

# Online Catalog Failure as Reflected through Interlibrary Loan Error Requests

Scott Seaman

*Using data obtained through the interlibrary loan process, this study examines why some users failed to find existing entries in the online catalog at The Ohio State University. Approximately 9% (1,369) of the 1988-89 interlibrary loan borrowing requests were returned to the patron because library staff found the items in the local online catalog. These requests represent either a failure of the user to search the system correctly or a failure of the catalog to retrieve the required record. A sample of the borrowing requests was sorted into user errors and catalog errors. User errors represented 49% of the sample and exhibited five characteristics: no apparent error (43%); spelling errors (7%); incorrect author or title (40%); abbreviations (9%); stoplist words (1%). Catalog failures represented 51% of the sample and five categories of error were identified: punctuation (22%); analytics (16%); corporate word order (28%); truncation (10%); file too large (24%).*



ibrary catalog use studies have a long history. As catalogs have moved from card format to online media, some of these studies have continued to focus on a consistent question: If a user has an accurate and complete author and title, why does a search of the catalog fail to retrieve the corresponding record?

This study examines why some users failed to find existing entries in the online catalog at The Ohio State University. However, rather than relying on the traditional methods of collecting data by interviews or observation of the catalog search process, this study relies on data obtained through the interlibrary loan process. The interlibrary loan process offers a unique perspective on catalog failure. Prior to requesting an item from another university, interlibrary loan staff

routinely search the online catalog to confirm that the item is not held locally. Those items found to be locally owned comprise the data for this study.

Data collected through interlibrary loan offer several advantages. Each request unquestionably represents what the user believes to be a known-item search. Because the data have been collected indirectly, the well-documented problems of interviews, questionnaires, and online observation methods are avoided. Also, it is reasonable to assume that the patron has gone to some length to search the item locally, since the item was considered important enough to wait three to six weeks for the item's delivery. Further, because undergraduate students are not serviced by Ohio State's Interlibrary Loan Office, the data are assumed to be from a population tending

to be experienced library users, i.e., faculty, graduate students, and university staff.

A substantial body of research discusses the interaction between the user and the card catalog.<sup>1</sup> In the tradition of this work, researchers have studied the performance of online catalogs. Primarily this research has focused on the searching of known items, not because known items are more important, but because the results can be more accurately measured.

Librarians generally agree that users have a high success rate searching the online catalog for known items. From questionnaires distributed to users of the Ohio State University Libraries' online catalog, Sammy R. Alzofon and Noelle Van Pulis found that 48% of the searches were for known items with an overall success rate of 81%.<sup>2</sup> At Ohio State, Mary Gouke and Susan Pease interviewed users to determine success or failure of known-item searches in the Libraries' online and card catalogs. They reported that patrons with more than one year's experience searching the online catalog had a 92% success rate when searching known items. The most important factor associated with success in the online catalog, they found, was the users' perception of the complexity of the title. That is, the more complicated the title appeared, the less successful the patron was in finding the existing catalog record. When Gouke and Pease analyzed the search failures, they found that titles appearing to be subject headings often caused users to perform the wrong search. Hyphenated words and stoplist words caused confusion over which words to search. Searches that produced too many matches, even though the correct record did appear, failed because the user did not thoroughly examine the search results. Also, searches performed using abbreviated titles failed because the correct record was not retrieved.<sup>3</sup>

Similarly, Renata Tagliacozzo, Lawrence Rosenberg, and Manfred Kochen interviewed users in four large libraries to analyze patterns of card catalog searching and success. Their findings indicated

that users often had correct titles but, nevertheless, preferred to search by author although it was less likely the user had a correct author statement. They found that, depending upon the library, 7 to 20% of all searches failed to find a corresponding entry. Also, they found that less than 50% of users went beyond the first attempt to locate a known item and that only 4% of users went beyond a second.<sup>4</sup>

Jean Dickson analyzed Northwestern University's NOTIS transaction log to classify "zero-hit" searches. Dickson found that 26% of zero-hit author searches were entered with forenames first, preventing the user from finding a title in the database. Nearly 64% of zero-hit title searches were caused by incorrect spelling of the first or second word of the title or by the inclusion of an initial article.<sup>5</sup>

