

Notes on Operations

Utilizing Z39.50 to Obtain Bibliographic Copy

A Cost-containment Study

Christine DeZelar-Tiedman, Cecilia Genereux, and Stephen Hearn

This paper looks at one approach to controlling costs when seeking cataloging copy. A small task group in the University of Minnesota Libraries Technical Services Department conducted a study to devise the most cost-effective strategy for searching for and importing bibliographic copy, by compiling costs and benefits of importing records from the OCLC Online Computer Library Center database, the Research Libraries Group Union Catalog (RLIN), and the Library of Congress (LC) catalog. Results of the study indicated that, although the LC database is smaller than the other two utilities, a sufficient portion of needed records were available from LC to more than offset the cost of re-searching in the other databases for records not found. In addition, due to differences in pricing structure, searching RLIN second was found to be more cost effective than going next to OCLC, even though a slightly larger proportion of items were found in OCLC than RLIN. This study may prove useful either as a research method or in terms of its findings for other libraries wishing to compare sources of cataloging copy.

Many libraries and their technical services units face similar challenges, including steady state or declining operating budgets, implementation of new integrated library systems (or version upgrades), and continuing interest in improving processes to gain efficiencies. One goal in cataloging units often is to locate and use as many copy cataloging records as possible in order to speed movement of items to the collection and reduce the number of items that require original cataloging. This paper reports the results of a research project at the University of Minnesota Libraries (UML) Technical Services Department that explored costs and workflow issues related to obtaining bibliographic copy from the OCLC Online Computer Library Center (OCLC) database, the Research Libraries Group (RLG) Union Catalog (RLIN), and the Library of Congress (LC) catalog.

The authors describe the record import process and workflow prior

to the project, present the questions that the project sought to answer, and report the research method and findings. A literature review places this project within the context of other cost studies. Implemented changes and the resulting financial impact are reported. Finally, the authors draw conclusions for other libraries considering a similar cost-benefit analysis.

Background

In spring 2003, a small group, the Source of Records Task Group (SRTG), was formed at UML to explore issues concerning cost and workflow for obtaining bibliographic copy from various sources. UML's technical services department was in the midst of revisiting and redefining workflow as a result of a July 2002 conversion from the NOTIS system to the ExLibris Aleph500 integrated library system, and the department was also seeking to

Christine DeZelar-Tiedman (dezel002@umn.edu) is Archives and Special Collections Catalog Librarian, **Cecilia Genereux** (gener002@umn.edu) Serials Cataloging Coordinator, and **Stephen Hearn** (s-hear@umn.edu) is Section Leader, Database Management and Authority Control Coordinator, at the University of Minnesota Libraries, Minneapolis.

reduce expenditures as part of a libraries-wide budget retrenchment.

UML has a longstanding policy of cataloging in the local system rather than directly in the OCLC database or RLIN and of using batch loading to contribute to the utilities. RLIN (now RLIN21) properly refers to the set of database services provided by RLG. In this paper, following common usage in the profession, RLIN refers more narrowly to the RLG Union Catalog. Because of the variety and breadth of UML's collections, the libraries have found using and contributing to both services beneficial. At the time of the system conversion in 2002, system implementers decided to discontinue using OCLC Passport and RLIN for Windows client interfaces (hereafter referred to as native clients) for the purpose of searching and downloading records into the local system. Instead, implementers explored the option of importing records through Z39.50, a protocol that allows a client to search multiple remote servers using a single search interface. One reason for this was a desire to provide staff members with a uniform interface for bibliographic searching so that they need only be trained in a single set of commands and search string formulation criteria. Additionally, both RLG and OCLC were in the midst of, or soon to be embarking upon, significant redesigns of their own search interfaces (RLIN21 and OCLC Connexion, respectively). Asking staff to learn three new search interfaces in rapid succession seemed like too much change in too short a time. The RLIN for Windows (later RLIN21) client would continue to be used for some specialized functions, such as East Asian vernacular cataloging and Name Authority Cooperative Project (NACO) authority work.

With these factors in mind, the local Aleph system was configured to import bibliographic copy from OCLC and RLIN using Aleph's Z39.50 interface. A basic outline of the workflow

for this procedure follows, based on Aleph500 version 14.2. In this outline, *user* refers to the staff member performing the searching and copy cataloging functions.

