

Library Resources & Technical Services

ISSN 0024-2527
April 2006
Volume 50, No. 2

From the Ubiquitous to the Nonexistent
Jay H. Bernstein

Perpetual Access to Electronic Journals
Jim Stemper and Susan Barribeau

Library of Congress Classification Numbers
Bhagirathi Subrahmanyam

Utilizing Z39.50 to Obtain Bibliographic Copy
Christine DeZelar-Tiedman,
Cecila Genereux, and Stephen Hearn

If You Buy It, Will They Use It?
Anna M. Ferris



Library of Congress
cvr 2
p/u Jan '06 cvr4

Library Resources & Technical Services (ISSN 0024-2527) is published quarterly by the American Library Association, 50 E. Huron St., Chicago, IL 60611. It is the official publication of the Association for Library Collections & Technical Services, a division of the American Library Association. Subscription price: to members of the Association for Library Collections & Technical Services, \$27.50 per year, included in the membership dues; to nonmembers, \$75 per year in U.S., Canada, and Mexico, and \$85 per year in other foreign countries. Single copies, \$25. Periodical postage paid at Chicago, IL, and at additional mailing offices. POSTMASTER: Send address changes to *Library Resources & Technical Services*, 50 E. Huron St., Chicago, IL 60611. Business Manager: Charles Wilt, Executive Director, Association for Library Collections & Technical Services, a division of the American Library Association. Send manuscripts to the Editorial Office: Peggy Johnson, Editor, *Library Resources & Technical Services*, University of Minnesota Libraries, 499 Wilson Library, 309 19th Ave. So., Minneapolis, MN 55455; (612) 624-2312; fax: (612) 626-9353; e-mail: m-john@umn.edu. Advertising: ACLTS, 50 E. Huron St., Chicago, IL 60611; 312-280-5038; fax: 312-280-5032. ALA Production Services: Troy D. Linker, Angela Hanshaw; Christopher Keech, Stephanie Kuenn, Kristen McKulski, and Christine Velez. Members: Address changes and inquiries should be sent to Membership Department—*Library Resources & Technical Services*, 50 E. Huron St., Chicago, IL 60611. Nonmember subscribers: Subscriptions, orders, changes of address, and inquiries should be sent to *Library Resources & Technical Services*, Subscription Department, American Library Association, 50 E. Huron St., Chicago, IL 60611; 1-800-545-2433; fax: (312) 944-2641; subscription@ala.org.

Library Resources & Technical Services is indexed in *Library Literature*, *Library & Information Science Abstracts*, *Current Index to Journals in Education*, *Science Citation Index*, and *Information Science Abstracts*. Contents are listed in CALL (Current American—Library Literature). Its reviews are included in *Book Review Digest*, *Book Review Index*, and *Review of Reviews*.

Instructions for authors appear on the *Library Resources & Technical Services* Web page at www.ala.org/alcts/lrts. Copies of books for review should be addressed to Edward Swanson, Book Review Editor, *Library Resources & Technical Services*, 1065 Portland Ave., Saint Paul, MN 55104; e-mail: swans152@umn.edu.

©2006 American Library Association

All materials in this journal subject to copyright by the American Library Association may be photocopied for the noncommercial purpose of scientific or educational advancement granted by Sections 107 and 108 of the Copyright Revision Act of 1976. For other reprinting, photocopying, or translating, address requests to the ALA Office of Rights and Permissions, 50 E. Huron St., Chicago, IL 60611.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992. ∞

Publication in *Library Resources & Technical Services* does not imply official endorsement by the Association for Library Collections & Technical Services nor by ALA, and the assumption of editorial responsibility is not to be construed as endorsement of the opinions expressed by the editor or individual contributors.

Library Resources & Technical Services

ISSN 0024-2527

April 2006

Volume 50, No. 2

Editorial

78

ARTICLES

From the Ubiquitous to the Nonexistent

79

A Demographic Study of OCLC WorldCat
Jay H. Bernstein

Perpetual Access to Electronic Journals

91

A Survey of One Academic Research Library's Licenses
Jim Stemper and Susan Barribeau

Library of Congress Classification Numbers

110

Issues of Consistency and Their Implications for Union Catalogs
Bhagirathi Subrahmanyam

NOTES ON OPERATIONS

Utilizing Z39.50 to Obtain Bibliographic Copy

120

A Cost-containment Study
Christine DeZelar-Tiedman, Cecilia Genereux, and Stephen Hearn

If You Buy It, Will They Use It?

129

A Case Study on the Use of *Classification Web*
Anna M. Ferris

FEATURES

Book Reviews

138

Edward Swanson, Editor

Index to Advertisers

119

ABOUT THE COVER

Cover photo by Angela Hanshaw shows antique tiles displayed in the American Library Association's lobby.

Association for Library Collections & Technical Services
Visit *LRTS* online at www.ala.org/alcts/lrts.

For current news and reports on ALCTS activities, see the *ALCTS Newsletter Online* at www.ala.org/alcts/alcts_news.

EDITORIAL BOARD

Editor and Chair
Peggy Johnson
University of Minnesota

Members

Christina Bellinger, *University of New Hampshire*

Tschera Harkness Connell, *Ohio State University*

Trisha L. Davis, *Ohio State University*

Magda A. El-Sherbini, *Ohio State University*

Karla L. Hahn, *Association of Research Libraries*

Sara C. Heitshu, *University of Arizona*

Sue Kellerman, *Pennsylvania State University*

Shirley J. Lincicum, *Western Oregon University*

Bonnie MacEwan, *Auburn University*

Norman S. Medeiros, *Haverford College*

Carolynne Myall, *Eastern Washington University*

Linda L. Phillips, *University of Tennessee*

Pat Riva, *McGill University*

James Stickman, *University of Washington*

Diane Vizine-Goetz, *OCLC, Inc.*

(Intern) Caroll Nelson Davis,
Library of Congress

Ex-Officio Members

Charles Wilt, Executive Director,
ALCTS

Mary Beth Weber, *Rutgers University*, Editor, *ALCTS Newsletter Online*

Edward Swanson, *MINITEX Library Information Network*, Book Review Editor, *LRTS*

Editorial

Peggy Johnson



Thomas Frey, the executive director and senior futurist at the DaVinci Institute, has identified ten trends that he believes are affecting the development of libraries.¹ I share a synopsis with you to encourage thought about how these trends will likely influence what you do in your work and how you plan for the future. Some of the trends Frey identifies might serve as starting points for thoughtful papers about the future of technical services. Does anything here speak to you?

1. Communication systems are continually changing the way people access information.
2. All technology ends. All technologies commonly used today will be replaced by something new. Frey suggests that writing and books are technologies with limited life spans.
3. We have not yet reached the ultimate small particle for storage—but will soon. Standards for information storage will become more critical as storage particles become smaller, and the most critical component of stabilizing information storage will surround the issues of findability.
4. Search technology will become increasingly more complicated. As searching technology expands to include the ability to search for such attributes as taste, smell, texture, reflectivity, opacity, mass, density, tone, speed, and volume, the role of the librarian becomes more and more important.
5. Time compression is changing the lifestyle of library patrons. It is affecting nearly every aspect of our lives. As human need grows, the opportunities for libraries to meet these needs is also growing. The library of the future needs to be designed to accommodate the changing needs of its constituency.
6. We will be transitioning to a verbal society. Keyboards remain as the primary interface between people and electronic information, yet the days of the keyboard are numbered. There will be a strong trend toward verbal information.
7. The demand for global information is growing exponentially.
8. The stage is being set for a new era of global systems. Our ability to learn about and understand the cultures of the world is central to the global society of the future. Libraries will play a central role in developing global systems because libraries will be charged with archiving and disseminating the information necessary for the new systems to flourish.
9. We are transitioning from a product-based economy to an experience-based economy. Books will transition from a product to an experience. As books change from words on a page to digital manifestations of information, they will be reviewed and evaluated by the experience they create. How patrons experience libraries, as well as books, will become a key measurement criteria.
10. Libraries will transition from a center of information to a center of culture.

Reference

1. Thomas Frey, "The Future of Libraries: Beginning the Great Transformation," DaVinci Institute (2006), www.davinciinstitute.com/page.php?ID=120 (accessed Jan. 9, 2006).

From the Ubiquitous to the Nonexistent

A Demographic Study of OCLC WorldCat

Jay H. Bernstein

Analysis of a random sample of bibliographic records from OCLC WorldCat finds that the great majority of items in WorldCat are held by very few participating libraries, and that an inverse geometric relationship exists between the number of libraries holding an item and the number of items with a given level of shared holdings. The findings provide a context for interpreting holding levels in WorldCat with regard to the proportion of widely shared items and the characteristics of items at various ranges of holdings. Used with other quantitative and evaluative measures, these findings will assist libraries in assessing their collections.

OCLC WorldCat is arguably one of the most valuable tools available to librarians, as it provides online, global access to the shared records of thousands of libraries around the world. This enables librarians, at a minimum, to verify the existence of an item mentioned by a patron and to provide correct bibliographical facts of publication for it. More generally, it helps librarians identify and locate items that may not be held in their own collections.

WorldCat is also a potentially powerful research tool for collection analysis because each of its bibliographic records indicates the number of member libraries holding that item. These data are tantalizingly provocative to scholars interested in patterns in the distribution of books, knowledge, and information.

The theoretical basis for considering WorldCat holding levels in library research is that the number of libraries holding a given title provides a score that can measure that title's influence or impact. Several studies conducted over the years on specific library material domains—adult fiction, books from small publishers, scientific journals, and award-winning monographic titles in the social sciences and humanities—have cited figures on holdings in WorldCat or its predecessor, the OCLC Online Union Catalog.¹ Similarly, Mirwis has included WorldCat holdings as one of several factors used to rate and rank encyclopedias.²

The difficulty with citing these holdings data lies in knowing how to interpret them. Without a clear picture of what exactly is in WorldCat and what the spectrum of distribution is, the raw holdings numbers mean little. Only by examining a sample of items from WorldCat itself, rather than beginning with particular

Jay H. Bernstein (jbernstein@kingsborough.edu) is Assistant Professor and Reader Services Librarian in the Robert J. Kibbee Library, Kingsborough Community College, City University of New York.

The author thanks Jennifer Coggon, Edward Martin, Allan Mirwis, Roberta Pike, Elizabeth Tompkins, and Stephen E. Wiberley Jr. for their helpful comments on earlier drafts of this paper, and Anna H. Perrault for her advice and encouragement.

known items or a particular library collection, can one develop an interpretive framework that could maximize the analytical value of these data.

In the interest of providing a context for interpreting WorldCat holding levels, I undertook what might be called a demographic study of WorldCat, focusing on its composition and major characteristics as a population by analyzing a random sample of 500 records. The goal in investigating the items of WorldCat as representatives of a population was to determine the categories of materials that are represented in WorldCat and their proportions, the proportion of widely held items, and the characteristics of those items.

Precedents for this study may be found in publications by White and Perrault.³ White demonstrated the connection between Research Libraries Group (RLG) Conspectus levels and OCLC holding levels. He found that titles at the research level had holding levels up to 150, and that the midpoint of holdings of titles in library collections at all levels was about 400.⁴ Perrault studied a systematic random sample of nearly 3.4 million WorldCat records to gauge trends in library collection building over time and assess WorldCat's potential for promoting shared access to scarce resources.⁵ While her study did not concern levels of shared holding in a manner comparable to White's work, she did uncover a finding that is striking when considered alongside White's results: more than 53 percent of the records in Perrault's sample were held by only one member library.⁶

The present study attempts to show the big picture of what is in WorldCat by considering its widely shared items against the background of a much larger number of extremely scarce items. Like Perrault's study, it employs an inductive approach to assess WorldCat as a whole, though it is based on a sample far smaller than hers. Like White's work, it relates its findings to those items in WorldCat that are frequently found in library collections. This study finds that 39.8 percent of items in WorldCat are held by only one library. This is significantly less than the 53.1 percent found by Perrault, though it is still a very considerable fraction. The study also finds that only 9.1 percent of items are held by more than 50 participating libraries. Perhaps most surprising, the study shows that 1.2 percent of items in WorldCat are not held by *any* libraries.

OCLC WorldCat in a Nutshell

Provided by OCLC Online Computer Library Center, WorldCat is an online bibliographic database containing records created and maintained collectively by more than 9,000 member institutions (the exact number is unavailable). Described by OCLC as "the largest and most comprehensive database of its kind," its resources "span thousands of years and nearly every form of human expression. Records

exist for everything from stone tablets to electronic books, wax recordings to MP3s, DVDs and Web sites."⁷ A million records were added in the first three-and-a-half months of 2004, bringing the total inventory to 55 million. On average, WorldCat adds a new record every 12 seconds, a phenomenally rapid rate of growth.⁸ The database has doubled in size in just more than eleven years.⁹

Not long after launching its Online Union Catalog in 1971, OCLC set forth initiatives throughout the world promoting input from other countries that would make its database global. In doing so, OCLC had to incorporate international variations in MARC format, authority rules, record quality, and library traditions, not to mention non-Roman alphabets and characters.¹⁰

The many libraries around the world contributing to WorldCat make it an unsurpassed data bank for research on the contents of libraries. It is probably a good measure of library holdings, especially in the English-speaking world. However, this does not necessarily mean that it accurately reflects the totality of world library holdings, much less the entire store of recorded human knowledge.

Most institutions that contribute to WorldCat are governing members of OCLC, meaning that they contribute all current cataloging to an OCLC-affiliated database. Two lower grades of OCLC membership involve less-than-full participation in WorldCat. More than 77 percent of governing members are located in the United States. WorldCat seems to represent academic libraries more so than other kinds of library. According to data in OCLC's 2003 annual report, 31 percent of governing member libraries are college and university libraries, compared to 22.2 percent of governing members that are public libraries, 12.8 percent that are federal, state, or municipal government libraries, and 10.2 percent that are corporate or business libraries.¹¹ Other categories of libraries making up OCLC's governing membership are community college and vocational (8.3 percent), school (7.4 percent), associations and foundations (4.0 percent), state and national (1.0 percent), and other (3.0 percent).¹²

Data Collection Procedures

The first step in data collection was to select 500 random numbers from 1 to 54 million using an online random number generating service.¹³ The maximum number of 54 million was chosen because the database contained "54 million quality records and counting" at the time research for the study began in July 2004, according to OCLC's WorldCat home page.¹⁴ Each randomly chosen number obtained using this method was matched with the bibliographic record for the item with that OCLC record number. Given the size of the chosen population, the sample of 500 provides a confi-

dence level of 95 percent and a confidence interval of 4.4. Therefore, conclusions about this sample are expected to be accurate plus or minus 4.4 percent, with 95 percent certainty, as statements about the entire WorldCat database. Such a sample is adequate for drawing general conclusions about the major outlines and proportions of items in WorldCat. It is not sufficiently large enough to include many outstanding but extremely rare types of items, such as incunabula, paleographic writing boards, or items in hieroglyphics.

Findings

Types of Documents and Materials

In order to comprehend WorldCat as a database, one must sort out the various kinds of items it contains. The *Anglo-American Cataloguing Rules*, 2nd edition, 2002 revision, differentiates in its table of contents the following categories of items for cataloging: books, pamphlets, and printed sheets; cartographic materials; manuscripts (including manuscript collections); music [printed notated music]; sound recordings; motion pictures and videorecordings; graphic materials; electronic resources; three-dimensional artifacts and realia; microforms; continuing resources [serials]; and analysis [monographic or journal analytics].¹⁵

Such a synopsis is typical. However, it is inadequate as a categorical scheme, as the individual categories are not mutually exclusive. Books, for example, can be manifested not only on printed paper but also in sound recordings, microforms, and electronic resources. Indeed, most document types can appear in a number of material formats. To avoid confusion in collection analysis, a system that creates nonoverlapping categories by distinguishing explicitly between document type and material type must be established.

The use of the terms *document type* and *material type* is intended to emphasize a sharp conceptual distinction that is absent from most discussions of the categories of items in catalogs and bibliographic databases. Without this distinction, any discussion of the categories of items in libraries and catalogs is bound to be muddled.

Document type is primary, and material type is subsidiary. Document type sums up the general, overall character of a work in a basic, salient, overriding category. The major document types, such as book, serial, manuscript, sound recording, and motion picture, are sociocultural categories characterized by prototypical cores but fuzzy and ambiguous boundaries.¹⁶ Material type denotes not only the physical medium (e.g., paper, magnetic tape, celluloid film), but also the mode of communication—the medium in which the content is presented. Thus, printed language material, printed cartographic material, and printed musical material are distinct material types. Similarly, atlases and scores on microfilm are separate material type categories from

microfilm books. Material type differentiates among various specific microformats (e.g., microfilm, microfiche, microopaque); among formats for reproducing sound, visual, or graphical data; and among various electronic formats and means of access to computer files (Web, file transfer protocol, Usenet, diskette, CD-ROM, and so on.).

Analyzing the makeup of OCLC WorldCat in terms of document type and material type involves both the aggregation and separation of categories. Thirteen document types and 22 material types may be identified in the sample. Combining document types and material types, bibliographic records from the sample of 500 can be placed into 30 mutually exclusive categories, plus a null category consisting of 5 numbers matching no records. Items are by no means distributed evenly among categories, but are heavily clustered in a few categories, with a few scattered entries among all the other categories (see table 1). Printed language books (hereafter print books) alone account for more than two-thirds of all items.

For both document type and material type, the predominating type is responsible for more than 70 percent of records and is about ten times more prevalent than the second most common type. The 2 leading types in both categories are closely related to each other: the leading document

Table 1. Document types and material types in WorldCat sample (n=495)

Document type	No.	%
Books	364	73.5
Manuscripts (text)	40	8.1
Serials	25	5.1
Scores	15	3.0
Musical recordings	12	2.4
All others	39	7.9
Material type		
Printed language	355	71.7
Original language	36	7.3
Microfilm language	21	4.2
Printed notated music	14	2.8
Microfiche language	10	2.0
All others	59	11.9
Combined document: material type		
Books: Printed language	335	67.6
Manuscripts: Original language	33	6.7
Serials: Printed language	18	3.6
Books: Microfilm language	15	3.0
Scores: Printed notated music	14	2.8
All others	80	16.2

type is the book, and the leading material type is printed language material, while in second place are manuscript and original language material respectively. The same ratio holds for the 2 leading combined document-material types: there are ten times as many print books as original language material manuscripts. Together, these 2 kinds of items account for almost 75 percent of all items in the database.