Traditionally, catalog failure studies have relied upon user questionnaires (Gouke and Pease, Alzofon and Van Pulis, Tagliacozzo et al.) or some form of unobtrusive observation (such as the transaction log analysis of Dickson) to compile data. But reliance on both methods of data collection have been questioned. Jerry Specht observed that transaction log data are inherently limited in their capacity to record successful searches and search failures.<sup>6</sup> James Krikelas, for example, asserts that relying upon observed behavior is not accurate and that many previous studies have inadvertently mixed known-item with what are truly subject searches. Like Gouke and Pease, Krikelas found that a user's search strategy is affected by the bibliographic information available and the user's perception of the complexity of that information. As the information about an item becomes complicated or less distinctive (such as with corporate authors or technical reports), there is a corresponding increase in users creating title or subject searches. Consequently, many users searching complex items resort to subject searches rather than to using the information at hand.<sup>7</sup> Similarly, Ben-Ami Lipetz demonstrated that users often employ a known-item search to locate a subject heading for browsing.<sup>8</sup>

## THE ONLINE CATALOG

Ohio State University's Library Control System (LCS) became operational in 1970 as a circulation system and in 1975 as a public catalog. Currently 235 dedicated public access terminals are available throughout the university's library system serving a population of 55,000 students, 4,500 faculty, and 16,500 staff.

Of the 4.3 million volumes in Ohio State's collection, 2.9 million are cataloged on LCS. Of the 2.9 million titles, about 1.3 million have only brief location displays that include author, title, date of publication, holding library, and availability statement. The remaining 1.6 million titles include a full bibliographic record with full descriptive information. Although the card catalog is available, it has not been maintained since 1982. Analytics for items cataloged before 1983 are not yet represented other than by a search of the serial title. Access to added entries and subject headings is available for Ohio State titles added since 1974.

LCS is a command-driven system searchable by author, title, author/title combination, subject heading, call number, uniform title, and series title. It creates a search algorithm from the full title or author and title the user enters. LCS ignores initial and internal articles. Boolean combinations or keyword searching is not available. The increased searching power of Boolean or keyword searching could, in cases of incorrect or incomplete information, lead to increased retrieval on the part of skilled searchers.

### HYPOTHESIS AND DESCRIPTION OF DATA

The hypothesis of this study is that certain characteristics of the online catalog, or its file structure, inhibit users from locating existing catalog records in the online database. Further, these characteristics are such that a user with complete and correct information will not, in some cases, locate the catalog record corresponding to a bibliographic search.

Data to test this hypothesis are available through the interlibrary loan process.

During the 1988-89 academic year, 15,300 requests to borrow books and articles were submitted to the University Libraries' Interlibrary Loan Office. Each request was searched in the local online catalog to verify that the item was not available locally before ordering from another library. About 9% (1,369) were returned to the requestor because interlibrary loan staff found the item in the local online catalog. These requests represent failure of the user to search correctly the system or a failure of the catalog to retrieve the required record.

### THE SAMPLE

The large number of requests (1,369) necessitated sampling for analysis. It was anticipated that the data would fall into two broad categories: user error and catalog error. The required sample size was determined by using a formula for estimating a population proportion (the formulas are available from the author upon request). It was found that to represent accurately the population, the sample size must be at least 226.

### DEFINITIONS AND METHODOLOGY

Using a table of random numbers, we generated 226 numbers and arranged them sequentially. The 1,369 error cards from the 1988-89 academic year are maintained in a file arranged alphabetically by main entry. Starting at the beginning of the file, the card that matched the random number was pulled, photocopied, and replaced. Online catalog searches were performed on each of the 226 photocopied cards and the printouts attached to provide correct bibliographic information.

The sample was sorted into two categories: those cards exhibiting user errors and those exhibiting catalog errors. User errors were defined to be mistakes for which a catalog could not be expected to compensate, such as spelling errors or incorrect titles. Catalog errors were defined to be those instances when the user possessed and presumably entered the correct bibliographic data, but the catalog structure impeded the user in locating the matching online record. The

objective in sorting by user and catalog error was to separate obvious spelling and citation errors from the failures that the catalog structure imposed.