- Upon logging in to Aleph's searching module, the user is automatically connected to the local catalog; a menu is available to allow the user to connect to other databases. The user selects an external database (such as RLIN or OCLC), which is accessed using the Z39.50 protocol.
- The user executes a search, using the same types of commands available to perform a staff search of the local catalog. While the choice of indexes is more limited than what is available for the local database, and slight differences in formulating searches exist in some cases, the process is relatively uniform across databases.
- If appropriate bibliographic copy is found, the user selects the record and moves it to the cataloging module. Users also may select several records from a single retrieval set and move them to the cataloging module in a batch.
- In the cataloging module, the user executes the duplicate command, which creates a working copy of the record on the user's hard drive. The user is prompted to select a format type (such as books, serials), and then is able to edit the record as needed, depending on the level of the staff member and the stage in the workflow.
- Once a record has been saved to the server, associated orders, holdings, and item records are created as needed.

Figure 1 presents a workflow chart of these procedures.

Some significant challenges arose in implementing this new process. First, while the Aleph system theoretically allowed for this method of obtaining bibliographic copy, no other North American library had yet implemented it. Therefore, UML systems staff had to work very closely with ExLibris programmers, as well as with staff at OCLC and RLG, to develop a workable configuration.

The initial Z39.50 configuration was confined to a few indexes, with little or no ability to limit or sort search results. This led to some frustration among staff members, particularly those who were searching for materials lacking standard numbers, such as International Standard Book Numbers, International Standard Serial Numbers, or Library of Congress Control Numbers, or for materials with common one-word titles, frequently found with serials. In this period of transition, implementers decided not to immediately remove staff access to the native clients while encouraging the transition to Z39.50. However, technical and anecdotal evidence indicated that many staff were searching the native client to obtain an RLIN or OCLC number, and then going back to the Z39.50 interface to retrieve the record or, because of unfamiliarity with the new interface and mistrust of their initial search results, were following up the Z39.50 search with another search in the native client to make certain they had retrieved a complete result set. The per search cost for Z39.50 searches in OCLC is higher than for RLIN Z39.50 searches or for native client searches in either utility. Consequently, staff were duplicating searches, resulting in double charges and increasing the bibliographic utility bills, particularly for OCLC.

Problem Statement

In an environment where cost cutting was needed, the trend of rising

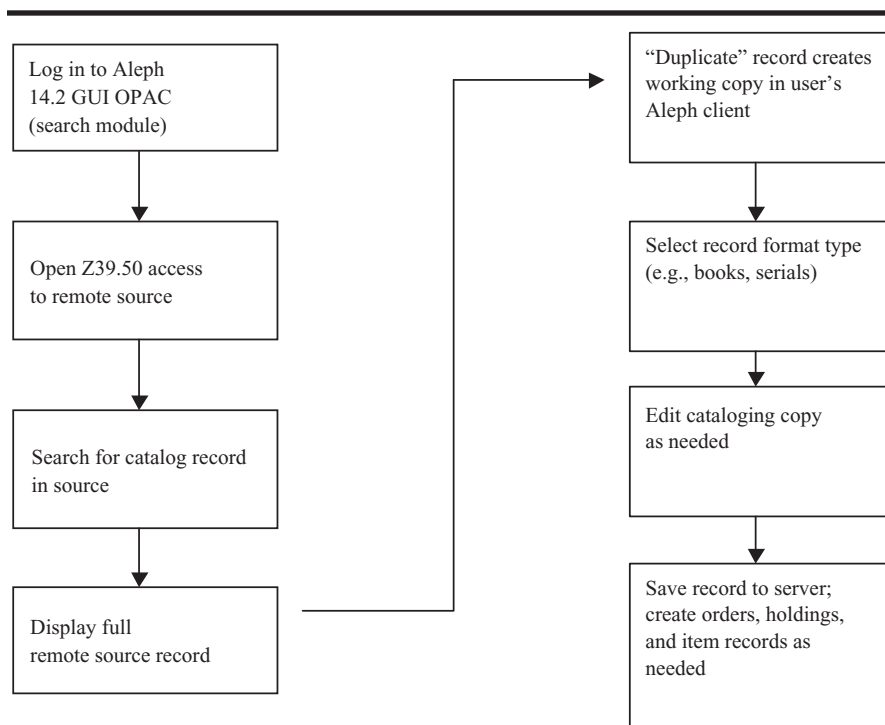


Figure 1. Workflow chart for importing bibliographic records in Aleph500 14.2 through Z39.50

bibliographic utility costs was a cause for concern. SRTG sought ways to dramatically cut searching costs. In order to achieve this, the task group used a two-pronged approach. The first recommendation, now successfully implemented, is not the focus of this study. In brief, it recommended that additional search indexes and search limits be made available for Z39.50 searching, and that in-depth training on searching each of the databases be provided to all technical services staff. This helped staff gain confidence that the Aleph client's Z39.50 search results are complete and accurate.