Categories of Holdings Levels

This study aims to determine the proportion of items in the catalog that are widely held and to analyze the characteristics of these widely held items. Determining a cut-off point for high holdings to define a category of widely held items is an arbitrary procedure that can lead to circular reasoning. One cannot know in advance what a high level of holdings is. Because this paper is a first attempt to determine the occurrence of widely held items in context of the totality of merged online catalogs, predefining the category seems unwise. It is somewhat easier on an intuitive level to grasp the meaning of *scarce* items than *common* items (though here too the cut-off point is arbitrary). Therefore, rather than attempt to provide a parameter for widely held items, I will set these items aside for the time being by dividing the sample into four categories:

- Nonexistent: items with 0 holdings
- Unique: items with 1 holding
- Scarce: items with 2 to 50 holdings
- Non-scarce: items with more than 50 holdings

In this typology, non-scarce is a residual category and will be addressed later in this paper.

The reader may be surprised at these categories, particularly the first. However, the typology arises from the data themselves, so the distribution of holding levels itself is the real surprise:

- Nonexistent items: 6 (1.2 percent)
- Unique items: 197 (39.8 percent)
- Scarce items: 247 (49.9 percent)
- Non-scarce items: 45 (9.1 percent)

Non-scarce items account for a relatively small fraction of the sample, and this finding justifies not subdividing the non-scarce portion into thinner layers at the outset.

These distributions vary by both document type and material type. Table 2 shows the dis-

tribution for all items and for the top five document types, which make up more than 92 percent of the sample. Government publications, which account for 7.3 percent of the sample, break down in approximately the same proportions as other items.

Based on the sample, 1.2 percent of items cataloged in WorldCat have absolutely no holdings; they are not held by the libraries that had cataloged the items and input the records. These appear to be items that once were held but that were deaccessioned, as none of the records are based on prepublication data or unexamined material. However, the bibliographic records that should have been deleted or reported as errors persist, and the items must be considered nonexistent in terms of library collections. For 5 of the 6 items with 0 holdings, other records with holdings for items with the same title are present; all items are extremely scarce, except for one that is cataloged in the sample as a book but that is held as a microform serial by 17 libraries. In addition to the 6 records that have 0 holdings, 5 numbers used to obtain sample records do not match any records. The percentages given here are percentages of the sample of 495 records, not of all the 500 random numbers used to generate the sample, unless otherwise noted.

In this paper, *unique* means that only one library that is a participating member of OCLC owns the item. *Unique* does not necessarily mean that no other library in the world owns the item, much less that it is an absolutely unique document. In analyzing the percent of unique items, the differentiation of original or archival materials, which one would expect to be unique from other categories of materials is advantageous. Original and archival materials in the sample include language and music manuscripts, mixed material collections, and a graphic collection. Forty-four items in the sample are archival, and 41 of these are unique. All but one of the language manuscripts are academic theses, and the exception appears to be an incorrectly cataloged book. The great majority of theses are for degrees from United States institutions and are in English. Non-archival materials account for 451 items in the sample and, of these, 156 (34.6 percent) are unique. Archival materials account for only 8.9 percent of all materials in the sample but 20.9 percent of the

Table 2. Distribution of holding levels categories by document type (n=495)

Document type	Nonexistent	Unique	Scarce	Non-scarce	Total
Books	4	118	200	42	364
Manuscripts	0	36	4	0	40
Serials	0	11	13	1	25
Scores	0	7	8	0	15
Musical recordings	1	3	7	1	12
All others	1	22	15	1	39
Total	6	197	247	45	495

unique content. Besides archival materials, the category of unique items includes many of the items outside the print book category as well as many technical reports and items of only local significance.

Table 3 shows the distribution of frequency by material type categories. It contrasts print language material, which is the predominant category, making up almost three-quarters of the total, against all archival materials and the residual category of all other materials, which includes microforms, computer files, maps, sound recordings, and so on.

Unique items are more likely than items in other holding level categories to be short in length. Of unique print books, 36.2 percent have fewer than 50 pages (or leaves), and 40.9 percent are more than 100 pages long. This compares to 27.2 percent and 56.1 percent, respectively, in all holding level categories. Table 4 shows the distribution of print books in each holding level category according to page length.

In a study of nearly 3.4 million items in WorldCat, Perrault found that 53.1 percent of all records for *monographs* were unique.¹⁷ Presumably, Perrault's *monographs* are identical to my *print books*, though the term also could refer to anything not cataloged as a serial. Only 39.8 percent of all items in the present research sample, including 31.3 percent of print books, are unique.

Scarce is defined as having greater than 1 but no more than 50 holdings; 49.9 percent of items are in the scarce category. Although a few items cataloged as archival materials have multiple holdings in very low numbers and thus may be considered scarce, most scarce items are non-archival.

Holding levels are strongly correlated with the presence of call numbers. One hundred fifty-eight records

(31.9 percent of the total) have no call numbers of any kind, but only 2 are non-scarce. Of unique items, 46.2 percent lack call numbers (very close to the 47.6 percent of unique monographic titles with no call number found by Perrault), and for scarce items, the figure is 25.5 percent.¹⁸ Half of all items have a Library of Congress Classification (LCC) number assigned either locally or by the Library of Congress (LC). Adding National Library of Canada numbers, 52.5 percent of items have at least one LCC-type call number. Dewey Decimal Classification (DDC) numbers, either LC-assigned or locally assigned, are present in 30.7 percent of items. For all items, the ratio of LC-assigned to locally assigned LCC numbers is approximately 1 : 1, but for DDC numbers the ratio is greater than 2 : 1 (see table 5). For unique items, however, locally assigned numbers greatly outnumber centrally assigned numbers in both classification systems.

Whereas scarce items frequently have incomplete cataloging, the records for non-scarce items tend to be the fullest, most complete records. They are far more likely than others to have been cataloged at the national level, meaning they have been both cataloged and transcribed by national bibliographic agencies, as indicated in the MARC 040 field (cataloging source) by codes for national cataloging agencies—the Library of Congress (DLC), the British Library (UKM), the National Library of Canada (NLC), the National Library of Medicine (NLM), and others. For example, an item cataloged and transcribed by LC would contain the code “DLC ‡c DLC” in the MARC 040 field. Compared to 18.6 percent of scarce titles and just 7.7 percent of unique titles, 58.7 percent of non-scarce titles are so cataloged. For titles with holdings of more than 100, the proportion

is 81.8 percent. Excluding archival materials, 20 percent of records are cataloged and transcribed by national bibliographic agencies. Print books account for 90.9 percent of items cataloged at the national level. Table 5 shows that call numbers for non-scarce items are far more often created by LC than locally; 78 percent of items have LC-created numbers (LCC or DDC), as compared to 9 percent with locally created numbers. Non-scarce items often have specialized call numbers in addition to ordinary call numbers. Reflecting the dominance of Anglo-American institutions in WorldCat participation, Universal Decimal

Table 3. Distribution of holding levels categories by material type (n=495)

Material type	Nonexistent	Unique	Scarce	Non-scarce	Total
Printed language	4	118	197	40	359
Archival	0	41	3	0	44
Other	2	38	47	5	92
Total	6	197	247	45	495

Table 4. Length in pages of print books categorized by holdings level category

Holdings level	1-50 pages	51-100 pages	More than 100 pages	Unknown page length	Total
Nonexistent	2	0	1	1	4
Unique	38	14	43	10	105
Scarce	50	21	109	7	187
Non-scarce	1	2	35	1	39
Total	91	37	188	19	335

Table 5. Distribution of types of call numbers in frequency categories

	No call number		050 (LCC)	055 (Can. LCC)	060 (NLM)	070 (NAL)	080 (UDC)	082 (DDC)	084 (other)	086 (Gov. doc.)	090 (local LCC)	092 (local DDC)
	No.	%										
Nonexistent	2	33	0	0	0	0	0	0	0	1	2	1
Unique	91	46	18	2	3	0	1	8	3	1	57	18
Scarce	63	26	70	9	3	0	0	63	0	7	64	23
Non-scarce	2	4	35	1	3	2	0	35	0	4	4	4

Note: Numbers in top row refer to MARC fields. Not all categories are mutually exclusive.

Classification (MARC 080 field) or “other” (MARC 084 field) numbers are found only in unique items.

Non-scarce items make up only 9.1 percent of the total; 93 percent of these are in the book document type, and 89 percent are print books. Besides print books, non-scarce items include 2 microfilm reproductions of original books dating from 1665 and 1894, a map, a musical recording on a compact disc, and an online journal. Considering that these document and material types occur far less frequently than print books, one cannot conclude from the low numbers that items of these kinds are scarce in greater proportions than are print books.

Countries and Regions of Origin

Given OCLC's stated objective of the global networking of information, United States dominance as the place of origin of items in WorldCat (despite the representation of many countries in all regions of the world) is noteworthy.¹⁹ Items originate in 54 countries, an impressive number for a sample this size. Twenty-four countries are represented by one item apiece, and 19 countries by between 2 and 5 items. The United States alone is responsible for 195 items, or 39.4 percent.

In second place, with 60 items, or 12.1 percent, are materials with no place or an unknown or undetermined place of origin, as indicated by “Ctry: xx” in the fixed field of the MARC record. Most of these are archival or unpublished materials, especially manuscripts, which as a rule do not name a country of origin. However, other items for which the place of publication or origin is not recorded or cannot be determined are coded the same way. Apart from manuscripts (language and music), 4.2 percent of records do not name a country of origin.

Limiting the materials to print books, 52 countries are represented in all (see figure 1). Twenty-three of these are responsible for only 1 book each, and another 16 are responsible for between 2 and 5 books each. The United States, England, and 8 other countries (mostly in Western Europe) are responsible for more than three-quarters of all

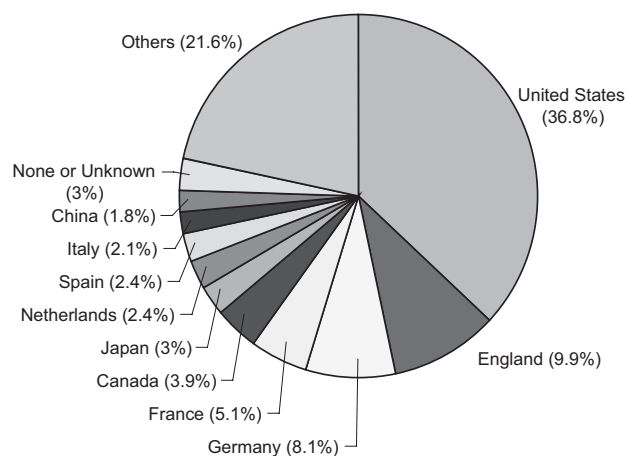


Figure 1. Distribution of print books in WorldCat sample by country of publication (n=335)

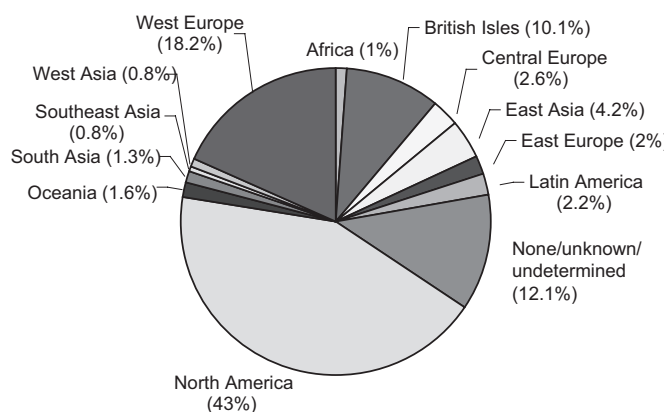


Figure 2. Distribution of items in WorldCat sample by region of origin (n=495)

print books in the sample. In the absence of manuscripts, archival materials, and computer files, only 3.0 percent have no known or determined country of origin, as compared to 30.6 percent in all other categories.

Looking at WorldCat in terms of regional representation (see figure 2), North America (defined as the United States and Canada), Western Europe, and the British Isles (including the Republic of Ireland) are dominant, while Africa and all of Asia except for the Far East are only very slightly represented.

Microfilm and other reproductions usually state the country of origin of the original document, not the copy, at least in the fixed field of the MARC record, which is the preferred source of information on countries for this study. Bearing in mind this proviso, 31 of 45 non-scarce items (69 percent) originated in the United States, followed by eight (18 percent) from England; the remaining items (1 each) originated in Canada, France, Germany, Japan, Spain, and Scotland. All nonexistent items originated the United States.

Languages

Official information on the number of items in various languages in WorldCat, as of July 30, 2004, is available on OCLC's Internet page, "WorldCat facts and statistics."²⁰ Based on figures for the total number of items in the database presented therein, one can calculate the fraction of items in each language. Comparison of these statistics to those for the top 9 languages in the research sample (see table 6) lends assurance that the sample is representative of the whole, though noticeable differences exist concerning the proportions of items in French and Japanese. The reader is referred to the Web site for figures for the top 53 languages (Serbo-Croatian is counted twice, differentiating between Romanic and Cyrillic forms). The sample of 495 includes items in 31 languages (including both forms of Serbo-Croatian) and 2 items in multiple languages. Three percent of the sample, chiefly musical scores and recordings, are nonlingual. Of the 27 items in those categories, 13 (48 percent) have no language.

The last column of table 6 provides percentages for print books. It shows that the percentage of print books in English is significantly smaller than the percentage of all items in English while the percentages in French and German are considerably higher for print books than for all formats. This appears to be the effect of the large number of records contributed by academic libraries in the United States and other Anglophone countries for academic theses that are overwhelmingly in English and excluded from the print book category as these are cataloged as manuscripts.

Intensity of language representation and holding levels are strongly correlated, with the most frequently occurring languages having the highest holdings. The 305 English-language items can be divided into the frequency categories as follows: 6 (2 percent) nonexistent, 127 (42 percent) unique, 130 (43 percent) scarce, and 42 (14 percent) non-scarce. By comparison, of 190 items not in English,

69 (36 percent) are unique, 117 (62 percent) are scarce, and only 4 (2 percent) are non-scarce, including titles in French, German, and Spanish, as well as one nonlingual musical recording.

Material Age

To analyze the age structure of WorldCat, all serial or continuing items as well as undated items are removed from the sample, resulting in 460 items. The sample is divided by century, further breaking down the twentieth century into time periods for the first half, third quarter, and fourth quarter. Separate categories are created for items that have been reproduced in periods different from the original period of publication. The results appear in table 7. The period of 1976 through 2000 is the largest category, with 208 items, which is 45.2 percent of dated, non-serial and non-continuing items in the sample. This time period accounts for 69.6 percent of non-scarce items.

Holdings Distribution

The simplest description of the holdings distribution of sample records is that the greater the number of holdings,

Table 6. Representation of languages in WorldCat and in research sample by percentage

Language	OCLC statistics (%)	Research sample (%)	Print books (%)
English	61.2	61.6	56.7
French	6.4	7.5	8.7
German	6.3	5.7	8.4
Spanish	4.6	4.2	4.2
Japanese	2.7	1.8	2.7
Chinese	2.3	2.0	3.0
Russian	1.9	1.6	1.8
Italian	1.8	1.4	1.8
Latin	0.9	1.2	1.2

Table 7. Date of publication of non-serial and non-continuing items (n=460)

Year of Publication	Originals	Reproductions
1401–1600	0	3
1601–1700	2	1
1701–1800	7	1
1801–1900	37	7
1901–1950	54	6
1951–1975	104	2
1976–2000	208	0
2001–2003	28	0

the fewer the number of items having that many holdings. There are many records with low holdings, with the greatest number having a single holding, and decreasing numbers of records with higher numbers of holdings, followed by a few records with a large number of holdings. The first part of the curve can be seen as a reverse J-shaped curve with a long tail, as shown in figure 3, which provides the frequency of records with 1 to 40 holdings. But to say there is a long tail is insufficient, as this does not begin to suggest the relatively small but significant number of records that have much higher levels of holding.

Although any cut-off point is arbitrary, one may visualize the sample in two sections, as in figures 4 and 5. The first group is the large number of items with few holdings, and the second is the group with few members, but with elevated numbers of holdings. Two-thirds of all the records (328 out of 495) have 0 to 5 holdings, while the remaining one-third have more than 5 holdings. These categories are analyzed separately.

For the first group, one can apply a modified version of Lotka's law, which states that "the number of authors who have published a specific number of papers [is] approximately equal to the inverse square of that number multiplied by the number of authors who had published one paper only, that is, [if one sets] $f(y)$ as the number of authors publishing y papers, . . . [then] $f(y) \approx 1/y^2 \times f(1)$."²¹ One can adapt this law by substituting records for authors and hold-

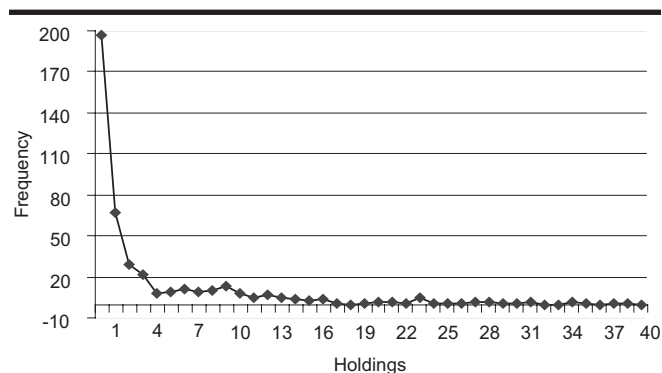


Figure 3. Frequency of records with 0 to 40 holdings (n=445)

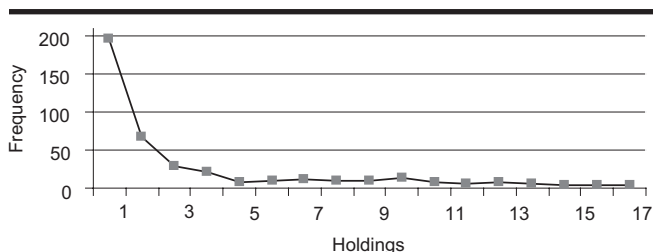


Figure 4. Frequency of records with 0 to 17 holdings (n=417)

ings for papers. Given the size of the sample, this formula provides a reasonably close approximation for the fraction of items with 2 to 5 holdings (see figure 6).