The user error category comprises:

1. **No Apparent Error.** Correct and complete author and title given on the interlibrary loan request, but the user failed to find the online record.
2. **Spelling Errors.** Errors either in the author's name or in the 4,2,2,1 title characters used as the search algorithm. An example is citing *Maria Luisa* as *Maria Lisa* or citing *South West Review* instead of *Southwest Review* on the interlibrary loan card.
3. **Incorrect Author or Title.** Personal author or title-main entry citations for which the wrong author or incorrect title was provided. An example is citing *The Social Effects of Mental Illness* rather than the correct title, *The Social Consequences of Psychiatric Illness*.
4. **Abbreviations.** Citations for which the title was abbreviated such that it could not be accurately searched. An example is *S-Foodserv.-J.* for *School Foodservice Journal*.
5. **Stoplist Words.** LCS, like most online catalogs, does not search frequently used words in the title or author/title search. Primarily these are initial articles such as: *a*, *an*, *from*, *ha-*, and *the*. The patron may omit the stoplist word or LCS will automatically read over the word.

After examining the data, five subcategories of catalog failure were identified:

1. **Punctuation.** In the title fields that affect the search algorithm. For the title, *Problem-Solving Processes of College Students*, *Problem-solving* must be searched as a single word in LCS.
2. **Analytics.** Series titles for which no analyzed entry exists in the online catalog. For example, *Anna Rutledge's Artist in the Life of Charleston* is volume 39 of the *Transactions of the American Philosophical Society*. Both titles are correct, but only the series title is

accessible on Ohio State's online system for titles cataloged before 1983.

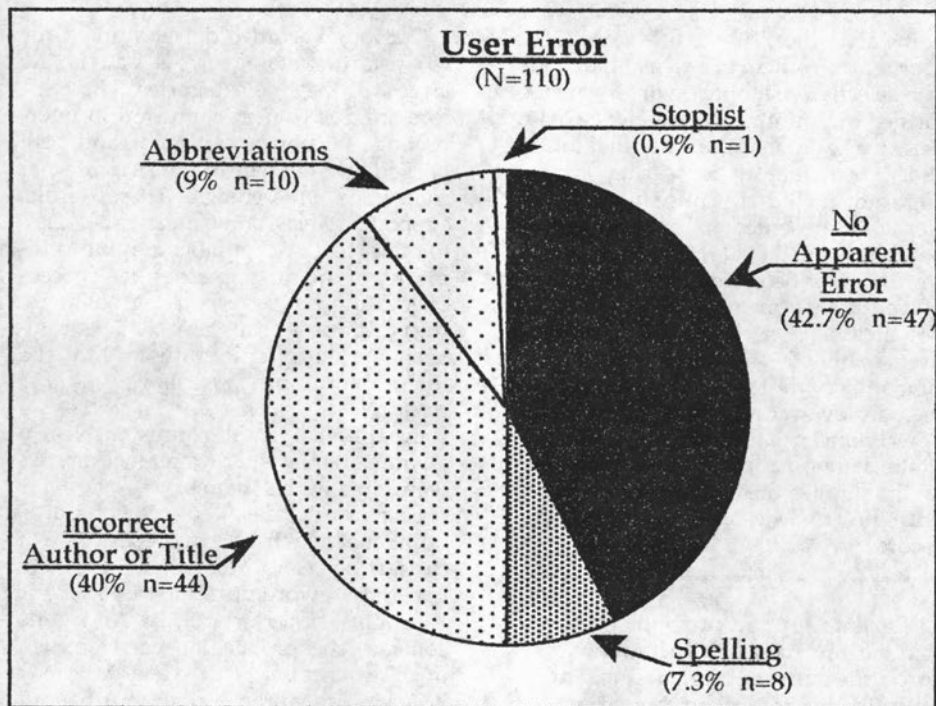
3. **Corporate Word Order.** Items produced by corporate bodies for which the author or title is not clear. For example, a request for *Proceedings of the International Conference on Robot Vision and Sensory Controls* consists of the author *International Conference on Robot Vision and Sensory Controls* and the title of *Proceedings*. Users often have difficulty distinguishing between the title and author and, consequently, fail to locate an item that is in the local catalog.
4. **Truncation.** Ohio State's online catalog allows for truncation of title entries by adding a hyphen to the end of a search. However, if the user fails to insert the hyphen after keying fewer than four words of the title, the online system will fail to retrieve the record. For example, a title search for *The Proud Heritage* will fail to retrieve *The Proud Heritage of Cleveland Heights, Ohio*.
5. **File Too Large.** Searches for which a large number of matches are retrieved interfering with locating the correct record. For the purposes of this study, it was assumed that whenever the search resulted in more than six screens of matches (approximately 50 titles), the file was too large for the user to interpret adequately. Examples are *Science*, *Scientific American*, and *Acustica*.