The second recommendation was to perform a study to explore whether providing Z39.50 access to a third bibliographic database, LC's online catalog, would significantly reduce searching and retrieval costs. Access to the LC database and the import of MARC records is available free of charge. However, it is a smaller database than either OCLC or RLIN. Would the

cost savings accrued by obtaining free records from LC be offset by the additional staff time spent searching in other databases when records are not found in LC? Or is it more efficient to search the larger databases first, accruing a higher search and download charge, but using less staff time? These factors had to be weighed in the task group's decision-making.

The study sought to answer the following questions:

- Overall, which is the most cost-effective order in which to search for materials through Z39.50 in the LC, RLIN, and OCLC databases?
- What are the relative levels of cost effectiveness in relation to searching and importing records from LC, RLIN, and OCLC through Z39.50 as compared to searching and importing through the OCLC and RLIN native clients?

- Based on such criteria as format, place of publication, and year of publication, are particular types of resources more likely to be found in specific databases, thereby precluding the previous searching order for these resources?

The data developed for the study and discussed in the following were able to provide working answers to the first two questions. Additional research will be needed to answer the third. However, because UML has chosen to maintain its policy of using the Aleph client for searching all three sources, no practical test or confirmation of predicted results can be reported for the use of OCLC and RLIN native clients with their more complex cost factors. Therefore, though the research method needed for exploring the second question is discussed, only the results relevant for searching with Aleph's Z39.50 client are reported here.

Literature Review

A survey of literature on the topic of cost studies, hit rates, Z39.50 searching, and the cost of obtaining cataloging copy yielded a number of articles, but few pertaining directly to the UML study. Several articles have been written about cost studies conducted by individual institutions. One of the best documented is Iowa State University's longitudinal cost and time study in technical services. Morris; Osmus and Morris; Morris, Rebarcak, and Rowley; Morris et al.; and Fowler and Arcand detail the Iowa findings at various points and relating to different categories of technical services work considered during the study.¹ The Iowa study breaks down technical services activities into product centers (products and services) and overhead centers (leaves, administrative and professional activities, meetings, and so on). Overhead

costs are calculated and then allocated to product centers, providing a look at the true cost of technical services activities. Product centers are separated into the five activities of acquisitions, cataloging, catalog maintenance, volume preparation, and preservation. The cataloging product center includes searching for cataloging copy, which is done primarily by cataloging staff. The Iowa reports indicate the number and percentages of staff hours spent on searching in the aggregate, but they do not provide detailed information about per-search costs.

Two of the Iowa study reports mention the percentage of bibliographic records found in OCLC, and one categorizes the type of cataloging records found. In the first article about the Iowa study, Morris writes that, "ISUL [Iowa State University Library] catalogs more than 90 percent of all new monographic titles at receipt with copy found in the OCLC database."² In a later report Morris, Rebarcak, and Rowley indicate that, "over 90% of the monographs ordered in 1994-95 had an OCLC cataloging record at the point of pre-order search."³ Providing the only information on the type of cataloging records obtained from OCLC and the PromptCat Service, Morris et al. report that in 1997-98, in comparison with 1990-91 data, LC copy rose from 46 to 59 percent, member copy rose from 27 to 40 percent, CIP declined from 27 to 1 percent, and 6,325 LC records were added through the PromptCat Service.⁴

Research and reports on hit rates in bibliographic databases have been published since the late 1970s. However, the authors were only able to find three studies comparing hit rates between RLIN and OCLC. In the 1990 article, "Chasing MARC: Searching in Bibliofile, Dialog, OCLC, and RLIN," Allan compared hit rates in Bibliofile, Dialog, OCLC, and RLIN by searching a sample of 1,000 English and foreign language monographs in April 1987.⁵ For items

with no exact match, a second search was conducted one year later. The first round of searching revealed that cataloging copy was found in OCLC for 86.3 percent of the titles and for 76.6 percent in RLIN. After the second round of searching a year later, the percentages of matches increased to 91.2 percent in OCLC and to 85.9 percent in RLIN.⁶