For the second group, with elevated numbers of holdings, a different form of analysis is necessary. Here, one may divide the group into zones containing equal numbers of records to see what the range of holding levels are for each zone. Keeping in mind the geometric growth seen in the curve in figure 5, and following Price's law, which states that "the number of prolific authors in a subject area (i.e., producing about half the publications in the field) [is] approximately equal to the square root of the total number of authors," a zone is defined as a group containing a number of items equal to the square root of the entire sample (500), which is slightly more than 22.²² Not all zones contain exactly 22 items, as the size is adjusted to fit all items with the same number of holdings. In all, there are 8 zones in all of the items with holdings of more than four, as seen in table 8. Unexpectedly, zone 1, the highest zone, defined as the zone with the 22 items with the highest level of holdings, consists of items (all print books) with more than 100 hold-

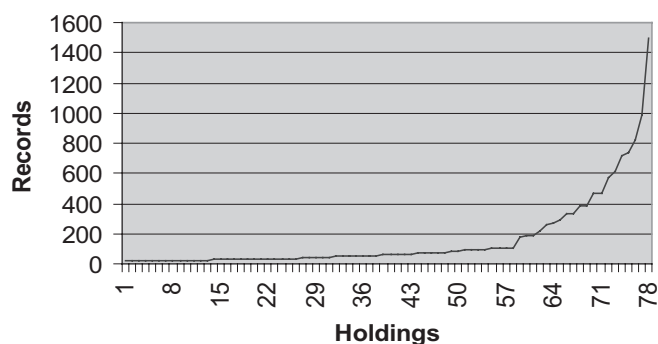


Figure 5. Occurrence of records with 18 or more holdings (n=78)

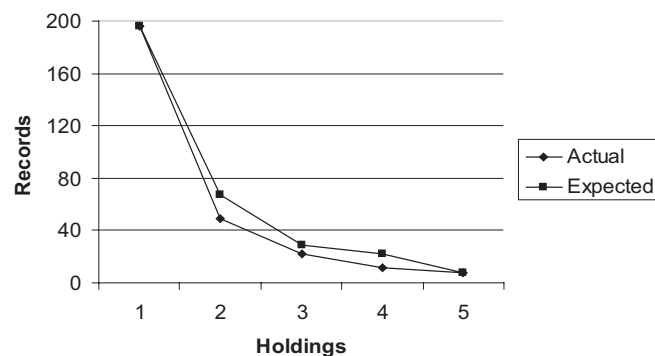


Figure 6. Distribution of items with 1 to 5 holdings (n=287)

ings. Zone 2 contains items with 52 to 100 holdings; zone 3 contains items with 25 to 51 holdings; and so on. In all zones, the majority of items are print books.

Table 8 shows the range of holdings levels in each zone, a range that becomes wider as the number of items increases. This analysis of zones helps segment the large and heretofore undifferentiated category of scarce items.

The number of widely shared items in OCLC WorldCat as a whole cannot be estimated from the square root of total number of items in the database. Given a database size of 54 million, this would mean that only 7,550 items have more than 100 holdings, an absurdly small number. On the contrary, the fraction of items in the sample with more than 100 holdings, 4.4 percent, is a much more reliable figure for projecting the number of items with more than 100. Using this percentage, one can estimate 2,376,000 records with more than 100 in a database of 54 million. The modified equivalent of Price's law is used simply to suggest that the number of items with more than 100 holdings is approximately the same as the number with 5 to 6 holdings, 7 to 8 holdings, and so on, up to about 52 to 100 holdings. The highest zone is limited only by the number of participating libraries; the second-highest zone has an interval of 49, and the third-highest zone has an interval of 27. As zones encompass materials with higher holdings, intervals of ranges grow exponentially. Excluding the first zone, whose wide range dwarfs all other zones, figure 7 graphs the interval sizes for 8 to 2.

Characteristics of Widely Held Titles

Forty-five items in the sample (9.1 percent) are non-scarce, with more than 50 holdings each. However, the previous analysis shows that this category constitutes quite neatly the top two zones, with a residue of just 1 item in the third zone. At this point one can focus on the highest zone, which consists entirely of print books in English, dividing it even further into two equal halves to arrive at a core consisting of the most ubiquitous items. Dividing the top zone into two halves of 11 items each results in group A, with 381 to 1491 holdings, followed by group B, with 105 to 335 holdings. The size of the sample does not enable pinpointing with precision the lower boundary of this highest fraction, and one ends up with a gap of 46 between the two sub-zones. Given a margin of error of 4.4 percent, one should probably say only that highly ubiquitous holdings begin at about 360, plus or minus 40. With 11 items, group A makes up 2.2 percent of the sample. Remarkably, this is the same percentage that has either no matching record or a record with no attached holdings. Even more striking is that White, in a study of library collections, found the midpoint for library holdings to be about 400; furthermore, he asserts that this figure of 400 library holdings represents the dividing line

Table 8. Zones of holding levels

Zone	Range	Interval	Total	Print books	Other formats
8	5–6	2	17	10	7
7	7–8	2	17	15	2
6	9–10	2	23	16	7
5	11–14	4	25	20	5
4	15–25	11	22	15	7
3	25–51	27	22	19	3
2	52–100	49	22	17	5
1	101–1,491	1,391	22	22	0

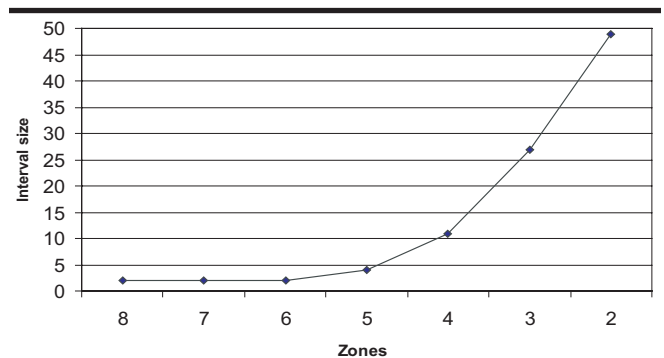


Figure 7. Increasing interval sizes of zones 8 to 2

between highly specialized books and books aimed at more general readership.²³ This suggests that fewer than 5 percent of items in WorldCat account for the great majority of library holdings.

A look at the items in groups A and B supports White's statement. All items in group A are published in the United States. Five are works of fiction. One is a translation. Publication dates range from 1954 to 2003. Publishers include industry giants—Warner Books, Houghton Mifflin, Viking, Basic Books, and St. Martin's (with two entries in this group)—along with a major scholarly publisher, Princeton University Press, as well as Orchard Books, a division of the Watts Publishing Group. The occurrence of two less-prominent publishers of group A books, the Center for Urban Policy Research at Rutgers University and John F. Blair (located in Winston-Salem, North Carolina), suggests that many relatively small publishers have at least some titles with high levels of holdings.

Most group B books seem to be located in a different sector of the book industry than group A books. Publication dates range from 1972 to 2002. Three of the 11 titles are published in England. This group includes 1 title each in the fiction, juvenile, and government publication

categories. In order of descending ubiquity, publishers of group B titles include Routledge, W. B. Saunders, Prentice-Hall, World Publishers, Kar-Ben Copies, Scholar Press/Ashgate, CRC Press, Overlook Press, Oneworld, Paulist Press, and the United States Government Printing Office. Most of these publishers are not in the trade or mass-market sectors of the publishing industry, but are dedicated to educational and scientific subjects.

Changes in Holdings over Time

By marking changes in the holdings of WorldCat records after a lapse of time, one can detect activity in the database. Ideally, one would like to know whether items circulated or were consulted, but WorldCat does not provide that information. Items whose holding levels have changed (upwards or downwards) over time are active; the others are static. As one might expect, items with the most holdings are the most active, while those with the fewest holdings are the least active. Surprisingly, however, items at all holding levels except nonexistent experience change over time.

Five months after first obtaining data on holding levels, each record in the sample was checked. Forty-six items (9.3 percent of the sample) had added holdings between July and December of 2004, while 24 (4.8 percent) had lost holdings during the same period of time. The largest number of holdings gained by any item was 91, while the largest number of holdings lost was 8.

Group A experienced the most marked change, with 5 items gaining and 5 losing. Only 1 item in this group did not change. In group B, 4 items gained and 3 lost. In the non-scarce category as a whole, 15 items gained in holdings, while 16 lost holdings. Therefore, 69 percent of non-scarce items changed in the absolute number of holdings in a five-month period. All but 3 of these items were print books.

Thirty items with scarce levels of holdings gained, while 8 lost. This means that 15.4 percent of scarce items changed in the number of libraries holding the item. Twenty-six of the scarce items that gained in holdings were print books. Only 2 unique items gained in holdings during the course of the study. While these items are exceptional, the study finds that activity is not exceptional at the level of 2 holdings. Indeed, 3 items that were not unique at the beginning of the study became unique by the end of the study by being deaccessioned from libraries that previously owned them. This study suggests that unique items are occasionally removed from holding libraries without deleting the records, resulting in nonexistent items.

Widely Shared Titles and Collecting Levels

This investigation finds that most items in WorldCat are rare and unusual items that do not have a place in most

library collections and are unavailable for acquisition by libraries other than those that already own them. To use an oceanographic metaphor, most known content in WorldCat consists of the small fraction of materials at or near the surface, while a far greater amount of content is unknown, lying undisturbed deep below the surface and toward the ocean floor. The majority of items in WorldCat are extremely scarce, with two-thirds held by no more than 5 participating libraries. Even excluding manuscript and other archival materials (which make up a considerable portion of WorldCat), nearly 75 percent of items are held by a maximum of ten participating libraries. Although extremely scarce items are sometimes added to library collections, only a small fraction of WorldCat items are widely distributed among libraries and play a role in selection, acquisition, and collection management.

White has suggested that OCLC (i.e., WorldCat) holdings data can be used as a guide for collection development by checking a sample of items against the number of holdings in WorldCat to determine those items' collecting levels in terms of the RLG Conspectus.²⁴ The Conspectus provides guidelines for the acquisition of library materials, differentiating between minimal, basic information; study or instructional support; research; and comprehensive levels. These levels refer to libraries' objectives for specific call number ranges and subject collections, not whole libraries, and it should be noted that some levels have gradations and qualifications.²⁵ White developed what he called rules of thumb for determining the collecting level in these terms of given items. He reckoned that books with fewer than 150 holdings were at the research level; those with between 150 and 400 holdings were at the instructional level; those with between 400 and 750 holdings were at the basic information level; and those with more than 750 holdings were at the minimal level.²⁶

From 2001 through 2002, Lesniaski reexamined the holding counts of items in White's sample lists, finding that the number of holdings for titles in each collecting level had risen significantly.²⁷ By adjusting each collecting level upward, he found that the approximate ratios between levels remained approximately constant. By Lesniaski's new rules of thumb, level 1 (minimal level) corresponds to WorldCat holdings above 1,000; level 2 (basic information level) corresponds to between 500 and 1,000 holdings; level 3 (study or instructional support level) corresponds to between 200 and 500 holdings; and level 4 (research level) corresponds to fewer than 200 holdings.²⁸ The average numbers of holdings of titles at these levels are 1,541, 751, 389, and 153 respectively, producing ratios of 10 : 4.9 : 2.5 : 1.²⁹

The cut-off point for group A (ubiquitous items) is near the average for the study or instructional support level. Because the RLG Conspectus allows for the division of this level into two sublevels, 3a (study or instructional

level, introductory) and 3b (study or instructional level, advanced), one might connect group B with collecting levels 3b and 4, and group A with levels 3a and below. The category of ubiquitous holdings, therefore, encompasses the Conspectus introductory instructional or study level, the basic level, and the minimal level.

One would like to know the relative proportions of items in WorldCat at the various collecting levels. Only 1 item in the sample has more than 1,000 holdings and thus qualifies through Lesnias's rule of thumb as minimal level. Six items are at the basic information level, having holdings between 500 and 1,000. Ten have holdings between 200 and 500, thus qualifying for the study or instructional support level. Although the maximum for the research level is set by Lesnias at 200, and the mean is known to be 153, the line dividing the research level and the comprehensive level (level 5 of the RLG conspectus) is not established. White suggests that the research level goes down to 1 holding, and does not define a range for Level 5 titles, as he views the comprehensive level as "gap-filling over Levels 1 to 4. . . . For libraries already at Level 4 and seeking comprehensiveness, 'Level 5' titles are simply any remaining desiderata."³⁰ But the present study suggests that most unique and scarce items in WorldCat have either been superseded by more recent publications or pertain to such localized interests that they have little research significance beyond the communities in which they were produced. Based on an admittedly subjective and casual examination of bibliographic records from the sample, most print books in English with holdings fewer than 40 are *not* at the research level: 40 seems to be a reasonable cut-off point for such items. However, several other items with holding levels as low as 17—print books in languages other than English, along with items in other document and material types—also seem to be at the research level. Finally, academic theses (manuscripts) at the doctoral level with 1 holding should by definition have research value, and thus count as being at the research level, though whether libraries would seek to collect them is another question. White is correct if he means that items with 1 holding *could* be at the research level, and support for his view can be found in Perrault's discovery that 63.5 percent of monographic titles in research libraries are indeed unique.³¹ But a qualitative examination of my sample suggests that, in practical terms, only a small fraction of the many items with fewer than 40 holdings actually *are* at the research level. If one accepts my suggestion of 40 as the lower cut-off point for the research level, and accepts that conspectus levels can in fact be equated with ranges of WorldCat holding levels, then the sample shows ratios for proportions of items at the various levels, from the minimal to the research levels, as 1 : 6 : 10 : 28. If, however, one accepts White's notion that the research level goes down to 1, the ratios are 1 : 6 : 10 : 478.

Summary and Conclusion

This study has analyzed a sample of bibliographic records in WorldCat to determine the proportion of items that have ubiquitous (widely shared) holdings versus lower levels of shared holdings. It left the term *ubiquitous* undefined, and worked with the heuristic categories of unique items, scarce items (items held by between 2 and 50 libraries), and non-scarce items (items held by more than 50 libraries). In these terms, the study finds that the large majority of items are either unique or scarce, with only 9.1 percent of items non-scarce.

Analysis of a random sample of 500 records in WorldCat shows that 2.2 percent of the sample is empty, matching no record or a record with no attached holdings. At the other end of the spectrum, the same percentage of the sample is matched by what may be called ubiquitous items. This uppermost fraction is obtained by taking the square root of the entire sample (22 for a sample of 500), using a formula devised by the information scientist Derek de Solla Price, to divide the sample into zones to account for items held by many libraries.³² The highest zone, using this formula, consists of items with more than 100 holdings. Dividing this zone into halves results in a subzone of ubiquitous items consisting of items held by more than 380 libraries (out of more than 9,000 member libraries contributing to WorldCat). The items in this category consist of print-format books in English, all published in the United States in the last half century, mainly by giant mass-market publishing houses. Virtually all are cataloged and transcribed at the national level. Following a rule of thumb stating that the more ubiquitous an item is, the more basic it is as an educational or informational resource, the dividing line between ubiquitous and nonubiquitous items corresponds approximately to the line between introductory and advanced levels of instructional support, by the standards of the RLG Conspectus. Only .02 percent of items are at the minimal Conspectus level.

The distribution of items in the sample shows an inverse geometric relationship between the number of items at a given level of holdings and the number of libraries holding items, with the largest fraction (39.8 percent) held by just one library. Proportions of items with higher holdings can be described by Lotka's law, using the inverse square of the items with one holding.³³ While the fraction of unique items seems high, it is substantially lower than was found in a recent and much larger study.³⁴ Moreover, 20.9 percent of unique titles are actually archival materials (mainly manuscripts); bracketing this portion of the sample, only 34.6 percent of remaining items are unique.

This paper presents a categorical scheme accounting for the distribution of items with various levels of holding by member libraries—unique, scarce, and non-scarce—

along with a system of zones accounting for levels of non-scarce and ubiquitous items. Used with other quantitative and evaluative measures, these categories and levels are helpful in assessing individual library collections and online union catalogs.

References

- Judith J. Senkevitch and James H. Sweetland, "Evaluating Public Library Adult Fiction: Can We Define a Core Collection?" *RQ* 36, no. 1 (Fall 1996): 103–17; Judith Serebnick, "Selection and Holdings of Small Publishers' Books in OCLC Libraries: A Study of the Influence of Reviews, Publishers, and Vendors," *Library Quarterly* 62, no. 3 (1992): 259–94; Debora Shaw, "An Analysis of the Relationship between Book Reviews and Fiction Holdings in OCLC," *Library & Information Science Research* 13, no. 2 (1991): 147–54; Danny P. Wallace and Bert R. Boyce, "Holdings As a Measure of Journal Value," *Library & Information Science Research* 11 (Jan. 1989): 59–71; Stephen E. Wiberley Jr., "The Humanities: Who Won the '90s in Scholarly Book Publishing," *portal: Libraries and the Academy* 2, no. 3 (July 2002): 357–74; Stephen E. Wiberley Jr., "The Social Sciences: Who Won the '90s in Scholarly Book Publishing," *College & Research Libraries* 65, no. 6 (Nov. 2004): 505–23.
- Allan N. Mirwis, *Subject Encyclopedias: User Guide, Review Citations, and Keyword Index* (Phoenix, Ariz.: Oryx Pr., 1999).
- Howard D. White, *Brief Tests of Collection Strength: A Methodology for All Types of Libraries*, Contributions in Librarianship and Information Science, no. 88 (Westport, Conn.: Greenwood, 1995); Anna Perrault, *Global Collective Resources: A Study of Monographic Bibliographic Records in WorldCat* (report of a study conducted under the Auspices of an OCLC/ALISE Research Grant, July 2002), www.oclc.org/research/grants/reports/perrault/intro.pdf (accessed Oct. 6, 2004).
- White, *Brief Tests of Collection Strength*, 135–36.
- Perrault, *Global Collective Resources*.
- Ibid.*
- OCLC Online Computer Library Center, "WorldCat," www.oclc.org/worldcat (accessed Aug. 11, 2004).
- OCLC Online Computer Library Center, "Watch WorldCat Grow," www.oclc.org/worldcat/grow.htm (accessed Aug. 11, 2004).
- OCLC Online Computer Library Center, "WorldCat Gold Records," www.oclc.org/worldcat/goldrecords.htm (accessed Aug. 11, 2004).
- Thompson M. Little, "OCLC's International Initiatives and the Online Union Catalog," *Cataloging & Classification Quarterly* 8, no. 3/4 (1988): 67–78.
- OCLC Online Computer Library Center, "OCLC Annual Report 2002–2003: Year in Review," www.oclc.org/news/publications/annualreports/2003/yearinreview.htm (accessed Aug. 16, 2004).
- Ibid.*
- Random.org—True Random Number Service, www.random.org (accessed Mar. 10, 2005).
- OCLC Online Computer Library Center, "WorldCat."
- Anglo-American Cataloguing Rules*, 2nd ed., 2002 rev. (Ottawa: Canadian Library Association; London: Chartered Institute of Library and Information Professionals; Chicago: ALA, 2002), v.
- See J. W. D. Dougherty, "Salience and Relativity in Classification," in *Language, Culture, and Cognition: Anthropological Perspectives*, ed. Ronald W. Casson (New York: Macmillan, 1981), 163–80; George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: Univ. of Chicago Pr., 1987); David B. Kronenfeld, *Plastic Glasses and Church Fathers: Semantic Extension from the Ethnoscience Tradition* (New York: Oxford Univ. Pr., 1996); Anna Wierzbicka, *Lexicography and Conceptual Analysis* (Ann Arbor, Mich.: Karoma Pub., 1985).
- Perrault, *Global Collective Resources*.
- Ibid.*, table 4-1a.
- Jay Jordan, "Global Networking of Information—OCLC's Strategy for the Future" (paper presented at 7th International Bielefeld Conference, Feb. 3–5, 2004, Bielefeld, Germany), <http://conference.ub.uni-bielefeld.de/proceedings/jordan.pdf> (accessed Oct. 12, 2004).
- OCLC Online Computer Library Center, "WorldCat Facts and Statistics," www.oclc.org/worldcat/statistics/default.htm (accessed Sept. 1, 2004).
- Concepción S. Wilson, "Informetrics," *Annual Review of Information Science and Technology (ARIST)* 34 (1999): 173; Alfred J. Lotka, "The Frequency Distribution of Scientific Productivity," *Journal of the Washington Academy of Sciences* 16, no. 12 (1926): 317–23.
- Dietmar Wolfram, *Applied Informetrics for Information Retrieval Research*, New Directions in Information Management, no. 36 (Westport, Conn.: Libraries Unlimited, 2003), 46; Derek J. de Solla Price, *Little Science, Big Science* (New York: Columbia Univ. Pr., 1963).
- White, *Brief Tests of Collection Strength*, 135–36.
- Ibid.*
- Anthony W. Ferguson, Joan Grant, and Joel S. Rutstein, "The RLG Conspectus: Its Uses and Benefits," *College & Research Libraries* 49, no. 3 (1988): 197–206.
- White, *Brief Tests of Collection Strength*, 127.
- David Lesniascki, "Evaluating Collections: A Discussion of Brief Tests of Collection Strength," *College & Undergraduate Libraries* 11, no. 1 (2004): 11–23.
- Ibid.*, 17.
- Ibid.*
- White, *Brief Tests of Collection Strength*, 129.
- Perrault, *Global Collective Resources*.
- Price, *Little Science, Big Science*.
- Lotka, "The Frequency Distribution of Scientific Productivity."
- Perrault, *Global Collective Resources*.