## RESULTS

It was determined that of the sample of 226 requests, 110 were user failures (49%). Catalog failure accounted for 116 requests (51%). Confidence limits were established for user and catalog failure. It was found with 95% confidence that the true proportion of user failures lies between 43% and 55%. Similarly, one is 95% confident that the true proportion of catalog failures lies between 45% and 57%.

### User Error

Figure 1 details the categories of user error. The largest, at 43%, were those requests with no apparent error. On



**FIGURE 1**  
User Error by Subcategory

those requests, the author and title were correct and complete. Each request was found in the local database using either the author, title, or author/title search with the information provided on the request card. That this statistic is so high may reflect Tagliacozzo, Rosenberg, and Kochen's findings that only 50% of users go beyond the first search when attempting to locate an item.

---

**The hypothesis of this study is that certain characteristics of the online catalog, or its file structure, inhibit users from locating existing catalog records in the online database.**

---

Requests having incorrect authors or titles represented 40% of the user errors. The user, most likely, brought to the online catalog an incorrect citation. For example, one user gave the title *Local Political Opposi-*

*tion in Japan* for a monograph titled *Political Opposition and Local Politics in Japan*. Often these errors were identified during the OCLC or RLIN searching process. They were then searched again with the correct information in the local catalog. While many of these failures are presumably due to faulty memory on the part of the user, the influence of prepublication information may have some influence on these data.

The other three categories—abbreviations, stoplist words used in the search, and spelling errors—account for only 17% of the user errors. That these categories are so low probably reflects the nature of the data. Presumably, by the time the user requests an item through interlibrary loan, he or she has searched the citation multiple times to minimize such errors.

Only one stoplist error was found among the data, suggesting the construction of the stoplist is adequate. LCS automatically omits stoplist words when

doing an author/title or title search. The user may put them in or leave them out. The selection of stoplist words was performed by counting the frequency of articles in the database. To keep the stoplist short, infrequently appearing articles were not included. Consequently, articles not represented on the stoplist must be included in the search by the user. The stoplist error detected by this sample involved the German article *das*, which is not included on the LCS stoplist. The user searched "Freundschaftsbild der Romantik" and found no matching LCS record. However, a matching record, titled "Das Freundschaftsbild der Romantik," is in the database. Because *das* is not listed on the stoplist, the title must be searched with the article to retrieve the matching record.

---

**These data support previous findings that patrons have difficulty with corporate authors, punctuation, and with finding records in large files.**

---

Previous studies have indicated that spelling errors represent a much larger proportion of catalog failure. Dickson reported that 64% of title failures were due to incorrect spelling, and Henty found 33% of unsuccessful keyword entries were spelling errors. In both cases, this information was compiled from transaction log analyses. One limitation of transaction log data is that they cannot identify if a search is eventually successfully completed or not. Data collected in this study suggest that a significant proportion of those were typographical errors, rather than spelling errors, which the user corrected.