The second article comparing RLIN and OCLC is "Cooperative Cataloging of Latin-American Books: The Unfulfilled Promise," in which Grover examined hit rates in the two databases and whether the cataloging copy found was LC- or member-provided.⁷ Grover used a sample of 298 humanities and social science monographs from Latin American countries from three different libraries, with the majority coming from a medium-sized academic library. Each item was checked at six-month intervals between November 1983 and May 1985. Grover concludes that, "There were no important differences between OCLC and RLIN, and in the end, both systems had cataloged almost the same number, although not the same books."⁸ Grover also notes that almost 50 percent of the books in both systems were first cataloged by LC.⁹

In a report of a study examining the availability of bibliographic records in RLIN and OCLC for Spanish- and Portuguese-language monographs, Erbolato-Ramsey and Grover find that the two bibliographic databases are comparable.¹⁰ Using a sample of books that ranged from just receipted to 46 months in a cataloging backlog, searches were conducted first in RLIN. If no cataloging copy was found, then the title was searched in OCLC. Hit rates were broken down into percentages of titles found in RLIN, on both RLIN and OCLC, and no copy found. The authors found that, "By the end of the sixteenth month, the percentages were 84% on RLIN, 91% on RLIN and OCLC combined, and 9% for items not found in either

system."¹¹ The authors also found that LC cataloging records were available for 53 percent of the titles in the first six months, peaking at 65 percent at twelve months.¹²

Little in the literature examines the benefits of obtaining cataloging records from national libraries. Only the recent article by Beheshti, Large, and Riva, reporting on the cost savings to Canadian universities and large urban public libraries by using MARC records provided by the National Library of Canada (NLC), relates closely to the UML findings.¹³ The Canadian study found that, "The average annual cost saving for a university library when using NLC MARC records for derived cataloging for Canadian monographs and federal government documents is \$16,400, while the average savings for a large urban public library is \$7,800."¹⁴ Not only do libraries save on the cost of original cataloging, but Beheshti, Large, and Riva also report that most libraries indicated that they obtain the NLC records through a free source. Of the sources mentioned by libraries, NLC's online catalog, Amicus Online, was the "most frequently cited *single* source," and "Web OPACs including Z39.50 servers are used by 76% of libraries."¹⁵

Few articles about the use of Z39.50 by technical services to search for bibliographic records have appeared in the library literature. Reporting on the benefits LC derives from the use of BIBCO records (bibliographic records contributed to the LC catalog by libraries participating in the Bibliographic Record Cooperative Program), Wiggins mentions that the records are searched for and obtained through Z39.50 from OCLC and the RLG union catalog.¹⁶ Wiggins also notes that cataloging teams will search for BIBCO or other source records again upon receipt of items, "following their hunches about titles that are likely to be represented in OCLC or the RLG union catalog."¹⁷

The first priority of the UML

study was to determine the most cost-efficient sequence for searching OCLC, RLIN, and the LC catalog for typical records using Z39.50. Along with the most cost-efficient search order, the study sought to suggest preferred sources of more specialized copy because, just like LC staff, staff at UML often use their intuition about where copy is likely to be found. If preferred record sources could be determined for specific types of materials, then staff could avoid unnecessary searching. The two approaches work to address both the efficiency and effectiveness of searching. This latter question was not adequately answered by the task group's sample data, so UML currently follows LC in relying on searchers' hunches when searching for specialized copy.

Method

Sampling Searching Hit Rates

To get an indication of the best source of records for different types of materials, SRTG developed an interactive project database to track hit rates for a sample of materials that staff might search in their daily work. Ten staff members were asked to keep track of their searches during a typical work-week, gathering statistics on what level of copy was found in the LC database, OCLC, and RLIN. The group defined *typical* to mean a week when the variety of items searched was not affected by such exceptional circumstances as a special ordering, a cataloging project focused on a particular type of material, system downtimes, or other factors. The task group relied on the judgment of staff doing the sampling that the week reviewed was not unusual. The project database allowed staff to indicate whether full copy, minimal copy, or no copy was found in each of the three databases, as well as the format, language, place and date of publication, and broad subject area of the materials they were searching.