Perpetual Access to Electronic Journals

A Survey of One Academic Research Library's Licenses

Jim Stemper and Susan Barribeau

A perpetual access right to an electronic journal, defined as the right to permanently access licensed materials paid for during the period of a license agreement (not to be confused with the right to copy journal content solely for preservation purposes), is a concern of increasing importance to librarians as academic libraries discontinue paper subscriptions and retain electronic-only access. This paper explores the current environment for perpetual access to electronic journals. The authors report on analysis of the contracts between a large, research-level university library and 40 publishers of electronic journals, as well as ten large electronic journal aggregators. The authors seek to determine the frequency of contractual provisions for permanent access rights for the years of active subscription in the event an electronic journal contract is terminated for any cause other than breach by the licensee. Costs and formats of any granted perpetual access are considered. The paper concludes with an exploration of the potential impact of the perpetual access clauses libraries are accepting in licenses, the possible lack of continuing access, and options for addressing the situation.

Many, if not most, academic research libraries are engaged in canceling print journals for various reasons, including saving money and coping with escalating journal subscription prices, providing more immediate access to journal content, and alleviating shelving space problems.¹ This cancellation raises inevitable questions for the library: whether the print journal's online counterpart also will be canceled in the future to save money and whether ongoing access to subscribed content can be guaranteed.

Statement of the Problem

Faculty View

Faculty are taking note of the issue of long-term journal access. In a survey of 7,400 American university faculty members conducted in fall 2003 by the non-profit Ithaka, "three-fourths said a journal should ensure that its archives will be preserved indefinitely" and "Eighty-four percent of the survey respondents said that archiving of electronic resources was very important to them."² This finding suggests that faculty do care about the long-term availability of electronic journals (e-journals). The survey also echoes the findings of a study commissioned by the Journal Storage (JSTOR) Project in 2000, in which more than 4,000 faculty members at American universities were anonymously surveyed about their use of electronic resources.³ Seventy-six percent of the respondents agreed

Jim Stemper (stemp003@tc.umn.edu) is Electronic Resources Librarian, University of Minnesota Libraries, Minneapolis; **Susan Barribeau** (sbarribeau@library.wisc.edu) is Electronic Resources Librarian for Collection Development, University of Wisconsin Libraries, Madison.

The authors wish to thank the following colleagues for their comments on an early draft of this paper: Becky Albitz, Kate Flanagan, Aimee Glassel, Janice Jaguszewski, Kate McCready, Marcia Pankake, Kathy Robbins, Charles Spetland, and Cynthia Teague.

The authors wish to note that terms and conditions may have changed since they conducted their research.

that it is important “that electronic journals be preserved for the future.”⁴

Research Library Mission

Given the traditional stewardship mission of research libraries, one can understand why university faculty would look to librarians to preserve access to subscribed journals. In their collection development textbook *Building Library Collections*, Curley and Broderick emphasize the research library’s stewardship role.⁵ Several library authors have urged the profession to maintain this stewardship role in the Internet age.⁶

The Digital Library Federation (DLF) Electronic Resource Management Initiative (ERMI) defines a perpetual access right as “the right to permanently access the licensed materials paid for during the period of the license agreement”; this right is different than an archiving right, which is defined as “the right to permanently retain an electronic copy of the licensed materials.”⁷ The emphasis of the first term is on retaining access, not on how such access is achieved (through the publisher’s site, a locally retained copy, or a third-party site). A literature review shows that people define e-journal archiving differently. Librarians and publishers frequently use the term *archival access* when often meaning *perpetual access*. The focus of this paper is not on preservation of subscribed e-journals per se, but on contractually assuring ongoing access to them. The challenge of ensuring continued access to subscribed information resources for the future is at least partly in the hands of research-level academic institutions; therefore, research libraries should address the issue, make their policies clear for all to see, and hold to them.

Previous License Reviews

In 2001, Millett conducted a review of 61 publisher licenses for electronic content to ascertain whether *continued access* was addressed.⁸ If available, the exact license wording was quoted, labeled “Yes” if perpetual access upon termination was granted and “No” if not; any explicit mention of financial costs was noted as well. Out of 44 licenses with a “Yes” or “No” label (no reason was noted for the blanks on the other publishers), 22 were considered a “Yes” (half of the reviewed licenses) and 9 of those carried a charge (fewer than half of the “Yes” licenses).⁹ Hughes analyzed the Web sites of 15 e-journal providers and found “no agreement on how to provide access to journals after a subscription is lapsed,” though she predicted that “some sort of extra access fee so that lapsed users can access the volumes for which they had subscriptions may become the norm.”¹⁰ The authors of this paper sought to update and expand this research by determining whether e-journal providers were

willing to add perpetual access rights to an existing license if not already present, and to examine how issues of costs and format were addressed, if at all, in the license itself.

Print Retention Projects

Various consortia and national library organizations are pursuing print retention projects to ensure that someone somewhere still has a complete run of a journal’s print copy, available through interlibrary loan (ILL) for consortial members, should a library’s licensed copy no longer be available. The Committee on Institutional Cooperation (CIC) retains print copies of Academic Press, Springer, and Wiley titles in at least one member library; the University of California (UC) system stores print copies of Association for Computing Machinery (ACM) and Elsevier Science titles in a centralized, shared archive.¹¹ To identify which print archive holds the last copies of specific titles, the Center for Research Libraries is developing a national registry system.¹²

McDonald cited the important role that consortial membership plays in retention decisions as such, “Many libraries consider themselves or are considered by a consortia or regional group to be the archive of record for all or a portion of their collective print collection. It will continue to be their responsibility to maintain print issues for those items that are their archival responsibility.”¹³ However, there are several drawbacks to print archives that must be considered. Due to print subscription costs as well as attendant hosting, processing, and tracking commitments, print retention projects may entail a significant monetary investment. Also, the *Library Journal Academic Newswire* paraphrased one participant (Baker, chair of the CIC Library Directors Group) as cautioning that “at some point, major publishers might choose to abandon print altogether, posing yet another challenge.”¹⁴ Such print retention projects are in their infancy, and the academic library community does not yet know whether document delivery from such repositories (perhaps delayed, perhaps completely unavailable due to copyright clearance requirements) will be acceptable to users. Libraries determined to go e-only to save money may wish to allocate money for print retention projects and seek membership in a consortia that has such projects.

E-journal Archiving Projects

Simultaneous with these efforts to preserve a print journal archive, various projects have studied e-journal archive creation and maintenance. Extensive research efforts have assessed the scope of the e-journal archiving issue and proposed strategies for preservation of and long-term access to full-text content.¹⁵ The Task Force on Archiving of Digital Information, under the auspices of the Commission

on Preservation and Access and the Research Libraries Group, issued an influential 1996 report.¹⁶ Based on the findings of the task force's report, the commission's successor organization, the Council on Library and Information Resources (CLIR), along with DLF and the Coalition for Networked Information (CNI), issued "Minimum Criteria for an Archival Repository of Digital Scholarly Journals," the sixth of which has special implications for licensing librarians: "A repository will make preserved information available to libraries, under conditions negotiated with the publisher."¹⁷ What type of access a lapsed subscriber might have at such an archive and what license wording is specific enough to ensure such access are the critical issues. The authors' analysis of one institution's licenses explores how these access issues manifest themselves in contracts.

As part of its Archiving Electronic Journals project, the Mellon Foundation funded seven year-long studies that aimed to meet these repository criteria, six of which involved large academic research libraries working in tandem with major journal publishers.¹⁸ Most projects addressed issues related to technological infrastructure and workflow, as well as sustainable business and access models. Except for the publishers who worked with Stanford on the Lots Of Copies Keeps Stuff Safe (LOCKSS) project, none seemed open to the idea of access to the archive by lapsed subscribers. Users could have access to e-journal archives only under certain conditions usually called *trigger events*; conditions discussed between Harvard and Blackwell Publishing, the University of Chicago Press, and John Wiley and Sons typify those cited in the project reports:

1. When material is no longer accessible online from the publisher.
2. When the publisher sells or otherwise transfers the rights to publish a given title to another body.
3. When the material has been in the archive for "n" years ("n" being a time period to be agreed to by Harvard and the publisher on a title-by-title basis).
4. When the title ceases to be published.
5. When the content enters the public domain.¹⁹

Elsevier stipulated that its prototype archive at Yale University Library could not compete with the publisher's revenue stream.²⁰ This archive's trigger events were correspondingly limited: users not affiliated with Yale could only access the archive on-site at Yale or in the event of a natural disaster, as ILL was not part of the archive's mission.²¹ To a faculty member of an American university whose library was forced to cancel a needed ScienceDirect title for budgetary reasons, the prospect of making a trip to Yale University Library to photocopy an article is not an attractive or likely option. More practical alternatives for research libraries are needed. The University of Pennsylvania recommended,

among other things, that further research in this area "determine the relationship between events that would 'trigger' nonsubscriber access rights versus paid subscriptions."²² Short of each library setting up its own archive of a publisher's journals, and in the absence of traditional ILL service, how could a lapsed subscriber obtain access to content for which the library had once paid without adversely affecting the publisher's bottom line?

The Yale/Elsevier e-journal archiving project, while drawing no final conclusions about economic models pending further study, identified preservation metadata as one way to add value to journal content, thus justifying some type of setup or maintenance fee.²³ The project report also observed that "the ongoing [archiving] costs, at least for standard publisher's journals, could be relatively predictable and eventually stable over time."²⁴ This outlook of feasibility and stability is at least partially good news for library administrators.

After sifting through the lessons learned from the Archiving Electronic Journals projects, the two models the Mellon Foundation felt worthy of continued funding were LOCKSS and Ithaka's Electronic Archiving Initiative. The key issues that led to this decision to fund two different archiving options revolved around cost and control—no one project's means of financing an archive was identified as being the clear solution, and neither publishers nor libraries were willing to cede control of an archive to the other.²⁵

The innovation of the LOCKSS model is its decentralized nature—allowing libraries to create locally cached copies of subscribed e-journal content that they could deliver to authorized subscribers as needed with the sanction of participating publishers, with other institutions' peer-to-peer network of caches acting as backups.²⁶ With LOCKSS staff negotiating publisher participation, all the individual library needs to do is to get the publisher's blanket permission to cache the title, which LOCKSS's already-negotiated license wording facilitates.²⁷ The LOCKSS site claims that "more than 80 libraries and 50 publishers . . . are using the software."²⁸ Like many projects, LOCKSS now needs secure funding to remain viable, so the new LOCKSS Alliance is seeking participation from new libraries who would pay annual membership fees.²⁹

The second archiving model that Mellon is funding is Portico, a recently launched project of the nonprofit organization Ithaka (itself an outgrowth of JSTOR, the respected e-journal back-file archiving service).³⁰ Also called E-Archive, Portico argues for a centralized archive of electronic scholarly literature, with distributed costs shared by participating publishers and libraries. Portico justifies the centralized approach on the grounds that most publishers will not see a long-term business interest in maintaining and migrating e-journal data as the content gets older, and strained library acquisition budgets will not be able to cover

such costs.³¹ Portico's proposed funding model is similar to JSTOR's; participating libraries are asked to pay a one-time archive development fee and an annual archive support fee. Participating publishers are asked to pay a one-time setup fee and then an annual deposit fee for each journal.³² The funding model thus reflects the findings of Mellon journal archiving projects that journal archives have both initial setup and ongoing maintenance costs. The Portico site lacks specifics about the conditions under which one can obtain access to content. The mission page's statement that "access to the contents of the archive will be available under limited circumstances" may be cause for concern.³³ Exactly how limited are the circumstances? The JSTOR moving wall of content is mentioned on the Portico site, which apparently means that a library that has canceled a title may have to wait a publisher-specified number of years before the once-subscribed issues become available through Portico.³⁴ What remains to be seen is whether most or all of a library's desired savings from going e-only is simply transferred to such an archiving service, and whether the limited years of Portico access are of sufficient value to users.

If a critical mass of publishers joins the Portico project, the approach may offer a compelling alternative to LOCKSS. Publishers would not have to worry about losing control of their content in distributed LOCKSS caches; libraries would not have to setup and maintain local LOCKSS caches. However, LOCKSS co-founder Reich maintains that "only the largest publishers have sufficient resource to implement (or negotiate with third parties to provide) archives for content they publish," which makes LOCKSS's low cost a compelling option for smaller publishers.³⁵ Some publishers may opt for only one of the LOCKSS or Portico approaches (or neither), meaning that libraries might well end up with a hybrid approach to perpetual access, with implications for budgeting and presenting a coherent user interface:

- local or regional print archive (when the publisher offers no perpetual access rights);
- local data loading (e.g., LOCKSS cache);
- access through a central subscription archive (e.g., Portico); or
- access through the publisher's server, which may or may not entail a maintenance fee.

Individual Library Actions

Many individual university libraries and consortia are attempting to address this problem systematically through collection development policies governing the move to electronic-only collections.³⁶ Whether labeled criteria, checklist, model license, or principles, these policies show that libraries need guidelines so they can safely make the move to e-only journals without jeopardizing long-term access. Such

guidelines typically state the need for an explicit guarantee of perpetual access rights in the license should the library cancel a title in print or electronic format. Potential costs and hosting options for continued access also are frequently mentioned. Both the University of California (UC) and the United Kingdom's National Electronic Journals Licensing Initiative (NESLi2) for higher-education institutions state that perpetual access must be free, while the Canadian National Site Licensing Project states that members will have to pay a fee to the publisher. UC and NESLi2 also state that three hosting options are acceptable: access through the publisher's site, the library's site, or some type of central archive hosted by a third party.³⁷ Policies from several libraries and consortia stipulate that perpetual access should be guaranteed even when a journal title ceases publication or is transferred to another publisher.³⁸

The University of Maryland's policy document provides a useful example in terms of detail and context. The criteria provide three scenarios under which one might consider going e-only (no publisher commitment, an expressed commitment from the publisher, and actual planning by the publisher for perpetual access), giving the selector progressively more leeway to go e-only as the publisher's commitment and the core nature of the title allow. Such tiered decision-making acknowledges that no single answer fits all situations, even within one library system.³⁹

Guidelines are only as effective as the enforcement behind them. Are they ideals for which to strive, or required elements of any license? Some libraries have demonstrated the willingness to apply the guidelines in practice. The California Digital Library (CDL) publicly rejected the Nature Publishing Group's 2005 renewal terms because of the lack of guaranteed perpetual access and their belief that "*Nature* may change hands, or *Nature* may be archived by a third party, and UC does not expect to pay repeatedly for the same content."⁴⁰ Because consortia represent so many libraries and so much subscription income for publishers, they are well-positioned to exert market power and apply criteria for the granting of perpetual access rights.

Surveys of Current Practice

Anecdotal evidence from postings to the Electronic Resources in Libraries (ERIL-L) and Liblicense electronic discussion lists indicate that librarians are not taking publishers up on their offer of LOCKSS access.⁴¹ Shreeves, collection development officer (CDO) and associate director of the University of Iowa Libraries, said:

I suspect that the "perpetual access" and ownership rights that publishers like Elsevier often include in their contracts are in some measure convenient fictions we use to convince ourselves that the old

ownership, or stewardship, model still works in the digital age. . . . But in practice the tremendous uncertainties of digital preservation thwart our ability to make useful distinctions between nominal ownership and leasing of information. This is not to minimize the fundamental importance of this distinction in principle—I assume we all try to negotiate permanent or archival access to any product we license.⁴²

The authors tested the latter assumption in August 2004 by posting a brief survey to the Chief CDO's electronic discussion list, which includes the CDOs of the largest Association of Research Libraries (ARL) member institutions. The authors' institutions were not included in the survey, nor were national and public libraries, which left a survey population of 38 potential respondents out of 47 members. CDOs at 25 institutions responded (for a response rate of 65.8 percent, or two-thirds of the authors' peers). Members were asked whether, during e-journal license negotiation, they ask for perpetual access rights to subscribed years upon cancellation, and whether they consider the lack of perpetual access to be a deal breaker. While this was not a scientific sample, the results were still informative and clear—a majority of respondents (76 percent) consistently ask journal publishers for perpetual access, yet a majority (also 76 percent) would still sign an e-journal license without assurances of such access.