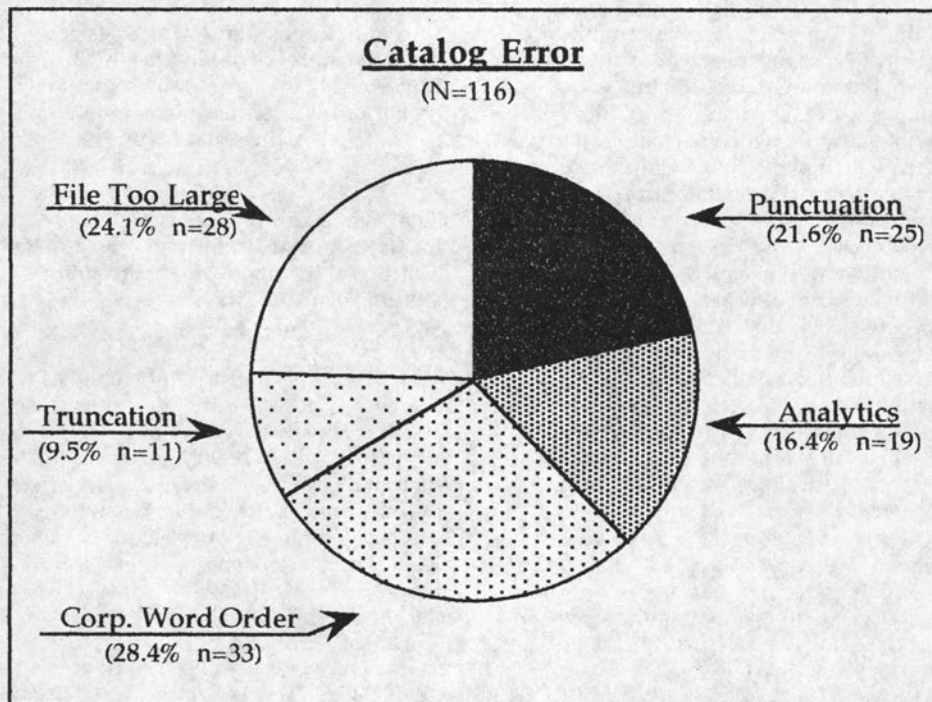
#### *Catalog Error*

Catalog error, those errors over which the patron had little or no control, accounted for 116 of the sample of 226 or 51% of the error cards. Unlike the user errors, catalog errors proved to be more evenly distributed throughout five sub-categories. Figure 2 depicts the type of errors found.

Corporate word order accounted for 28% of the catalog errors and is the largest of the five categories. These errors are due to user confusion in interpreting corporate authors and are particularly evident in the searching of conference proceedings. These results support Krikelas' and Gouke and Pease's findings that, as the bibliographic complexity of the title increases, the success of locating the corresponding online record decreases. That many citations combine the corporate author into the title—for example, *Proceedings of the First International Conference on Fracture*—is a source of considerable confusion. Nearly all of the requests for conference articles were cited in this manner.

Searches that resulted in a very large number of matches accounted for 24% of the catalog errors. Most often these were one- or two-word titles such as, *Scientific American*, *Science*, or *Aperture*. And while some searches produce too many matches, other searches produce too few. If the user is lacking either the third or fourth word or perhaps a subtitle, LCS will not retrieve the record without truncation. Without truncating, for example, the search for *Euphorion* will not retrieve any records for the title cataloged as *Euphorion: Zeitschrift Für Literaturgeschichte*. In this instance LCS requires that the subtitle be included to retrieve the record. This type of truncation problem accounted for 10% of the catalog errors.

Punctuation problems appeared in 22% of the catalog error items. Of the twenty-five errors, ten were hyphenation problems in the first four words of the title. Examples are: *HaydnStudien*, *Problem-Solving Processes of College Students*, and *Robotics and Computer-Integrated Manufacturing*. The remaining 15 are foreign-language titles, including punctuation, that, if keyed into the search string, cannot be interpreted by LCS. *Late Ch'ng Views on Fiction* is an example of this. Users' difficulty with hyphenation was noted by Gouke and Pease. Henty found that 14% of keyword failures were due to punctuation errors. Unlike spelling errors, however, these data suggest that users are not correcting



**FIGURE 2**  
Catalog Error by Subcategory

punctuation errors—particularly those involving foreign names and titles.