To achieve a broad cross-section of searching, the task group deliberately chose staff who work with a variety of materials. As student workers perform initial searching in some units, student supervisors were recruited for the project as well. However, a simplified definition of *full copy* for students gathering statistics was provided (see appendix), as they could not be expected to make the more refined judgments that a permanent staff member would be likely to make regarding level of copy.

Staff were instructed to search the LC database first to locate available copy. Searchers recorded whether they found no copy, minimal copy needing additional work, or full copy. The task group provided detailed definitions for these terms, as defined in the appendix.

The searchers then looked for each title in OCLC and RLIN, recording the results. Although in an actual workflow staff would stop searching once a full record is found (or, if the search was being performed at time of order, when any level of copy is found), searchers were asked to perform each search in all three databases to make full comparisons of the hit rates per database.

Per-Item Cost Estimates

The task group also was aware of the need to consider staff time and the cost of trying more than one search in a source and of repeating searches in multiple sources as factors in determining preferred sources for materials. To address these cost components, sample data was gathered for the time required to search and download a record from each source. Estimates were made of the proportion of instances in which a single search would not suffice and two or three searches in a single source would be required. Lastly, the costs charged per search by each source were determined. For Z39.50 searches,

these per-search charges were fairly straightforward for UML in the year of the study: \$1.00 for OCLC, \$0.59 for RLIN, and \$0.00 for LC. More involved calculations were required for estimates of the cost of native client interface searches done in OCLC and RLIN, as standard number searches are charged differently from other kinds of searches. Details of these cost calculations can be found in the final report of the task group.¹⁸

These data and estimates were then combined into formulae to express the cost per item for searching in each of the three sources. For each source, the cost of each type of search and the staff time required for each search were multiplied by the estimated frequency of that search type in the overall searching process. The cost of the total number of estimated searches was then divided by the number of searches to produce a figure for the average cost of searching each source using the Aleph Z39.50 client or using the native client interfaces (OCLC Passport and RLIN Terminal for Windows). The latter figures are included in the task group report. However, as both utilities have since converted to new native interfaces (OCLC Connexion and RLIN21), these calculations will not be reported here.

Per-Search Cost Averages

By combining these per-item costs with the results of the study sample, the task group was able to determine the comparative costs of searching all items in one source, then all unfound items in the next source, and all remaining unfound items in the third source. Because LC's full records are distributed to both OCLC and RLIN, a sequence calling for searching LC after searching either OCLC or RLIN was considered superfluous. Given that each sequence was known to produce the same number of total records found, the task group could

then calculate the most cost-effective sequence of sources for searching the sample set of records. This was the primary goal of the study—to recommend a preferred sequence for searching these three sources in terms of cost effectiveness, not to measure the coverage of the sources as compared to each other.

Results

Sampling Searching Hit Rates

Data was submitted for a total of 433 items searched, of which 300 were books and 133 were nonbook formats. While the data gathering process allowed results measurement by format, language, publication place and date, and subject, the complexity of the results and the relatively small sample size for particular characteristics make reliably predicting the best source of copy based on each of these separate criteria difficult. However, some generalizations can be made from the data.

Full copy was found in the LC database for 39 percent (168) of the 433 titles searched, as compared to 57 percent (246) in OCLC and 52 percent (224) in RLIN. When positive hits for either full or minimal copy are combined for each source, the results were 51 percent (220), 81 percent (353), and 74 percent (320), respectively. Because OCLC and RLIN both contain overlapping but different sets of full and minimal records not available in LC, UMLs overall hit rate using all three sources was 78 percent (340) for full records and 91 percent (394) for full plus minimal records. Table 1 and figure 2 illustrate the hit rates for the sample set searched in each of the three databases and the overall hit rate for the three databases combined.