An ARL survey on e-journal licensing, conducted the previous year by Case, was published in the same month the authors' survey went out to the Chief CDOs list and found similar results.⁴³ The ARL survey documented a gradual but definite move towards e-only access. In 2002, only a few of the 40 respondents were going e-only; in 2003, 25 of 57 respondents (44 percent) were doing so for new packages, and "another six [respondents] indicated . . . they seemed to be moving in that direction."⁴⁴ Further, a minority of ARL institutions view "Archival Access/Perpetual Rights/'Own' the Content" as a deal breaker in license negotiations (14.6 percent of the respondents for ARL, where the Chief CDOs survey found 24 percent).⁴⁵ Put another way, using Case's survey, the remaining 85.4 percent of ARL institutions evidently do not view lack of perpetual access as a contractual deal breaker; most likely, many of this clear majority of the respondent institutions were among the 44 percent who were going e-only for new packages. One must take note that so many research level libraries are not assigning greater weight to perpetual access guarantees when making the big move to e-only collections. Taken together, both surveys suggest that fewer than 1 in 5 large research libraries say no to a license that does not provide perpetual access assurances.

A subsequent survey sent to the e-journals discussion list, to which primarily United Kingdom academic librar-

ies responded, echoes the previous findings.⁴⁶ A substantial number of the institutions (80 percent) had "already moved to e-only for at least part of their collections," but a lack of prior planning was occasionally reported: "Sometimes this [decision] is underpinned by a strategy or policy, but more often it seems to [be] happening as a reaction to other pressures," such as "shrinking bookfunds and lack of space."⁴⁷

With the large numbers of libraries going e-only, as demonstrated by these surveys, one cannot realistically expect many libraries to rely on a print copy as the primary means of perpetual access. For example, Bracke and Martin maintained that, for the University of Arizona's Science-Engineering Library, "shrinking budgets and space constraints make print retention an untenable and non-strategic plan of action."⁴⁸ At the University of Minnesota and University of Wisconsin, the move to e-only journal access is encouraged for the same reasons. In cases where a specific license does not permit perpetual access rights, retaining a print subscription for specific titles may be an affordable option, especially if the library belongs to a consortia that participates in a print retention project.

Access through the Publisher When the Journal Changes Hands

The transfer of journal title ownership from one publisher to another is an increasing problem. McDonald defined the problem thus:

Libraries must have a secure sense that the publisher or the project is reputable and likely to retain their electronic publishing program at a reasonable cost in future years. The purchase of Academic Press by Elsevier has laid a heavy economic burden on many libraries and future purchases of commercial publishers by others can wreak havoc on the information resources for any library. . . . Libraries that long ago canceled print issues of journals from Academic Press in favor of electronic only access [through the International Digital Electronic Access Library (IDEAL) service] are now faced with huge cost increases when renewing these licenses through Elsevier.⁴⁹

Eells noted that Elsevier did not honor some perpetual access commitments for former IDEAL customers, meaning that customers essentially had to pay for access to some years twice (once through IDEAL, then through ScienceDirect).⁵⁰ Bird and Waller examined the licenses for 6 publishers that offered what the authors termed "big deals" (Academic Press, American Chemical Society [ACS], Elsevier, Institute of Physics, Royal Society of Chemistry, and Springer), with particular attention to perpetual access

clauses for journals that these publishers sold and that were no longer part of the current package. All of the surveyed publishers except Springer offered perpetual access rights to the Canadian Research Knowledge Network as part of the license and delivered ongoing access upon the sale of ten randomly chosen titles.⁵¹

Access through Local Data Loading

During the last decade, academic libraries have been moving from locally loaded databases in favor of Web access through the publisher's site. The primary reason is the high infrastructure cost for librarians to network content stored on disc, but the user preference for the convenience of the Web and the potential lack of enhanced functionality, such as cross-links are factors as well.⁵² LOCKSS addresses many of these needs, but the lack of widespread adoption by librarians and publishers merits the consideration of other options. While the OhioLINK consortium has been able to secure perpetual access rights to locally loaded journal content in licenses from all publishers except ACS, Diedrichs noted that "the investment required in the archiving process is considerable and the question of who best should perform that function—the library, the publisher, some middleman—is yet to be resolved for most libraries."⁵³

Grudzien, head of collection development at Central Michigan University Libraries, illustrated that even if one is able to obtain perpetual access assurances, the option of local data loading is not to be taken lightly, citing late delivery of Elsevier journal content, and subsequently the need to both locate outdated tape drives for accessing the content stored on digital linear tape and to organize the content itself.⁵⁴ The fact that only 10 Elsevier customers worldwide have opted for the locally loaded version of ScienceDirect may be an indication of the preference for accessing e-journals at the publisher's site and thus avoiding the attendant infrastructure and staffing costs of local data loading.⁵⁵

Access through Third Party Solutions

The e-Depot at the Koninklijke Bibliotheek (the National Library of the Netherlands) acts as an archive for BioMedCentral, Blackwell Publishing, Elsevier Science, Oxford University Press (OUP), Sage, Springer, and Taylor and Francis journals.⁵⁶ However, for institutions outside the Netherlands who cancel any subscriptions to these journals, e-Depot is only an archive of last resort. The head of e-Depot has said that these depositing publishers allow three limited types of access:

- on-site access for registered users;
- ILL supply in the Netherlands; and

- access for any licensee in case publishers cannot meet obligations (calamities, bankruptcy).⁵⁷

Subscribers of major e-journal packages outside the Netherlands cannot cancel titles and expect to get free ILLs of articles for their users; they will need to make other arrangements. Still, the experiment may show promise as a model for other countries, such as the United States (perhaps in conjunction with consortial site-licensing projects at a regional or national level), and bears watching.

A research library's mandate to provide current access to journals for today's scholars can be at odds with the mandate to keep those journals available to be accessed by scholars in the future. Librarians still value their stewardship role in the digital realm, but they perhaps fear that pressing the issue contractually is commercially and financially unrealistic at this time. If the level of public support for higher education continues to erode, acquisition budgets continue to dwindle accordingly, and research libraries continue to go e-only, this practice will have major implications for the type of library collection scholars will see in the future. Research libraries will have to grapple with the way they now interpret their traditional stewardship role, and come to grips with the issue of perpetual access to subscribed e-journals.

Objectives and Methods

The goal of this paper is to determine the frequency of contractual provisions for perpetual access rights for the years of active subscription in the event an e-journal contract is terminated for any cause other than breach by the licensee. If a library asks for perpetual access rights in a license, how often will such access provided? Has the access-ownership model triumphed, or can libraries promise users that the electronic content for which the libraries have paid will still be available in the future? The authors' working hypothesis was that the majority of e-journal provider licenses do not provide explicit allowance for perpetual access rights.

Toward this end, the authors analyzed a representative set of journal publisher and aggregator licenses signed by the University of Minnesota as of the start of the fall 2005 semester (40, or about one-third, of the library's journal publishers), looking for perpetual access clauses. September 1 is the usual time the university's journal cancellation decisions are due to its serials agents. Separately, the authors analyzed the licenses for 10 major e-journal aggregators (defined as "a bibliographic service that provides online access to the digital full-text of periodicals published by different publishers").⁵⁸ A smaller number of aggregators was analyzed because there are not as many in the marketplace from which to choose, and a few of these

aggregators control a significant market share. Aggregators were included because selectors can be tempted to cancel print versions of journals covered in aggregated databases just as they would be for an individual journal publisher; aggregators were analyzed separately to see if they had any unique issues or findings. Thus a total of 50 e-journal licenses were analyzed.

There were three criteria for selecting the journal publisher set and the aggregator set:

- a mix of commercial and society or university press publishers;
- a mix of large and small publishers; and
- journal collections likely to be held by large libraries.

By examining collections likely to be subscribed to by most research libraries, the authors hoped to get a better sense of the implications for lost national access should libraries cancel the paper versions of these journals with no print backup or ensured contractual access upon termination. By examining equal numbers of commercial and society publishers, the authors hoped to ascertain whether one type of publisher was more or less likely to grant such access. The providers surveyed are listed in the appendix.

For each e-journal collection from these entities, the authors looked at the library's existing license and asked the following questions:

- In the event that the library terminates the license, is the publisher or aggregator willing to provide some form of perpetual access to the years licensed during the subscription?
- If the publisher or aggregator grants perpetual access rights, does the license specify an associated cost?
- If the publisher or aggregator grants perpetual access rights, does the license specify a format or access mode (e.g., continuing access through the provider's Web site, local data loading through a CD- or DVD-ROM or a LOCKSS cache, or access through a third party, such as Portico)?

At the University of Minnesota–Twin Cities and the University of Wisconsin–Madison, the practice is to ask for a contractual guarantee of perpetual access rights when licensing electronic access to journals. In some cases, the publisher or aggregator generic license did not offer such access, and the library had to ask for the right during the license negotiation process. Sometimes the library had to ask a publisher representative for clarification when the license wording did not clearly answer the perpetual access question. As any licensing librarian knows, this negotiation process is central to contract work. The University of Minnesota adapted language from a generic academic library single-institution license template recommended by ARL: “On termination

of this License, other than for cause, the Publisher shall provide continuing access for Authorized Users to that part of the Licensed Material that was published within the Subscription Period.”⁵⁹ Note that this language does not mention costs. Also, wording in the template related to choice of format (“either from the Server, or by supplying {electronic files} {CD-ROMs} {microfiche} to the Licensee”) was deleted, as the University of Minnesota did not want to invite the prospect of locally loading journal content.⁶⁰

Each final, signed license was examined for any clause related to perpetual access rights. If a provider's generic contract explicitly granted a perpetual access right (as earlier defined by DLF's ERMI), or if the university was successful in getting the provider to include this right during negotiations, this was counted as a “Yes.” Often, if included, the clause was under such a heading as “Termination,” “Usage Rights of Lapsed Subscribers,” or “Perpetual License.” The relevant clause was copied and pasted into a master document so all provider wording could be compared side by side. Decisions on availability of access rights for each provider were recorded in an Excel spreadsheet, as were any contractual mentions of cost or access format.

If a license was silent on the issue of perpetual access rights, this was counted as a “No.” The authors followed the ARL's Strategic and Practical Considerations for Signing Electronic Information Delivery Agreements, which advise the licensing librarian to get desired rights in writing: “If the language of a contract leaves you feeling unsure how users may use the materials, clarify the contract by writing in what you need.”⁶¹ Because every license has a specified term or period of performance, once that time has past, a university has no rights that the agreement does not expressly grant. This is why one sometimes sees contractual provisions stating that the rights and obligations of section X shall not be terminated upon the conclusion of the term of the agreement. So, if the license renews annually and a university wishes to have access to the material after the term's end date, then the license needs to include a provision granting the continued right of access. While a library could attempt to renegotiate an existing license that does not explicitly grant perpetual access rights in the future, the authors were only interested in the licenses that grant such access at the present time. This paper thus affords a glimpse of what the University of Minnesota would have perpetual access rights to the day after the library canceled its e-journal subscriptions with the surveyed publishers (through its serials agents or directly), barring any later renegotiation with these publishers.

Results

Overall, most publishers—by almost 2 to 1 (64 percent to 36 percent)—in the sample grant perpetual access (see table 1).

This initial finding makes the prospect of continued access sound realistic *if* librarians ask for the right and *if* librarians vote with their dollars. For years, librarians have been discussing the pros and cons of access versus ownership of material; however, in the electronic content delivery climate of today, that unembellished model is too simple to cope with the related complexities and ramifications in ongoing discussions of the current e-journal licensing environment. In the authors' perspective, perpetual access should be the clear objective. To that end, examining and defining *how* such access will be provided, to *what* content and by *whom*, and under what *cost* and *format* conditions, is necessary in library license negotiations.

Within the scope of the licenses examined, more of the commercial publishers' agreements allow or provide for perpetual access rights than do those of society publishers—72 percent of commercial publishers and aggregators, compared to 56 percent of society publishers and aggregators. Almost three-quarters of commercial publishers' licenses provide for perpetual access, while more than half of societies' licenses do not mention or explicitly decline to provide perpetual access. Libraries' past experiences with scholarly publishers versus commercial publishers (with regard to pricing issues and a general cooperative spirit) created an expectation that the society publisher agreements would naturally be more likely to provide for these access rights. The actual outcome was unexpected, considering the belief (from the perspective of the authors) that society publishers will be more responsive and responsible players in the scholarly communication realm. The publications of scholarly societies are essential components of academic and research library collections, and these publications present a balance to the commercial publications also relied upon to disseminate scholarly research.

Some of the declining publishers give their "No" answer right in the contract and did not rely on silence alone. For example, the following clauses seem to say "Don't even ask":

Walter de Gruyter does not warrant that the Licensed Materials will be made available permanently.⁶²

In the event of termination or expiration, the Subscriber may not retain any portion of the [Association of Computing Machinery Digital Library].⁶³

Some publishers with free online access with print subscription arrangements simply tell libraries that they lose online access when they cancel print subscriptions. More and more, though, these publishers are reversing the traditional cost base, making print an add-on to electronic access, rather than the other way around. In this new environment, perpetual access rights would then have to be negotiated as part of a renewal. Haworth Press goes even farther, advising librarians not to cancel print prematurely. Because Haworth Press considers "the electronic/print serial scene [as] still too volatile and too early in an infancy stage to allow [Haworth] to dismantle [its] entire in-house printing operation," they regard the print version as the "archival back-up" to the electronic version.⁶⁴ In the absence of an explicit contractual guarantee of perpetual access rights for the online version, or a collaborative print retention program for this publisher, one should weigh the issues carefully before canceling the print version of a Haworth journal.

The American Medical Association's representative, in an e-mail, explained her society's stance this way: "We do not allow perpetual access. A site license is like getting cable—you are paying for access to the programming, not the ownership of the actual programming."⁶⁵ This statement seems to preclude any possibility of entering into a discussion of the issues as part of a contractual negotiation.

Some publishers go so far as to require the licensee to destroy all copies of downloaded articles. For example, one commercial publisher says, "Licensee agrees to destroy, and will use reasonable endeavors to instruct Authorized Users to destroy all Licensed Material stored on any digital information storage media, including, but not limited to, system servers, hard discs, diskettes, and back up tapes."⁶⁶ Librarians try to negotiate the deletion or rewording of such clauses, due to privacy and enforceability concerns, but again the implication is that the library does not own this data and thus does not have a continuing right to access the data. Implicit in this clause is a major shift in the concept

of ownership that needs to be examined carefully, not adopted out of short-term expediency.

A handful of providers note specific expiration parameters for perpetual access in the licenses. Springer specifies two years, the Online Computer Library Center (OCLC) five years, and Karger ten years.⁶⁷ Perpetual access may not be

Table 1. Publisher's contractual position on perpetual access rights in some form (n=50)

Publisher type	% yes overall		% yes within this publisher type	% no overall		% no within this publisher type	Total
	Yes	No		Yes	No		
Commercial	18	36	72	7	14	28	25
Society	14	28	56	11	22	44	25
Total	32	64	n/a	18	36	n/a	50

perpetual at all; the clock may be ticking. Along with cost and format, this is yet another aspect that must be considered by selectors when deciding whether to cancel a title. If nothing else, the selector may need to reconsider reinstating such a title or package, or renegotiating some form of “subscribed years” access, probably for a price.

Separately, the authors analyzed the licenses for 10 full-text database providers (as earlier defined), 5 of which are commercial providers and 5 of which are society providers. Several of these 10 are often referred to as aggregators, in that they provide journal content from publishers, but are not the publishers themselves and generally do not provide for continued access should the subscription be terminated. The commercial providers in this group are from EBSCO Industries, Gale, H. W. Wilson, Ovid, and ProQuest; the society providers are BioOne, Highwire Press, JSTOR, OCLC Electronic Collections Online (ECO), and Project Muse. Few offer any contractual guarantee of perpetual access. Canceling a print title included in such a package is risky. Their title lists and coverage frequently change, depending on their contracts with the publishers. If a library's agreement is not renewed, one cannot count on any continued access to the journals therein.

Cost Issues

In a parallel finding that tempers the two-to-one result about perpetual access being available, the authors found that almost half of the providers that allow perpetual access specify that a charge will or may be associated with this access (14 of 32 granting providers, or 43.8 percent); see table 2. A roughly equal percentage of commercial and society providers who allow perpetual access charge for such access (44.4 percent versus 43 percent). As the Mellon Foundation's journal archiving projects demonstrated, hosting and maintaining content comes with an attendant cost. Thus one can reasonably expect a cost to be associated with a perpetual access service, whether handled locally, by the publisher, or by a third party.

The licenses that mention cost do so in a variety of ways and at differing levels of specificity. In the majority of the licenses studied, the fee is to be paid to the publisher (75 percent of commercial publishers, 100 percent of society publishers). In a handful of cases, the license specifies a fee be paid to a third party or simply is not clear. The library's own setup costs for loading data locally are implied in many cases and are not explicitly stated. More study is needed in this area for

purposes of budgeting for perpetual access. Examples of contract language about costs follow.

Some publishers distinguish between *content* costs and *access* or *setup* costs. The American Institute of Physics (AIP) exemplifies a situation where the library might be assuming both types of costs: one for the publisher and one for library technology infrastructure investments: “The Subscriber will be given the option to purchase a physical archive copy, for example a CD-ROM. . . . Any hardware or software required to distribute content from the archive copy will be the responsibility of the Subscriber.”⁶⁸

Most licenses are not as specific about a library's internal costs but assume that the library knows about, can cope with, and will absorb such hardware, software, programming, interface design, connectivity, and staffing costs. Similarly, the American Psychological Association (APA) distinguishes between content fees and access fees: “Although [the customer] would not pay ongoing data fees, they would pay for the delivery of the content.”⁶⁹ Either way (whether a library is paying the publisher or library or technology vendor support staff), sometimes multiple and ongoing costs for perpetual access will exist.