Analytics for items cataloged before 1983 are not yet represented in LCS. For example, Anna Rutledge's *Artist in the Life of Charleston* is volume 39 of the *Transactions of the American Philosophical Society*. Only the pre-1986 serial title is accessible on Ohio State's online system. The user must not only know the title of the series and the volume number but must also know how to search by the series title. Failure to identify successfully these items represents 16% of the catalog error. While this represents a source of error unique to Ohio State's LCS, it does suggest that analgesics are an important part of the bibliographic information available in the catalog.

### CONCLUSIONS

Overall, the catalog failure data collected through interlibrary loan supplement the findings based on interviews,

questionnaires, and online observation methods. These data support previous findings that patrons have difficulty with corporate authors, punctuation, and with finding records in large files. However, the data suggest that spelling errors, although they may be common among search transactions, do not prevent users from finding the correct online record. Further, it appears that stoplist words and truncation are being used successfully by at least some users. It can only be speculated as to why such a large proportion of users made no apparent errors but failed to locate the online record. However, anecdotal evidence from the ILL staff suggests that many of these users failed to search the online catalog before submitting an ILL request.

A significant portion of the user population, apparently, comes to the online catalog with an incorrect citation. Most likely this source of error is from the user's relying upon memory of a cita-

tion. However, incorrect prepublication information or incorrect citations are also possible sources of this error. Although not available at Ohio State, Boolean or keyword searching, particularly on the full bibliographic record, may retrieve the desired record although complete bibliographic information is lacking.

However, Boolean or keyword capability may not affect all types of searches. Because users typically search corporate authors by title rather than author, it is doubtful that such capability searching would improve retrieval. That even experienced users fail to locate corporate authors suggests that additional bibliographic instruction should focus on this issue.

These data suggest that punctuation errors are most likely to occur while searching foreign authors or titles. However, when punctuation errors ap-

peared in English-language titles, it was often the presence of a hyphen in the first four words that aborted the search. Perhaps more added entries could reduce the number of punctuation errors.

---

**That even experienced users fail to locate corporate authors suggests that additional bibliographic instruction should focus on this issue.**

---

Finally, the number of matches a search retrieves is a function of catalog size, search strategy employed, and the structure of the search algorithm. Although altering the algorithm may affect the number of matches, the search strategy the user employs can profoundly affect what is retrieved. Consequently, user education could significantly influence this failure also.

---

#### REFERENCES AND NOTES

1. Ruth Hafter, "The Performance of Card Catalogs: A Review of Research," *Library Research* 1:199-222 (Summer 1979); E. A. Montague, "Card Catalog Use Studies 1949-1965," (Master's thesis, Univ. of Chicago, 1967).
2. Sammy R. Alzofon and Noelle Van Pulis, "Patterns of Searching and Success Rates in an Online Public Access Catalog," *College & Research Libraries* 45:110-15 (Mar. 1984).
3. Mary Noel Gouke and Sue Pease, "Title Searches in an Online Catalog and a Card Catalog," *Journal of Academic Librarianship* 8:137-43 (July 1982); Sue Pease and Mary Noel Gouke, "Patterns of Use in an Online Catalog and a Card Catalog," *College & Research Libraries* 43:279-91 (July 1982).
4. Renata Tagliacozzo, Lawrence Rosenberg, and Manfred Kochen, "Access and Recognition: From User's Data to Catalogue Entries," *Journal of Documentation* 26:230-49 (Sept. 1970).
5. Jean Dickson, "An Analysis of User Errors in Searching an Online Catalog," *Cataloging and Classification Quarterly* 3:19-38 (Spring 1984).
6. Jerry Specht, "Patron Use on an Online Circulation System in Known-Item Searching," *Journal of the American Society for Information Science* 31:335-46 (Summer 1980).
7. James Krikelas, "Searching the Library Catalog—A Study of Users' Access," *Library Research* 2: 215-30 (Fall 1980).
8. Ben-Ami Lipetz, "User Requirements in Identifying Desired Works in a Large Library. Final Report," Yale University Library, New Haven, Conn., ED 042 479.