSH prior to beginning their subject search, only 11 percent responded affirmatively. This was surprising in view of the libraries' user education program, which emphasizes use of the LCSH (and has done so since 1977), and the wide availability of printed aids in public areas.

This behavior may reflect a tendency to be satisfied with an initial positive response in searching. If, as in this study, 74 percent of the searches result in an exact or close match and another 9 percent result in a cross-reference, then the majority of searches find a subject on the first attempt. Users may believe that they have no reason to employ more sophisticated, exhaustive searching patterns. Pauline Cochrane, commenting on users' impatience at the catalog, suggested that a "more important feature in online catalogs than Boolean operations or proximity searching may be subject search prompts and aids."15 Carefully constructed, directive prompts and aids may be needed to encourage pursuit of potentially valuable information in the catalog.

Use of the Card Catalog

Related to online searching behavior is the supplementary use of the card catalogs. About 71 percent of titles at OSUL at the time of this study did not have full catalog records in LCS and so lacked online access by subject heading. Over half of the users surveyed were not aware of this fact. Again, the libraries' user education program includes instruction on the importance of the card catalogs for subject searches of OSUL materials cataloged before 1972. More than half of the participants in this study said that they did not intend to search the card catalog. They-
Subject Searching in an Online Catalog

<table>
<thead>
<tr>
<th>Reason</th>
<th>No need</th>
<th>No time</th>
<th>Dislike</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduate</td>
<td>N = 89</td>
<td>43%</td>
<td>43%</td>
<td>21%</td>
</tr>
<tr>
<td>Graduate</td>
<td>N = 23</td>
<td>65%</td>
<td>22%</td>
<td>17%</td>
</tr>
<tr>
<td>Other</td>
<td>N = 9</td>
<td>56%</td>
<td>33%</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>N = 121</td>
<td>48%</td>
<td>38%</td>
<td>19%</td>
</tr>
</tbody>
</table>

Some respondents checked more than one option.

FIGURE 8
Supplementary Use of the Card Catalog

ffered a variety of reasons, as shown in figure 8. The most frequent reason, chosen by 48 percent of the searchers, was that they did not need older materials. It could be that this reflects the predominance in the sample of undergraduate students, whose research needs may be limited. However, this reason was selected by more than half of the graduate students and fewer than half of the undergraduates.

The next most frequently chosen response (38 percent) indicated that lack of time was the reason for not pursuing a topic in the card catalog. This reason was more often given by undergraduates (43 percent) than graduate students (22 percent). It is not known if these users perceive the card catalog as time-consuming to search by its nature, or if they simply do not wish to spend their time checking an additional source. Another choice, “don’t like to use the card catalog,” was selected by 19 percent of the respondents, with a lower percentage of graduated students (17 percent) than undergraduates (21 percent) indicating this reason.

Question eleven asked if the searcher considered looking elsewhere, such as periodical indexes, for information on the subject. Most respondents (80 percent) said that they had considered other sources. Although it is not known if users actually used these sources, it is impressive that such a large number of users, predominantly undergraduates, said that they would not rely solely on the library catalog for access to information. These users do not seem to be as catalog-dependent as Evan Farber fears online catalog users may become.17

User Satisfaction

Although this study did not directly address issues of search success, user satisfaction with results, and vocabulary failure, it did solicit users’ evaluation of their search results, both content and presentation. Fifty-nine percent of the users were satisfied with the titles found; 26 percent expected more; 10 percent did not find their topic; and 5 percent did not answer this question. The twenty search terms listed by those answering “I could not find my subject listed in LCS” were examined by the authors. Half of the terms entered by these persons were in fact on LCS. Most are, however, general terms (e.g., “Police, Italy”) and apparently do not reflect the topics actually sought. Because the survey asked users only about access terms chosen and not another expression of topic, the specific interest of the user who searched “Police” or “Italy” cannot be known. This suggests an important area for further study: How do users refine their searching in light of seeing their original search term in an alphabetic list (often with further subdivisions) and the availability of see also references and scope notes?

Only 27 percent of those surveyed responded to the question on clarity of instructions in the screen displays. Of these, 71 percent indicated that the instructions

“This study confirmed the findings of other studies that users who conduct subject searches tend to enter single, often broad terms.”
were clear and easy to follow. However, LCS screen displays and prompts have been revised since the study as part of continuing refinement and improvement of the online catalog.

CONCLUSION

This study confirmed the findings of other studies that users who conduct subject searches tend to enter single, often broad terms. They found subject search terms in the online catalog about 80 percent of the time, and 14 percent of them were LCSH see references. The inclusion of these references in the catalog does clearly lead some searchers to terminology used in the catalog. Most of the terms searched were also in the catalog with further subdivisions. About half of the terms displayed a syndetic structure, alerting users to related material. Although this structure was examined for only 27 percent of the headings for which it was available, 42 percent of the users surveyed had examined the structure at some point in their previous use of LCS, and most of them considered the additional information useful.

Ohio State is continuing development of its online catalog subject search features, including assessment of content and display formats. At the same time that programming priority is being given to improvements in LCS authority control, attention is being focused on the availability and wording of prompts and other online aids to assist users in making more effective use of available information.

Ohio State is considering the feasibility of introducing keyword search capabilities in LCS. For example, the user who entered Ethics—Abortion might have found the topic if keyword access has been available for the correct heading. Abortion—Moral and Ethical Aspects (but only if truncation also was available and used on the term Ethic). If the user has entered either Ethics or Abortion (as single topics), there would have been ways to find the desired topic within the current syndetic structure. Under Ethics, the syndetic structure refers to the subdivision—Moral and Ethical Aspects under topics. Searching Abortion, the user could page through the alphabetic headings index display until the subdivision was reached. However, instruction is necessary to inform the user of search strategies that take advantage of the available dictionary arrangement and syndetic structure.

With LCS and other online catalogs, the challenge remains: What is the best way to create access to subject headings and to display the syndetic structure so that users are led from their search term to the approved subject heading(s) and related terms? Boolean search capability and word truncation are only partial solutions. Software and hardware features (such as windowing and use of color in displays) may also serve as user aids. And unless typical user behavior changes, these features must take into account the tendency of the typical user to enter single words and to be satisfied with an initially successful response.

REFERENCES AND NOTES

5. Lorene E. Ludy, "OSU Libraries' Use of Library of Congress Subject Authorities File," Information Technology and Libraries 4:155-60 (June 1985). The development of subject access in LCS is de-