These numbers suggest that a sufficient portion of copy is available free of charge from LC to justify using this as the first source of copy. This assumption will be further explored in

Table 1. Hit rates by utility and level of copy found

Database	Full copy		Copy needs work		No copy	
	No.	%	No.	%	No.	%
LC	168	39	52	12	213	49
OCLC	246	57	107	24	80	19
RLIN	224	52	96	22	113	26
Combined	340	78	54	13	39	9

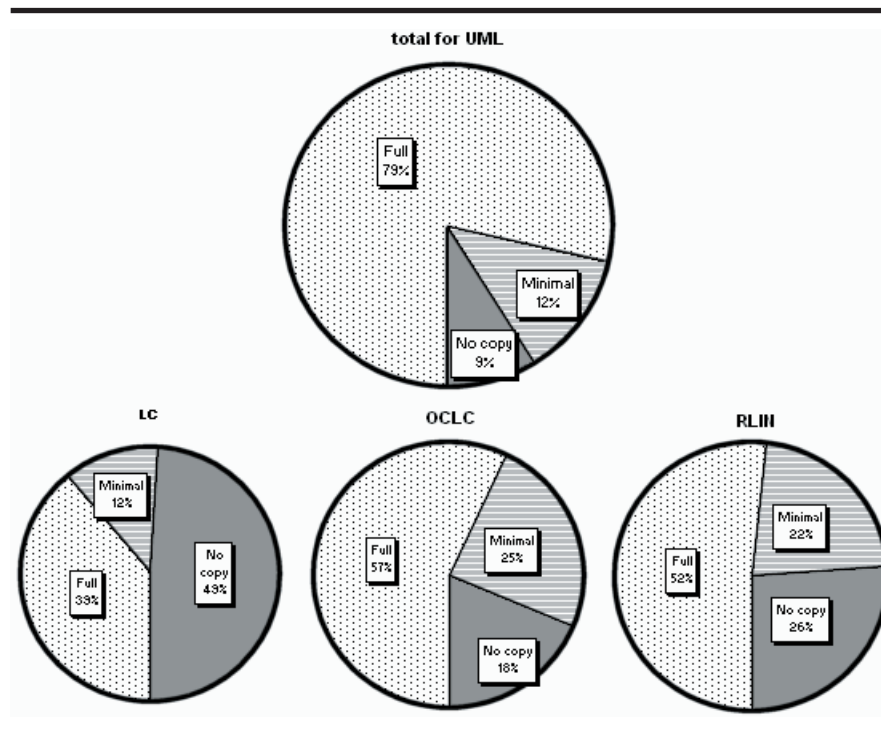


Figure 2. Hit rates by utility and level of copy found

the next section, which also considers time spent by staff in searching multiple databases.

To define more precisely the best source of copy for materials based on such specific criteria as language or format, the task group would need to collect additional data to build up a more reliable sample size. For example, the data gathered indicates that the LC database may be a poor source of records for music scores, but a good source for maps. However, due to the small number of data points gathered, the task group cannot conclude this reliably.

Per-Item Cost Estimates

The per-item searching transaction costs represent an average of various kinds of search outcomes, based on the working assumption that in 10 searches:

- 5 will end after 1 standard number search;
- 3 will end after 1 standard number and 1 additional author/title search; and
- 2 will end after 1 standard number and 2 additional author/title searches.

These proportions were based on limited sampling, and would need to be modified if a particular kind of material, such as older materials lacking standard numbers, were expected to be a significant part of the workflow. These proportions are also an acknowledged but necessary simplification of the options encountered in the searching process. Such simplifications were considered necessary to provide a basis for estimating the number of searches generated by a given number of items so that that could become a factor in the per-item cost of searching.

In calculating staff costs, the task group assumed:

- The average time for each search transaction is one minute from the point a search is entered until a record is selected and displayed for review. The average transaction time for copying a record to the Aleph server was twelve seconds. Based on the limited sampling done, these time figures were true for all three record sources.
- The average cost of staff time spent on obtaining a record is estimated at \$15.00 per hour, or \$0.25 per minute.
- Multiple searches incur a higher per-item cost; hence the approximation of the number of searches needed per item discussed previously must be a factor in calculating staff time as well.
- Occasional exceptional delays in response time from Aleph or the source systems were not considered. None of the sources has been found to be delay-prone.

The task group formulated the study's derived per-item cost figures for searching OCLC, RLIN, and LC by combining figures for the three sources' charges for different kinds of searches, the estimate of the vary-

ing number of each kind of search required, and the staff cost component of each search, and then dividing by the number of searches. The per-item cost factors for using Aleph's Z39.50 client are shown in table 2.

Per-Search Cost Averages

The per-item costs were then combined with the sample set data to yield average costs per search when different assumptions are applied about the order in which the three sources are searched. Because each item had been searched in all three sources, the task group could determine for each sequence, beginning with LC, how many records would be found in the first source, in the second source, and in the third source. Searches beginning with OCLC or RLIN were assumed to include all the records available from LC, making a separate search of the LC file superfluous.