Specific costs usually are not detailed in the license, even when the concept is mentioned and can be subject to change at the publisher's discretion. Typical wording is:

There may be fees associated with these options.⁷⁰

Wiley will provide the Licensee with access to the full text of the Licensed Electronic Journals published during the Term of this License . . . at a cost-based fee agreed by both parties.⁷¹

Upon termination of this Agreement, Licensee may retain the right to use in archived form the content of the Database provided that Licensee . . . pays all costs associated with providing the Database content to Licensee on a mutually agreeable media type . . . Licensee acknowledges that the terms and conditions applicable to Licensee's archiving rights under this paragraph including, but not limited to, the media type and annual fee or fee

Table 2. Granting publishers' contractual position on cost for perpetual access rights (n=32)

Publisher type	Yes	% yes overall	% yes within this publisher type	No	% no overall	% no within this publisher type	Total
Commercial	8	25.0	44.4	10	31.3	55.6	18
Society	6	18.8	43.0	8	25.0	57.0	14
Total	14	43.8	n/a	18	56.3	n/a	32

per year of archive material may be modified by [Institute of Electrical and Electronics Engineers] at its discretion.⁷²

Of the three previous examples, the likelihood of a fee increases with each succeeding quote. The first says there *may* be a fee; the second says that there *will* be a fee and the library has a *say* in what that fee might be; and the third says there *will* be a fee and the amount is *open-ended*. So the potential exists for perpetual access fees of undetermined amounts that may change from year to year. Because perpetual access is a relatively new and undefined service, this is not surprising, but libraries must be aware of the budgetary implications of open-ended financial commitments. One could plausibly argue that the uncertainty surrounding perpetual access fees may not be substantially different than the fluctuations in journal prices; still, this fluidity will make budget projections even more challenging to produce than they are now. The license negotiation process is the library's opportunity to participate actively and creatively in further definition of cost assessment, perhaps in consultation with colleagues at peer institutions.

Sometimes the cost is specified, at least with respect to existing subscription fees. For example, East View specifies a fee of 10 percent of the current annual subscription cost.⁷³ Sometimes the cost depends on the format, and the effort required to convert the journals to or deliver the content in that format.⁷⁴

In addition to LOCKSS access, the American Society for Microbiology (ASM) allows open access to its journals after a specified time, but ASM reserves the right to charge a fee (for current and back-file issues) in the future. This license serves to remind one of the unknown long-term potential of relatively new open access models.⁷⁵

The offering of separate back-file journal sets for years not covered in a current subscription is increasingly common (e.g., from ACS, Annual Reviews, Elsevier, Nature, Wiley). One might easily find oneself in the situation of having bought a journal back-file set, but being without access to current issues due to cancellation of a title for which perpetual access rights are not contractually guaranteed. In cases where perpetual access to a canceled journal is not allowed in the license, a library might opt to retain the subscription but would need to pay more to retain access to the back files. The American Association for the Advancement of Science (AAAS), publisher of the popular journal *Science*, says that "because AAAS cannot be certain of future technology, storage, or maintenance costs, AAAS cannot guarantee [back-file] access" and it "reserves the right to remove all or portions of the archive of past issues, or to institute a charge for access to it in the future."⁷⁶

JSTOR has been a unique scholarly project, embraced by many research libraries as a trusted repository of key

scholarly journals, as evidenced by their extensive list of participants that have licensed their journals. The expectation of the authors, based on JSTOR's good reputation in the library community, was that the JSTOR agreement would be an exemplar for perpetual access. However, the agreement at the time of this research does not allow for access to content after a JSTOR subscription termination.⁷⁷ JSTOR charges an archive capital fee; the authors would posit that the use of the term *capital* in reference to this fee implies a capital investment by the institution. Learning that the archive capital fee does not entitle one to permanent access to the collections as defined at the time of one's subscription is disconcerting. JSTOR's two-fold pricing model, reflecting the need to recoup up front and ongoing maintenance costs, is a common one. The expectation is that this is a trusted repository that a library can depend upon for continued access. While libraries might be surprised to learn that their initial capital investment alone does not guarantee them perpetual access rights to JSTOR content in the future, they should expect this separation of setup and maintenance costs from other publishers.

Format and Hosting Issues

When one looks at the types of formats and hosts that publishers offered for perpetual access, one finds that the option of continued access through the publisher's own server is offered just as often as the option of local data loading (i.e., at least one of these two options was offered in 53.1 percent of the studied licenses); see table 3. Also, the commercial publishers and society publishers studied are similar in the percentage that offer such access through their own server, with a slightly greater percentage of society publishers offering this option (57.1 percent to 50 percent).

Sometimes one finds a clear correlation between current access practice and perpetual access. Six of the 32 (18.8 percent) granting publishers that provide Web access will continue that access as the sole format. However, as noted earlier, Springer puts a time limit on perpetual access to its server, after which other options must be considered: "Access is granted for two years after cancellation; access and storage options are available for subsequent years."⁷⁸

By a large margin, more society publishers than commercial publishers allow the library to load data locally (71.4 percent to 38.9 percent), perhaps reflecting a greater concern among commercial publishers about retaining control of their data. AIP allows the library to copy data from a *physical archive copy* to the library's secure network.⁷⁹ Project Muse also allows this practice, on DVD, though the *non-searchable* format of their content does not sound very useful and may well imply significant, hidden setup costs for the library to make the dataset searchable (such as metadata tagging): "Approximately 90 days after the expiration of an

annual subscription term, Project Muse will provide, upon request, an archival (non-searchable) file on DVD-ROM or other appropriate media as determined by Muse, containing the content of all issues published online during the 12-month subscription term.”⁸⁰ Serving up this kind of data will not be plug-and-play for the library and, judging from the recent library literature, not enough institutional reporting on local experiences exists (beyond anecdotal electronic discussion list postings) to tell if this is a cheaper option than paying the publisher an annual maintenance fee.

Of the publishers included in this survey, 8 commercial publishers (Berkeley Electronic Press, Blackwell, British Medical Journal [BMJ] Publishing Group, Cold Spring Harbor Laboratory Press, Emerald, Lippincott Williams and Wilkins, Nature Publishing Group, Springer) and 8 society publishers (AAAS, ASM, BioOne, Cambridge University Press, HighWire Press, Massachusetts Medical Society, OUP, and Project Muse) are listed on the LOCKSS site as having agreed to make content available through LOCKSS.⁸¹ The LOCKSS option is not always stated in their licenses. However, just because a publisher does not mention LOCKSS in their license does not mean they will not consider the option or have not already done so. Silence is another reminder of the need for negotiation or for checking the current list of LOCKSS participating publishers. Overall, 16 of 50 (32 percent) of the studied publishers are represented as partners on the LOCKSS Web site. In addition, 6 of the studied publishers are listed as partners on the Portico site.⁸² Together with the LOCKSS partners, this adds up to 21 of the 50 studied publishers (42 percent) publicly opting for at least one of these perpetual access mechanisms. That is a respectable amount, considering that so many of these publishers’ journals are collected by research libraries. These two projects merit consideration as potential solutions to the perpetual access problem. Of these 21 publishers, only 1 (Blackwell) is a partner in both LOCKSS and Portico, indicating that publishers may opt for only one of these approaches. On the other hand, 2 major publishers included in this study, Elsevier and APA, are not LOCKSS or Portico partners, which shows the need to make sure such important publishers have perpetual access assurances in their standard licenses (as they do) or to negotiate their inclusion.

Of the publishers that allow perpetual access, 21.9 percent specify that a third party will provide the access and deal with format conversion issues. The

numbers for commercial publishers and society publishers are very close in this area. Blackwell’s wording is typical in this regard:

The Publisher shall . . . provide the Licensee with assistance in obtaining continuing access . . . from a third party’s server provided that the third party shall be responsible for any content conversion from the format in which the Publisher provides the material.⁸³

The American Geophysical Union’s stance is that any third party is solely “for the purposes of long-term preservation of the Licensed Materials,” and the publisher’s server and local data loading are the perpetual access options, thus making a distinction between *preservation of* versus *access to* the subscribed content.⁸⁴

Only 1 of the studied publishers names their chosen third party in the license; as noted earlier, Elsevier’s is the National Library of the Netherlands.⁸⁵ There are not enough specifics in the studied licenses or enough of a track record with third-party access to know if this is a valid alternative to paying the publisher an annual maintenance fee, or rather just an emergency mirror site for such paid access. Care must be taken in negotiation to clarify the exact role of the third party and whether the library would really get limited access only in an emergency situation, or true perpetual access based solely on one’s prior subscription history.

A minority of the publishers who grant some form of perpetual access rights present more than one future option for access (34.4 percent). Presumably this is at their discretion, not the library’s. For example, Wiley says that it will:

provide the Licensee with access to the full text of the Licensed Electronic Journals published during the Term of this License, *either* by continuing

Table 3. Granting publishers’ contractual options for perpetual access (n=32)

	Publisher	Library	Third party	Unclear	Multiple options
Commercial publishers	9.0	7.0	4.0	5.0	5.0
% of all granting publishers	28.1	21.9	12.5	15.6	15.6
% within this publisher type	50.0	38.9	22.2	27.8	27.8
Society publishers	8.0	10.0	3.0	1.0	6.0
% of all granting publishers	25.0	31.3	9.4	3.1	18.8
% within this publisher type	57.1	71.4	21.4	7.1	42.9
Total	17.0	17.0	7.0	6.0	11.0
% of all granting publishers	53.1	53.1	21.9	18.8	34.4

online access to the same material on Wiley's server or by means of an archival copy in the electronic medium selected by Wiley (emphasis added).⁸⁶

Sometimes the choice of format is at the library's discretion, as with APA:

APA is committed to providing customers options for delivery for site licenses. Currently, those options include customer loading, access through several vendors, or access directly from APA, and there is a separate cost for delivery. The customer who has stopped paying for an annual site license may also choose one of these options for the segment of content for which they retain rights.⁸⁷

Several publishers leave the format intentionally unclear. This is more publishers (18.8 percent) than offer a third party option. This uncertainty is understandable, given how quickly technology becomes obsolete. Slightly more commercial publishers than society publishers are unclear as to format (27.8 percent versus 7.1 percent). For example, the IEEE suggests "a mutually agreeable media type," while BMJ offers simply that its journals will be available to lapsed subscribers "in electronic form."⁸⁸

Full functionality is yet another concern, no matter which party is hosting the content. The Royal Society of Chemistry (RSC), despite having the most convincing assurance with regard to continuing access and multiple formats, remains troubling because access to the hypertext markup language (HTML) versions of its articles, with reference linking, is lost, and only portable document format (PDF) versions remain.⁸⁹

As with cost, this uncertainty as to format makes planning and thus budgeting difficult. Will any of these publishers decide that their own server is not an option and that the library must invest in its own infrastructure? Happily, for more than half of the studied publishers who allow perpetual access, the publisher's server may yet be an option.

Publisher Continuity Issues

Few publishers address what will happen if their business closes or if they are bought by another publisher—a concern of libraries in light of such recent events as Elsevier's purchase of Academic Press, Taylor and Francis's purchase of Dekker, and Springer's purchase of Kluwer. Elsevier guarantees that the National Library of the Netherlands testbed archive can be used by institutions in case the publisher is no longer able to provide ScienceDirect.⁹⁰ Elsevier is also the most explicit of the studied licenses regarding the potential transfer of ownership of specific titles to new publishers; in such a scenario, Elsevier will try to secure

access rights to subscribed issues for existing subscribers from the new publisher.⁹¹

RSC commits: "If rights in all or any part of Publisher Content are assigned to another publisher, Publisher shall use its best endeavours to ensure that the terms and conditions of this Agreement are maintained."⁹² This expression of good intentions is well-meaning, but a publisher who sells a title or is purchased outright has little power to compel the new owner to honor its previous agreements.

Vagueness of Wording

Librarians have rightly noted the vagueness and unenforceable nature of the contractual assurances in this area. Some illustrative cases:

Licensor and Licensee shall discuss a mechanism satisfactory to the Licensor and Licensee to enable the Licensee to have access . . . and the terms of such access.⁹³

AIP will use reasonable efforts to retain in an archive all electronic information published by the American Institute of Physics.⁹⁴

What constitutes a *discussion*? What is a *reasonable* effort to retain an archive? Just because an archive has been retained, is it *reasonable* for a past subscriber to expect access to the archive? In fairness, librarians take advantage of vague wording themselves to avoid unrealistic obligations or legal risk (e.g., the reliance on the word *reasonable* to characterize the extent of library efforts to inform users of license terms). Libraries need to take note when publishers use the same tactic in areas where libraries need rights that will stand up in court.

OUP's license has the most striking example of vague wording:

On expiry of the Subscription Period, the Licensee shall be entitled to continue to exercise the non-exclusive rights granted herein (subject to the terms and conditions hereof) but only in respect of Material published for the first time during the Subscription Period. Nothing in this subclause requires the Licensor to continue to host the Material on its servers after the expiry of the Subscription Period or to make the Material available in any other form to the Licensee.⁹⁵

To what are libraries legally entitled here? Arguably, nothing. The second sentence negates any entitlement in the first.

Ovid's contract language demonstrates that, while one may be able to obtain a general assurance about perpetual

access to a journal package, in cases where multiple publishers are part of an aggregator package a provider may be unable to consistently assure perpetual access at the journal title level:

Ovid's contracts with *nearly all* journal publishers state that when an institution subscribes to a journal through Journals@Ovid, the institution then has the right to perpetual access to that journal. The *majority* of the journals offered through Journals@Ovid include this provision. *We are optimistic that the few outstanding publishers will soon officially adopt this policy as well* (emphasis added).⁹⁶

Further down in the contract is the caveat that "Ovid's archival access may be subject to change without notice," which calls the whole issue of the enforceability of perpetual access guarantees into question.⁹⁷ At a minimum, one should negotiate the standard "this contract is amendable only in writing signed by both parties" clause, so that one's concerns have a chance to be adequately addressed.

While initially unsettling because the wording appears to favor the publisher, such vagueness in license agreements is understandable because of the flux in both the perpetual access arena (noted earlier in Haworth's license) and the technology. Such wording can also be favorable to the library. Open-ended wording can give the library an opportunity to negotiate more specific terms in the future, by which time issues of trust, costs for libraries and economic security for publishers, and technology standards may be closer to resolution.

Conclusion

To obtain useful contractual guarantees of ongoing access to purchased electronic content, librarians must work with content providers and communicate with peers to create and develop robust license language and stable options and procedures for perpetual access to subscribed material. The primary finding of this license survey is encouraging. The majority of individual publishers studied are willing to grant libraries perpetual access rights in some form, often continuing access through the publisher's server.

If libraries accept e-journal licenses in their generic form (often with no guaranteed perpetual access clause) and sign them without negotiation, libraries and their patrons risk losing access in the future should the online subscription be canceled. What seems unimaginable now might be very real a few budget cycles into the future. This conclusion may seem like 20/20 hindsight and not an option for those who have already signed restrictive licenses. Nevertheless,

annual renewals, changes in subscription levels (e.g., from a complete package to a limited, title-by-title package), entry into a new consortial deal, and publisher notifications that they have revised their license for subscribers present opportunities to renegotiate a license.

The experience of the authors' institutions shows that perpetual access rights can be gained by asking, and that the effort can pay dividends for users. The University of Minnesota has availed itself of Elsevier's perpetual access rights clause. When the university canceled some ScienceDirect titles in 2003, the library was left with 1998–2002 access to these titles according to its license. The titles are not just linked with specific coverage dates in the library's MNCAT catalog, but are also activated in its link resolver program for those dates, so users can get to the titles either directly through ScienceDirect, or indirectly through other databases that cite articles in them. This example shows the importance to users of perpetual access to a canceled title. Library users do not care at the moment of need whether a library has a current subscription to a title; their concern is whether they can access the article they need right at that moment.

Anecdotal evidence from licensing librarians, such as the postings to ERIL-L and Liblicense cited earlier, suggests that some publisher sales representatives decline license requests for perpetual access on the basis that the demand from other institutions does not exist. More institutions asking for perpetual access rights can only help libraries' collective case. But, as standard negotiation wisdom agrees, one has to be willing to walk away from a deal. Academic libraries proved the potential efficacy of this approach if employed in significant numbers when, in 2001, Nature Publishing Group proposed an embargo on current issues of *Nature* (and subsequently offered a revised license without the embargo).⁹⁸ After the license review for this paper was completed, *Nature* reversed itself and began offering perpetual access rights to lapsed subscribers for an annual access fee, in what may be in part a response to the UC's aforementioned unwillingness to sign its renewal without such assurances.⁹⁹

If libraries are successful in negotiating perpetual access clauses in e-journal licenses, they still may be faced with uncertain ongoing costs to maintain such access—perhaps substantial costs, even if they are not annual content subscription fees. Libraries also may be faced with the need to load and serve up the content locally, a practice most libraries have been trying to discontinue since at least the late 1990s. If libraries cancel print journals carried in large full-text aggregator databases, they risk losing online access in the future, as few of these databases analyzed here allow for perpetual access upon termination in their agreements. Due to the varying coverage between aggregators, switching between large e-journal databases, such as Expanded

Academic Index and Academic Search Premier, can result in a significant loss of full-text journal coverage.

A research library would be wise not to rely too much on the prospect of obtaining once-subscribed articles through ILL. As with perpetual access rights, many libraries do not negotiate or succeed in inserting language protecting their ILL rights in licenses, resulting in what some have termed the “death of ILL”—institutions unable to loan articles from licensed databases due to contractual restrictions.¹⁰⁰ If more institutions go e-only via licenses with ILL prohibitions, such restrictions on lending will affect access for peer institutions. A 2003 survey of ILL offices in CIC and their e-resource licenses found that most CIC member libraries agreed that ILL rights were “important but not the first consideration” and that most libraries encountered cases when “ILL was prohibited entirely [and] when the delivery mechanism was restricted, such as in electronic delivery.”¹⁰¹

The recent research report from CLIR, *The Non-subscription Side of Periodicals*, proposed that perpetual access can be paid for, at least in part, by the cancellation of print, and the resulting savings to be accrued from the nonsubscription side of periodical operations (e.g., the binding, shelving, and maintenance of print volumes).¹⁰² CLIR rightly warned, though, that “some provosts might argue that savings should be returned to the general fund rather than be redirected within the library,” so the library will need to educate administrators and funders about the real costs of continued electronic access, including setup and maintenance fees for the publisher or third party and local data loading costs.¹⁰³ Note, however, that CLIR did not mention the fact that many publishers do not offer perpetual access and, upon cancellation of the print copy from such a publisher, no guaranteed copy would be available.

The International Coalition of Library Consortia (ICOLC) declared that, as “permanent access and archiving are of paramount importance,” licenses “must include cost-effective provisions to purchase and not just to lease or provide temporary access” and recommends that libraries and consortia “explore new options,” such as working with publishers to create national and regional repositories.¹⁰⁴ The question is whether libraries and consortia adopting this statement will interpret *must* literally and refuse to sign agreements without such guarantees.

Publishers cannot be expected to solve the problem of perpetual access on their own, but they are participating in the definition of roles, however unclear those roles currently are. Blackwell representatives believe that “the ‘archival’ role of the publisher is ill-defined but shifting from finite to that of open-ended responsibility”; they define appropriate publisher roles as ensuring that content is available in the “short to medium-term,” protecting subscribers from “harmful changes,” and “[working] with key stakeholders” to

make content available in the long-term.¹⁰⁵ Libraries are one of these stakeholders and should partner with publishers as ICOLC recommends.

While institutional repositories are important developments, they alone cannot be expected to solve the problem either, at least in the short term. Even if faculty can be convinced to deposit their research output in such repositories (perhaps in addition to the versions available at commercial journal sites), these projects are still in their infancy and need to be tested. In a recent overview of Ohio State University’s (OSU) institutional repository project, the Knowledge Bank, OSU’s director of libraries Branin projected that the project “will likely be underway for five to ten years before it is mainstreamed or institutionalized at the University.”¹⁰⁶

Concluding Recommendations

Drawing on the results of this license analysis and on the review of the literature on perpetual access, the authors offer the following recommendations:

1. Libraries should not assume that they may safely cancel the print version of a journal published by a society publisher and retain online access to subscribed issues if the electronic version is later canceled. Based on the research reported in this paper, society publishers are not more likely to provide perpetual access to subscribed articles than are commercial publishers.
2. Libraries should not assume that they can safely cancel the print version of a journal included in a full-text aggregator database and retain access.
3. Libraries should consider making the lack of perpetual access rights a deal breaker (i.e., a valid justification for not signing an e-journal license). This is especially important if the library is also canceling a print version. The University of Maryland and the University of California–Berkeley (cited in the introduction to this paper) offer helpful precedents for large research libraries. The same access terms should apply during perpetual access as during subscription access (e.g., fair use, rights of walk-in users, remote access). This may be a difficult stance to take for journals that are considered so central to a collection that one would never cancel them. Holding firm for the inclusion of perpetual access rights is where consortia, which broker many of an academic library’s major e-journals, can bring their considerable economic leverage to bear.
4. Libraries should find a way to budget for perpetual access. These can be one-time purchase fees (e.g., back files from Annual Reviews or Wiley) or annual access or maintenance fees (e.g., established packages, such as the ACS Journal Archives and JSTOR), or might be

a combination of the two (e.g., third parties such as Portico). Care must be taken to ensure perpetual access rights to current issues in the future even if a back-file package is separately purchased. Many packages only go back to the mid- to late-1990s, and back files are not automatically added to one's package but must be licensed separately and are not always available as one-time purchases. Librarians must recognize that such fees are often justifiable requirements from publishers. As libraries give up the binding, shelving, and maintenance costs of print journals, publishers of e-journals have delivery costs that must be borne. Related to the cost issue, the profession should advocate for and participate in the development of standardized, affordable pricing models for such access. Here, too, consortial purchasing power can play a role.

5. Libraries should explore the potential for trusted third parties to host journals for libraries that have subscribed to them in the past.
6. Libraries should include wording about publisher mergers in perpetual access clauses.
7. Libraries should ask their legal counsel whether the publisher's perpetual access clause is clear and specific enough to stand up in court.

References

1. Mary M. Case, "A Snapshot in Time: ARL Libraries and Electronic Journal Resources," *ARL Bimonthly Report*, no. 235 (2004), www.arl.org/newsltr/235/snapshot.html (accessed Sept. 15, 2005). See also Michael Litchfield, "ATG Annual Survey Report," *Against the Grain* 15, no. 1 (2003): 53–55; Roger C. Schonfeld et al., *The Nonsubscription Side of Periodicals: Changes in Library Operations and Costs between Print and Electronic Formats*, CLIR Report No. 127 (Washington, D.C.: Council on Library and Information Resources, 2004), www.clir.org/pubs/reports/pub127/pub127.pdf (accessed Sept. 15, 2005).
2. Vincent Kiernan, "Professors Are Unhappy with Limitations of Online Resources, Survey Finds," *Chronicle of Higher Education* (Apr. 30, 2004), <http://chronicle.com> (accessed Sept. 15, 2005). See also Kevin Guthrie and Roger C. Schonfeld, "What Do Faculty Think of Electronic Resources?" (presentation, semiannual Task Force Meeting of the Coalition for Networked Information, Alexandria, Va., Apr. 16, 2004), www.cni.org/tfms/2004a.spring/presentations/CNI_Guthrie_What.ppt (accessed Sept. 15, 2005).
3. Kevin M. Guthrie, "Lessons from JSTOR: User Behavior and Faculty Attitudes," *Journal of Library Administration* 36, no. 3 (2002).
4. *Ibid.*, 118.
5. Arthur Curley and Dorothy Broderick, *Building Library Collections*, 6th ed. (Metuchen, N.J.: Scarecrow, 1985); See also Peter Graham, "Building Research & Action Agendas for Digital Archiving," in *Preservation of Digital Information: Association of Research Libraries Proceedings of the 131st Meeting*, ed. Gloria Werner, Jaia Barrett, and Karen A. Wetzel (Washington, D.C.: Association of Research Libraries, 1997), www.arl.org/arl/proceedings/131/graham.html (accessed Sept. 15, 2005).
6. Ross Harvey, "The Preservation of Electronic Records: What Shall We Do Next?" in *Collection Management for the 21st Century: A Handbook for Librarians*, ed. G. E. Gorman and Ruth H. Miller (Westport, Conn.: Greenwood, 1997), 181; Association of Research Libraries, *Strategic Plan 2005–2009* (Washington, D.C.: Association of Research Libraries, 2004), 8, www.arl.org/arl/strategic_taskforces/strategic_planning/ARLPlan.pdf (accessed Sept. 15, 2005).
7. Angela Riggio et al., *Report of the Digital Library Federation Electronic Resource Management Initiative*, Appendix D: Data Element Dictionary (Washington, D.C.: Digital Library Federation, 2004), 26, www.diglib.org/pubs/dlfermi0408/dlfermi0408appd.pdf (accessed Sept. 15, 2005).
8. Tony Millett, comp., "Survey of Database Providers' Policies Regarding Continued Access to Full-Text Databases for the Period Subscribed to Following Cancellation of Subscriptions" (working paper, Council of Australian University Librarians, Canberra, May 25, 2001), www.caul.edu.au/datasets/ElectronicJournalContinuedAccess.doc (accessed Sept. 15, 2005).
9. *Ibid.*
10. Janet A. Hughes, "Issues and Concerns with the Archiving of Electronic Journals," *Science & Technology Libraries* 22, no. 3/4 (2002): 131.
11. Thomas A. Peters, "Collaborative Print Retention Pilot Projects," *Against the Grain* 15, no. 2 (Apr. 2003): 18, 20; Francine Fialkoff, ed., "CIC Libraries in Deal with Wiley, Springer to Archive Print Journals," *Library Journal Academic NewsWire: The Publishing Report* (May 5, 2005), www.libraryjournal.com (accessed Sept. 15, 2005) [subscription-based electronic news service]; Systemwide Operations and Planning Group (SOPAG), "SOPAG Meeting, October 25, 2002, Action Minutes," <http://libraries.universityofcalifornia.edu/sopag/min102502.html> (accessed Sept. 15, 2005).
12. Peters, "Collaborative Print Retention Pilot Projects."
13. John McDonald, "No One Uses Them So Why Should We Keep Them?" Scenarios for Print Issue Retention," *Against the Grain* 15, no. 2 (Apr. 2003): 24.
14. Fialkoff, "CIC Libraries in Deal with Wiley, Springer to Archive Print Journals."
15. Maggie Jones, *Archiving E-Journals Consultancy—Final Report* (London: Joint Information Systems Committee, 2003), www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf (accessed Sept. 15, 2005); Christine Maher, David Groenewegen, and Gail James, "Preservation of Access to Subscription Electronic Journals in Australian University Libraries: A Discussion Paper for the CAUL Electronic Information Resources Committee" (Council of Australian University Librarians, Canberra, Australia, Nov. 2002), www.caul.edu.au/caul-doc/AccessPreservation_Nov2002.doc (accessed Sept. 15, 2005); Neil Beagrie, "The Continuing Access and Digital Preservation Strategy for the UK Joint Information Systems Committee (JISC)," *D-Lib Magazine* 10, no. 7/8 (July/Aug. 2004), www.dlib.org/dlib/july04/beagrie/07beagrie.html (accessed Sept. 15, 2005); Catherine Ayre

- and Adrienne Muir, *Copyright and Licensing for Digital Preservation—Final Report of the Copyright and Licensing for Digital Preservation Project* (Leicestershire, England: Loughborough University, 2004), www.lboro.ac.uk/departments/ls/disresearch/CLDP/Project_reports.htm (accessed Sept. 15, 2005).
16. John Garrett and Donald Waters, *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* (Mountain View, Calif.: Commission on Preservation and Access, Research Libraries Group, May 1, 1996), www.rlg.org/legacy/ftp/pub/archtf/final-report.pdf (accessed Sept. 15, 2005).
 17. Linda Cantara, ed., *Minimum Criteria for an Archival Repository of Digital Scholarly Journals, Version 1.2* (Washington, D.C.: Digital Library Federation/Council on Library and Information Resources, May 15, 2000), www.diglib.org/preserve/criteria.pdf (accessed Sept. 15, 2005).
 18. Linda Cantara, introduction to *Archiving Electronic Journals: Research Funded by the Andrew W. Mellon Foundation* (Washington, D.C.: Digital Library Federation/Council on Library and Information Resources, 2003), 4, www.diglib.org/preserve/introduction.pdf (accessed Sept. 15, 2005).
 19. Harvard University Library Mellon Project Steering Committee, Harvard University Library Mellon Project Technical Team, *Report on the Planning Year Grant for the Design of an E-journal Archive* (Washington, D.C.: Digital Library Federation/Council on Library and Information Resources, Apr. 1, 2002), 11–12, www.diglib.org/preserve/harvardfinal.pdf (accessed Sept. 15, 2005).
 20. Yale University Library and Elsevier Science, *YEA: The Yale Electronic Archive: One Year of Progress* (Washington, D.C.: Digital Library Federation/Council on Library and Information Resources, Feb. 2002), 13, www.diglib.org/preserve/yalefinal.pdf (accessed Sept. 15, 2005).
 21. Ibid., 61.
 22. John Mark Ockerbloom, *Report on a Mellon-Funded Planning Project for Archiving Scholarly Journals* (Washington, D.C.: Digital Library Federation/Council on Library and Information Resources, Sept. 16, 2002), 19, www.diglib.org/preserve/upennfinal.pdf (accessed Sept. 15, 2005).
 23. Yale University Library and Elsevier Science, *YEA*.
 24. Ibid., 29.
 25. Cantara, *Archiving Electronic Journals*.
 26. Stanford University Libraries, *LOCKSS: A Distributed Digital Archiving System* (Washington, D.C.: Digital Library Federation/Council on Library and Information Resources, Oct. 8, 2002), www.diglib.org/preserve/stanfordfinal.html (accessed Sept. 15, 2005).
 27. Victoria Reich, "Lots of Copies Keep Stuff Safe As a Cooperative Archiving Solution for E-journals," *Issues in Science & Technology Librarianship* 36 (Fall 2002), www.isl.org/02-fall/article1.html (accessed Sept. 15, 2005).
 28. LOCKSS Program, "About LOCKSS," Stanford University, <http://lockss.stanford.edu/about/about.htm> (accessed Sept. 15, 2005).
 29. LOCKSS Program, "LOCKSS Alliance," Stanford University, <http://lockss.stanford.edu/alliance/alliance.htm> (accessed Sept. 15, 2005).
 30. Portico, www.portico.org (accessed Feb. 10, 2006); Ithaka, www.ithaka.org (accessed Feb. 10, 2006).
 31. Ithaka Harbors, "The Archiving Problem," www.ithaka.org/e-archive/problems.htm (accessed Sept. 15, 2005).
 32. Ithaka Harbors, "Sustaining the Archive," www.portico.org/about/sustain.html (accessed Sept. 15, 2005).
 33. Ithaka Harbors, "E-Archive Approach," www.ithaka.org/e-archive/approach.htm (accessed Sept. 15, 2005).
 34. Ithaka Harbors, "Key Archival Terms," www.portico.org/terms.html (accessed Sept. 15, 2005).
 35. Reich, "Lots of Copies Keep Stuff Safe As a Cooperative Archiving Solution for E-journals."
 36. University of York, Information Committee, Library Committee, "Electronic Only Access to Journals Trial Final Report" (Jan. 2005), 6, www.york.ac.uk/services/library/lib-docs/e-journalfinalreport0502.pdf (accessed Sept. 15, 2005); William H. Walters, "Criteria for Replacing Print Journals with Online Journal Resources: The Importance of Sustainable Access," *Library Resources & Technical Services* 48, no. 4 (Oct. 2004): 300–4; Marianne Stowell Bracke and Jim Martin, "Developing Criteria for the Withdrawal of Print Content Available Online," *Collection Building* 24, no. 2 (2005): 61–64; University of Maryland, University Libraries, "Policy on Collecting Only Electronic Versions of Journals," www.lib.umd.edu/CLMD/e-verpol.html (accessed Sept. 15, 2005); Joan Emmet, NorthEast Regional Libraries Consortium, "NERL Principles for Electronic Journal Licenses" (Nov. 30, 2004), www.library.yale.edu/NERLpublic/EJrnlPrinciples.html (accessed Sept. 15, 2005); Regents of the University of California, "Checklist of Points to be Addressed in a California Digital Library License Agreement," California Digital Library, www.cdlib.org/vendors/checklist.html (accessed Sept. 15, 2005); NESLi2, "Model NESLi2 Licence for Journals" (2005), www.nesli2.ac.uk/NESLi2Licence_new_final280405.htm (accessed Sept. 15, 2005); Canadian National Site Licensing Project, "Canadian National Site Licensing Project License Agreement," Canadian Research Knowledge Network (Feb. 12, 2001), 10–11, www.crkn.ca/pr/achievements/CNSLP-License-12Feb01.pdf (accessed Sept. 15, 2005).
 37. Regents of the University of California, "Checklist of Points to be Addressed"; NESLi2, "Model NESLi2 Licence for Journals"; Canadian National Site Licensing Project, "Canadian National Site Licensing Project License Agreement."
 38. University of York, Information Committee, Library Committee, "Electronic Only Access to Journals Trial Final Report," 6; NESLi2, "Model NESLi2 Licence for Journals"; Joan Emmet, "NERL Principles for Electronic Journal Licenses."
 39. University of Maryland, University Libraries, "Policy on Collecting Only Electronic Versions of Journals."
 40. Regents of the University of California, "Challenges to Licensing from Some Publishers," California Digital Library, www.cdlib.org/news/barriers.html (accessed Sept. 15, 2005).
 41. Elizabeth Winters, "Perpetual Access," e-mail to the eril-l@listserv.binghamton.edu mailing list, Sept. 9, 2004, <http://listserv.binghamton.edu/archives/eril-l.html> [list archives] (accessed Sept. 15, 2005); Eric Lease Morgan, "Encourage

- Publishers to Make Their Content 'LOCKSS-able,'" e-mail to the liblicense-l@lists.yale.edu mailing list, June 9, 2005, www.library.yale.edu/~llicense/ListArchives/ [list archives] (accessed Sept. 15, 2005).
42. Edward Shreeves, "Selectors, Subject Knowledge, and Digital Collections," in *Improved Access to Information: Portals, Content Selection, and Digital Information*, ed. Sul H. Lee (Binghamton, N.Y.: Haworth, 2003), 69–70.
 43. Case, "A Snapshot in Time." See also Lynn N. Wiley, "License to Deny? Publisher Restrictions on Document Delivery from E-licensed Journals," *Interlending & Document Supply* 32, no. 2 (2004): 96.
 44. Case, "A Snapshot in Time."
 45. Ibid.
 46. Nicholas Lewis, "Are We Burning Our Boats?" Survey on Moving to Electronic-only," *SCONUL Newsletter* 31 (Spring 2004), www.sconul.ac.uk/pubs_stats/newsletter/31/19.pdf (accessed Sept. 15, 2005).
 47. Ibid., 61.
 48. Bracke and Martin, "Developing Criteria for the Withdrawal of Print," 61.
 49. McDonald, "No One Uses Them," 24.
 50. Linda L. Eells, "For Better or for Worse: The Joys and Woes of E-Journals," *Science & Technology Libraries* 25, no. 1/2 (2004): 33–53.
 51. Gwen Bird and Andrew Waller, "We Own It: Dealing with Perpetual Access in Big Deals" (presentation, North American Serials Interest Group Annual Conference, Minneapolis, Minn., May 20, 2005).
 52. Maggie Jones, "Continuing Access and Long-term Preservation" (presentation, Association of Learned and Professional Society Publishers Seminar: Is It Time to Drop Print? Oxford, England, Feb. 2, 2005), www.alpsp.org/events/previous/jon020205.ppt (accessed Sept. 15, 2005); Timothy D. Jewell, *Selection and Presentation of Commercially Available Electronic Resources: Issues and Practices*, CLIR Report no. 99 (Washington, D.C.: Council on Library and Information Resources, 2001), www.clir.org/pubs/reports/pub99/pub99.pdf (accessed Sept. 15, 2005).
 53. Carol Pitts Diedrichs, "E-journals: The OhioLINK Experience," *Library Collections, Acquisitions & Technical Services* 25, no. 2 (2001): 196.
 54. Pamela Grudzien, "CDL: DLT Tapes from Elsevier," e-mail to the colldv-l@usc.edu mailing list, July 29, 2004, www.infomotions.com/serials/colldv-l/04/0169.shtml (accessed Sept. 15, 2005).
 55. Joep Verheggen, "Digital Archiving at Elsevier" (presentation, International Council for Scientific and Technical Information Public Conference: Technical and Economic Challenges of Scientific Information [STM Content Access, Linking and Archiving], London, May 17, 2004), www.icsti.org/forum/46/Joep_Verheggen.ppt (accessed Sept. 15, 2005).
 56. Erik Oltmans, "The e-Depot at the Koninklijke Bibliotheek: Building a Digital Depository to Meet the Needs of a National Library" (Research Libraries Group Forum: To Have and to Hold: Metadata & Institutional Repositories, The Hague, NL, May 18 2004), 4, www.rlg.org/en/pdfs/2004membforum2/oltmans.pdf (accessed Sept. 15, 2005); also see Maggie Jones, "Continuing Access and Long-term Preservation"; Valerie Johns, "SAGE Publications Partners with KB, the National Library of the Netherlands, for Permanent Digital Archiving," e-mail to the liblicense-l mailing list, Apr. 30, 2005, www.library.yale.edu/~llicense/ListArchives/0505/msg00017.html (accessed Sept. 15, 2005).
 57. Oltmans, "The e-Depot at the Koninklijke Bibliotheek."
 58. *ODLIS Online Dictionary for Library and Information Science*, s.v. "aggregator" (by Joan M. Reitz), Libraries Unlimited, <http://lu.com/odlis> (accessed Sept. 15, 2005).
 59. Trisha Davis, "Terms and Conditions," in *License Review and Negotiation: Building a Team-Based Institutional Process* (Atlanta, Ga.: Association of Research Libraries Office of Scholarly Communication, 2001), 32.
 60. Ibid.
 61. Patricia Brennan, Karen Hersey, and Georgia Harper, "Strategic and Practical Considerations for Signing Electronic Information Delivery Agreements," Association of Research Libraries, <http://arl.cni.org/scomm/licensing/licbooklet.html> (accessed Sept. 15, 2005). See also Riggio et al., *Report of the Digital Library Federation Electronic Resource Management Initiative*, 26; Karen Coyle, "Digital Rights Management—Part 4," Writings on the Digital Age, www.kcoyle.net/drm_basics4.html (accessed Sept. 15, 2005).
 62. Walter de Gruyter, "Terms and Conditions of Use for Online Access to Walter de Gruyter Journals," Accessed Sept. 15, 2005, www.degruyter.de/journals/terms_e.html.
 63. Association for Computing Machinery, "Terms and Conditions for Institutional Use of ACM Digital Library," www.minitex.lib.umn.edu/cpers/ACM/license.pdf (accessed Sept. 15, 2005).
 64. The Haworth Press, "The Haworth Press Online Terms," www.haworthpress.com/pdfs/Single-SiteLicense.pdf (accessed Sept. 15, 2005).
 65. Saskia Volkers, "Re: License for AMA E-journals," e-mail to Jim Stemper, Dec. 5, 2003.
 66. Nature Publishing Group, "Nature Academic Licence Agreement Schedule," http://npg.nature.com/npg/servlet/Content?data=xml/07_lic.xml&style=xml/07_lic.xml#site (accessed Sept. 15, 2005).
 67. Springer, "SpringerLink Licensing Options for Library Administrators 2005," www.springeronline.com/sgw/cda/frontpage/0,11855,4-117-2-128726-0,00.html (accessed Sept. 15, 2005); Online Computer Library Center, "Archival rights," www.oclc.org/electroniccollections/archiving/rights (accessed Sept. 15, 2005); S. Karger AG, "Terms and Conditions for Use of Karger Online Publications by Academic Subscribers," http://content.karger.com/copy/sitelicense_4.asp (accessed Sept. 15, 2005).
 68. American Institute of Physics, "Scitation Institutional Single-Site User License," http://scitation.aip.org/jhtml/scitation/UserLic_Inst_2004-01.pdf (accessed Sept. 15, 2005).
 69. American Psychological Association, "APA Archiving and Access Policy For PsycARTICLES," www.apa.org/psycarticles/archiving.html (accessed Sept. 15, 2005).
 70. Ovid Technologies, "Journals@Ovid Electronic Archive Policy," www.ovid.com/site/products/journals_archive_policy.jsp (accessed Sept. 15, 2005).