The average per record costs for each of the four search sequences are found in table 3.

Based on the study's data and estimates, the fourth sequence, starting with LC, looking next at RLIN, and last at OCLC, was found to be the most cost effective by a small margin. A much clearer margin separated the average costs of searching OCLC or RLIN as the first source from either of the sequences beginning with LC. The goal of the study was to recommend to staff a preferred searching sequence for the three available sources. The data clearly supported recommending that LC's database be searched first. Though the preference of RLIN over OCLC is not as well-supported by the small margin of difference between them in the study, the need to give staff clear direction weighed in favor of recommending the LC/RLIN/OCLC order.

Table 2. Per-item cost figures for searching bibliographic databases using Z39.50

Database	Cost per item (\$)
OCLC	1.825
RLIN	1.428
LC	0.425

Note: The formula used for these and other cost per-item estimates in the original report is: .5 (one standard number search cost + staff cost for 1 search) + .3 (standard number search cost + title search cost + staff cost for 2 searches) + .2 (standard number search cost + 2 title search costs + staff cost for 3 searches) = average cost per item searched.

Table 3. Average per record cost by search sequence

Search sequence	Cost per record (\$)
OCLC/RLIN	3.46
RLIN/OCLC	3.16
LC/OCLC/RLIN	2.57
LC/RLIN/OCLC	2.53

Discussion

The advantages of searching LC first as a source of records are clear in the study's findings, despite the fact that LC's database is smaller than either OCLC or RLIN. If the proportion of records found in LC's database had been significantly smaller, this would not have been the case. The staff cost of fruitless searching would have outweighed the benefits of finding a few records at lower expense. However, the study's sample data indicate that LC is actually a good source for a substantial portion of the items for which UML needs records. Therefore, the cost savings realized by conducting an initial search in LC more than offset the costs of failed searches and of repeating searches in OCLC or RLIN for unfound items. This result may seem counterintuitive, as it leads to more failed searches and more staff time expended than would be

required if the richest source, OCLC, were searched first. However, the costs of searches in the two utilities, especially of Z39.50 searches in OCLC, are so high that the savings realized by searching LC first and reducing the number of OCLC searches required more than covers the cost of the additional staff time spent.

Similarly, RLIN was recommended as a preferred second source, despite having an overall lower hit rate than OCLC for the sample set. The significant difference in the per-search charges for Z39.50 searches set by the two utilities when factored against the sample data indicated that marginal savings could be realized by seeking records first from RLIN when comparable records might also have been obtained from OCLC. This difference was much less decisive than that which determined that LC should be the first source searched; but as the study's practical goal was to guide UML library staff in selecting a default order for searching the three sources, the task group recommended the order indicated. Cost efficiency was greatest when the source with the highest per-search charge, OCLC, was used the least.

Conclusion and Suggestions for Further Study

Based on the study's findings, UML implemented a policy of searching for copy first in LC, then in RLIN, and last in OCLC. This searching strategy enabled the libraries to significantly reduce their yearly costs for cataloging searches in OCLC and RLIN in the fiscal year 2003–2004, where estimates prior to the study had been for increased costs. The libraries have continued to adhere to the searching strategy, and in the fiscal year 2004–2005 were able to maintain lower levels of utilities costs for technical services while significantly increasing overall cataloging production (56,918

titles cataloged in 2003–2004, compared to 73,529 in 2004–2005).

UML upgraded to Aleph version 16 in early 2005. A major change in this version of the software was the elimination of the separate searching module. Full search functionality, including Z39.50 searching and importing of records, is now available in all the modules, including Acquisitions and Cataloging. The basic procedures remain the same, and this change has had no noticeable impact on the efficiency of searching and importing records.

The calculations required for this study are complex and based on a number of estimates and limited data samples. One would hesitate to claim that the numbers the task group reports are accurate cost measures in every detail; nevertheless, the task group is satisfied that the estimates are sound approximations of highly elusive figures, and that their practical value as comparative measures for guiding library decision-making will stand up to scrutiny. That scrutiny is strongly encouraged for any library intending to make use of the group's findings. Libraries should also consider factors that may not have been addressed in the UML study. For example, if maximizing the occurrence of other system numbers—for example, the OCLC system number—is important for other processes the library is engaged in, that would need to be factored into the equation.

The amount of data collected on hit rates for particular types of materials was not large enough to draw any conclusions regarding adjustments to the preferred search order for those materials. Since the initial study, additional data on hit rates for serials has been collected, and the same could be done for other formats, languages, subject areas, or older materials. The task group recognizes that this would be a useful area for further study.

Other libraries seeking to determine the most cost-effective and effi-

cient sources for copy may wish to replicate this study for their individual environments, noting the cautions identified in the previous paragraphs. The authors encourage others to explore similar projects and to report on their findings.

References

1. Dilys E. Morris, "Staff Time and Costs for Cataloging," *Library Resources & Technical Services* 36, no. 1 (1992): 79–95; Lori L. Osmus and Dilys E. Morris, "Serials Cataloging Time and Costs: Results of an Ongoing Study at Iowa State University," *Serials Librarian* 22, no. 1/2 (1992): 235–48; Dilys E. Morris, Pamela Rebarcak, and Gordon Rowley, "Monographs Acquisitions: Staffing Costs and the Impact of Automation," *Library Resources & Technical Services* 40, no. 4 (1996): 301–18; Dilys E. Morris et al., "Cataloging Staff Costs Revisited," *Library Resources & Technical Services* 44, no. 2 (2000): 70–83; David C. Fowler and Janet Arcand, "Monographs Acquisitions Time and Cost Studies: The Next Generation," *Library Resources & Technical Services* 47, no. 3 (2003): 109–24; and David C. Fowler and Janet Arcand, "A Serials Acquisitions Cost Study: Presenting a Case for Standard Serials Acquisitions Data Elements," *Library Resources & Technical Services* 49, no. 2 (2005): 107–22.
2. Morris, "Staff Time and Costs for Cataloging," 85.
3. Morris, Rebarcak, and Rowley, "Monographs Acquisitions," 312.
4. Morris et al., "Cataloging Staff Costs Revisited," 74.
5. Ann Allan, "Chasing MARC: Searching in Bibliofile, Dialog, OCLC, and RLIN," *Journal of Academic Librarianship* 15, no. 6 (1990): 339–43.
6. Ibid., 341.
7. Mark L. Grover, "Cooperative Cataloging of Latin-American Books: The Unfulfilled Promise," *Library Resources & Technical Services* 35, no. 4 (1991): 406–15.
8. Ibid., 409.
9. Ibid., 411.

10. Christine Erbolato-Ramsey and Mark L. Grover, "Spanish and Portuguese Online Cataloging: Where Do You Start from Scratch?" *Cataloging & Classification Quarterly* 19, no. 1 (1994): 75–87.
11. Ibid., 80.
12. Ibid., 81.
13. Jamshid Beheshti, Andrew Large, and Pat Riva, "Cost Savings to Canadian University and Large Urban Public Libraries from Their Use of National Library of Canada MARC Records," *Library Resources & Technical Services* 47, no. 2 (2003): 44–57.
14. Ibid., 55.
15. Ibid., 53.
16. Beacher J. Wiggins, "Library of Congress Benefits from BIBCO," *RLG Focus* 47 (2000), www.rlg.org/r-focus/i47library.html (accessed July 23, 2004).
17. Ibid.
18. University of Minnesota Technical Services, "Source of Record: Final Report," Sept. 8, 2003. The report can be downloaded from the UML Technical Services Web site, <http://staff.lib.umn.edu/mac/sourcerec/index.phtml> (accessed Aug. 18, 2005).

Appendix: Definitions of Level of Copy Found by Bibliographic Searchers

Copy Level	Definition
No copy	No matching record is found. Even if a similar or related edition is found, the piece you are searching for differs bibliographically and a new record would need to be created in the local catalog.
Copy needs work	A minimal record is available, containing at least a matching 245, 260, and 300 field, but for the collection for which you are cataloging, the record would need additional authority, classification, and/or subject work.
Full copy	<p>A record is available that would need no additional cataloging work before the item is placed on the shelf, as defined by the criteria of the specific collection.</p> <p>For example, periodicals are classed in some collections at UML and not others. If an otherwise full record were found that lacked a call number, it would be considered "Full copy" for an unclassified periodical collection, but "Copy needs work" for a classed collection.</p>
Full copy (student)	Simple definition for student workers: Full copy has matching 245, 260, and 300 fields, at least one subject heading (6xx field), and an LC call number (050 field).