71. Wiley Subscription Services, "Basic Access License for Wiley InterScience," http://www3.interscience.wiley.com/spreadsheet_documents/sample_ballicense.pdf (accessed Sept. 15, 2005).
72. Institute of Electrical and Electronics Engineers, "IEEE Online Products License Agreement," www.ieee.org/products/onlinepubs/licenses/OnlineLicenseAgrmtAcademic.doc (accessed Sept. 15, 2005).
73. "License Agreement, China Academic Journals Web Database," East View Publications, signed by Cindy Clennon, Committee on Institutional Cooperation, July 1, 2002.
74. Blackwell Publishing, "Blackwell Publishing Site License Agreement," www.blackwellpublishing.com/license/license.pdf (accessed Sept. 15, 2005); American Geophysical Union, "Academic: Single Institution License (Exclusive of 2-year Colleges)," www.agu.org/pubs/institution_forms/Academic_non_2-year.pdf (accessed Sept. 15, 2005).
75. American Society for Microbiology, "Institutional Online Site License Agreement for Full-Text Access," www.journals.asm.org/subscriptions/sitelicense.shtml (accessed Sept. 15, 2005).
76. American Association for the Advancement of Science, "Science Online Publications Institutional Site-Wide Subscription Agreement," www.sciencemag.org/subscriptions/terms-unlim.dtl (accessed Sept. 15, 2005).
77. JSTOR, "JSTOR Archive License Agreement," www.jstor.org/about/license.DV.9.1.05.pdf (accessed Sept. 15, 2005).
78. Springer, "SpringerLink Licensing Options for Library Administrators 2005."
79. American Institute of Physics, "Scitation Institutional Single-Site User License."
80. Project MUSE, "Terms and Conditions of Project MUSE Subscriptions," <http://muse.jhu.edu/about/muse/terms.html> (accessed Sept. 15, 2005).
81. LOCKSS Program, "About: Current LOCKSS Publishers and Titles," Stanford University, <http://lockss.stanford.edu/about/titles.htm> (accessed Sept. 15, 2005).
82. Ithaka Harbors, "E-Archive in the Community," www.ithaka.org/e-archive/community.htm#Publishers (accessed Sept. 15, 2005).
83. Blackwell Publishing, "Blackwell Publishing Site License Agreement," www.blackwellpublishing.com/license/license.pdf (accessed Sept. 15, 2005).
84. American Geophysical Union, "Academic: Single Institution License (Exclusive of 2-year Colleges)."
85. "Elsevier License Agreement," Elsevier B.V., signed by Wendy Pradt Lougee, Univ. of Minnesota, Dec. 13, 2004.
86. Wiley Subscription Services, "Basic Access License for Wiley InterScience." See also Internet-Journals, "Berkeley Electronic Press Journals Order Form," www.bepress.com/bepress_order_form.pdf (accessed Sept. 15, 2005); BioOne, "BioOne Subscriber License," www.bioone.org/images/sub-license.pdf (accessed Sept. 15, 2005); Ovid Technologies, "Journals@Ovid Electronic Archive Policy."
87. American Psychological Association, "APA Archiving and Access Policy For PsycARTICLES."
88. Institute of Electrical and Electronics Engineers, "IEEE Online Products License Agreement"; BMJ Publishing Group, "Online Journal Single Site Licence Agreement," www.bmjournals.com/subscriptions/sslicense.shtml (accessed Sept. 15, 2005).
89. Royal Society of Chemistry, "Electronic Access Licence Agreement," <http://chemistry.rsc.org/Publishing/librarians/LicensingRegulations.asp> (accessed Sept. 15, 2005).
90. "Elsevier License Agreement."
91. Daviess Menefee, "Elsevier Science Archiving Policy When Journals Are Sold or Transferred," e-mail to the liblicense-l@lists.yale.edu mailing list, July 18, 2002, www.library.yale.edu/~llicense/ListArchives/0207/msg00079.html (accessed Sept. 15, 2005).
92. Royal Society of Chemistry, "Electronic Access Licence Agreement."
93. The publisher is not named because of a confidentiality requirement in the negotiated agreement.
94. American Institute of Physics, "Scitation Institutional Single-Site User License."
95. Oxford University Press, "Oxford University Press Online Journals: Institutional Online Agreement," www.oxfordjournals.org/help/instsitelicense.pdf (accessed Sept. 15, 2005).
96. Ovid Technologies, "Journals@Ovid Electronic Archive Policy."
97. Ibid.
98. Bernd-Christoph Kämper, "Nature—What Other Libraries Say," Stuttgart University Library, www.ub.uni-stuttgart.de/ejournals/Nature_andere_Univ.html (accessed Sept. 15, 2005).
99. Nature Publishing Group, "NPG Introduces Post-cancellation Rights to Licensed Web Content," e-mail to Jim Stemper, Sept. 28, 2005.
100. Kenneth Frazier, "Collection Development and Professional Ethics," *Journal of Library Administration* 28, no. 1 (1999): 43.
101. Wiley, "License to Deny?" 95–96. See also Jeffrey C. Carrico and Kathleen L. Smalldon, "Licensed to ILL: A Beginning Guide to Negotiating E-Resources Licenses to Permit Resource Sharing," *Journal of Library Administration* 40, no. 1/2 (2004): 41–54; Mary E. Jackson, "Research Collections and Digital Information: Will There Be a Role for Interlibrary Loan and Document Delivery Services?" *Journal of Library Administration* 31, no. 2 (2000): 15–25.
102. Schonfeld et al., "The Nonsubscription Side of Periodicals."
103. Ibid., 39.
104. International Coalition of Library Consortia, *Statement of Current Perspective and Preferred Practices for Selection and Purchase of Electronic Information (Update No. 2, Pricing and Economics, October, 2004)* (New Haven, Conn.: Yale Univ., 2004), www.library.yale.edu/consortia/2004currentpractices.htm (accessed Sept. 15, 2005).
105. Paul Calow and Ian Bannerman, "A Publisher's Perspective" (presentation, Association of Learned and Professional Society Publishers Seminar: Archiving—Whose Problem Is It? [West Sussex, England, Nov. 29, 2002]) www.alpsp.org/events/previous/cal291102.ppt (accessed Sept. 15, 2005).
106. Joseph Branin, "Portals, Access, and Research Libraries," in *Improved Access to Information: Portals, Content Selection, and Digital Information*, ed. Sul H. Lee (Binghamton, N.Y.: Haworth, 2003), 53.

Appendix. Publishers and Aggregators Whose Licenses Were Analyzed

Commercial Publishers

Berkeley Electronic Press
 Blackwell
 BMJ Publishing Group
 Cold Spring Harbor Laboratory Press
 De Gruyter
 East View Information Services (China Academic Journals)
 Elsevier
 Emerald
 Haworth Press
 Hindawi Publishing Corporation
 Karger
 Lippincott Williams and Wilkins
 Mary Ann Liebert
 National Journal
 Nature Publishing Group
 Palgrave Macmillan
 SAGE Publications
 Springer (includes Kluwer)
 Taylor & Francis Group (includes Dekker)
 Wiley

Society Publishers

American Association for the Advancement of Science (AAAS)
 American Chemical Society
 American Geophysical Union
 American Institute of Physics*
 American Mathematical Society
 American Medical Association
 American Psychological Association
 American Society for Microbiology

American Society of Civil Engineers
 American Society of Mechanical Engineers
 Association for Computing Machinery
 CAB International
 Cambridge University Press
 IEEE
 Massachusetts Medical Society (New England Journal of Medicine)
 Oxford University Press
 Royal Society
 Royal Society of Chemistry
 University of California Press
 University of Chicago Press

Commercial Aggregators

EBSCO Industries
 Gale
 H. W. Wilson
 Ovid
 ProQuest

Society Aggregators

BioOne
 Highwire Press
 JSTOR
 OCLC (Electronic Collections Online)
 Project Muse

*American Institute of Physics (AIP) Scitation; the authors included AIP as an *individual* publisher (not an aggregator) because its hosted publishers have their own licenses (e.g., American Society of Mechanical Engineers)

Library of Congress Classification Numbers

Issues of Consistency and Their Implications for Union Catalogs

Bhagirathi Subrahmanyam

This study examined Library of Congress Classification (LCC)-based class numbers assigned to a representative sample of 200 titles in 52 American library systems to determine the level of consistency within and across those systems. The results showed that under the condition that a library system has a title, the probability of that title having the same LCC-based class number across library systems is greater than 85 percent. An examination of 121 titles displaying variations in class numbers among library systems showed certain titles (for example, multi-foci titles, titles in series, bibliographies, and fiction) lend themselves to alternate class numbers. Others were assigned variant numbers either due to latitude in the schedules or for reasons that cannot be pinpointed. With increasing dependence on copy cataloging, the size of such variations may continue to decrease. As the preferred class number with its alternates represents a title more fully than just the preferred class number, this paper argues for continued use of alternates by library systems and for finding a method to link alternate class numbers to preferred class numbers for enriched subject access through local and union catalogs.

As long ago as 1968 the possibility of searching for resources in online catalogs through assigned call numbers as access points was raised.¹ A call number is a notation that uniquely identifies an item within a collection and on the shelf. It is a composite number containing a class number that indicates the class that its subject belongs to as per a classification scheme, such as the Library of Congress Classification (LCC) scheme, and a book or an item number that uniquely identifies an item within that class, derived usually from the author's name and other physical features (e.g., volume number).² Today that possibility has been realized; many library catalogs provide call number searches. Some even allow searching by the class number part of the call number. For instance, OCLC's WorldCat, originally a cataloging resource, is becoming increasingly accessible to the public and allows searching and browsing by the class number part of the assigned call number. Moreover, networked resources are often assigned class numbers using the same classification schemes that are being used by the libraries to classify their resources.³

A searcher who chooses to search by call or class number can reasonably expect all institutions to have the same class number for a particular title, given that the same classification scheme was used as a source for class numbers, and only one class number was assigned, even for a multi-foci title. In this context, the objective of this study was to ascertain the degree of uniformity of class number assignment that exists within libraries and across different libraries and,

Bhagirathi Subrahmanyam (bsubrahm@kent.edu) is Assistant Professor, School of Library and Information Science, Kent State University, Columbus, Ohio. The author wishes to thank the editor, the reviewers, and Elaine Svenonius for their valuable comments and suggestions.

where deviations exist, to analyze the reasons for deviations and their implications for local and online catalog design in general, and in academic library catalogs in particular.

Background

The many advantages of call numbers in general, and class numbers in particular, as access points over author and title access points have been well-documented.⁴ Unlike vocabulary searching, call number searching is precise and unambiguous. Unless a cataloger provides for various vocabulary approaches, such as variant spellings and synonyms, a user may miss relevant items. In contrast, call numbers consist of unambiguous sequences of letters, numerals, or both. There are no two ways to represent the same class number. Thus, call numbers are more precise than vocabulary terms as search tools. Moreover, they are becoming more available as search tools as more and more network resources are being classified using general library classification schemes.

More importantly, call numbers serve as starting points for browsing shelves. Arrangement by call numbers on the shelves results in the collocation first by class numbers and then book numbers. The underlying principle of any general library classification scheme from which class numbers are derived is the principle of collocation, which aims to bring like things together by dividing the universe of knowledge hierarchically into classes and subclasses. Users who know a general class number can browse the shelf in any library without going to a catalog under the assumption that like items will be grouped together by class number. Such shelf browsing, from general to specific and vice versa, also can be simulated in an online catalog. A user can browse a catalog using either the call or class number of a known item wherever he or she is, at home or in a library. Such classificatory access helps overcome the problem of browsing a divided collection by simulating shelf display online, irrespective of the physical location of material. However, browsing will be meaningful in an online catalog only if the class number assignment for a given topic is consistent within and across libraries.

Contributing to inconsistency in the assignment of call or class numbers is that their collocation function is often overlooked in favor of using them as mark-and-park devices for shelf arrangement. Some materials (e.g., videotapes) may not be assigned subject-based class numbers. Libraries may find it difficult to keep up with changes that occur in classification schemes.⁵ As Taylor noted in a study on copy cataloging in libraries, many librarians find keeping up with the successive editions of classification schemes difficult, especially when changes between editions was drastic in the form of complete revisions or new editions.⁶ In order to keep pace with the changes in classification, libraries must have the resources—time, finances, and personnel—to ret-

respectively classify collections to reflect the changes as and when they occur. In the absence of such resources, libraries must reconcile themselves to giving items on the same topic two different class numbers. Older collections retain class numbers from an earlier edition of the scheme, while newly acquired materials are assigned class numbers from the latest edition of the scheme. This can result in inconsistency of class numbers across libraries.⁷

One would expect that centralized agencies that provide catalog entries to member libraries for copy cataloging, such as OCLC, and such programs as Cataloging in Publication (CIP) would help to standardize the process of class number assignment across libraries, but this is not always the case. Although OCLC records generally carry both LCC and Dewey Decimal Classification (DDC) numbers, records may be lacking one or both. For example, OCLC records created before 1968 in the field of law may lack Library of Congress call numbers because the Library of Congress schedule for law was not available then. In the absence of call numbers found on OCLC records, libraries have to assign their own call numbers, and these may vary from library to library.

Several studies have investigated the connection between copy cataloging and assigned call numbers. Younger and Nichols found that many libraries make minor adjustments to the call numbers available on cataloging copy for one or more of the following reasons: (1) to match the numbers of earlier collections; (2) to meet the special needs of a given segment of the collection; (3) to add Cutter numbers or prefixes; (4) to avoid shelf-list conflicts; (5) to bring like subjects together; (6) to make the copy fit the latest edition of the classification scheme; (7) to fix errors on the copy; and (8) to use alternative numbers that fit the local collection better.⁸ Circumstances such as these result in the separation of books on the same topic on the shelf. As a result, call numbers tend to become mere mark-and-park devices.

O'Neill and McCain, in a joint OCLC–Library of Congress project, conducted a study of library Cuttering practice to learn the extent and types of changes libraries made to call numbers, particularly Cutter numbers.⁹ Using a sample of 895 records created from October 1994 to May 1995, O'Neill and McCain compared the DDC call number from the OCLC record and the final call number assigned by the member library that used that record for copy cataloging. They concluded that almost 80 percent of the DDC class numbers assigned by the Library of Congress were accepted by the local library and another 10 percent were accepted with minor revisions.¹⁰

In another study, Massey and Malinconico examined the University of Alabama's originally assigned call numbers for shelflisting errors, such as wrong class or wrong Cutter number.¹¹ They examined items cataloged from October 1990 through March 1995 to determine whether the differences

in shelflisting errors between these call numbers and the class numbers copied from OCLC were significant enough to justify investment in original cataloging. They concluded that revising call numbers from a copy for local use was not a cost-effective operation, at least for their university. They did not study the nature of variations and the reasons for such variations.

Methods

The objectives of the present study were to answer the following two research questions:

1. Among academic library systems using the LCC scheme for classifying items, does a given title (where title can be defined as the distinguishing name of a work) have the same class number across different library systems?
2. What are the reasons for any observed variations in assigned class numbers for the same titles?

The study focused on the class number part of the call number, as it is generated from a standard classification scheme to represent the subject matter of the title under consideration. The book number part, which is derived locally by individual libraries and thus subject to a high degree of variation, was not considered. LCC-based class numbers were chosen because the LCC scheme is widely used in academic libraries and, as an enumerative scheme, provides little idiosyncrasy in assigning class numbers. Academic library systems were investigated because a majority of them use the LCC schemes and their catalogs were, for the most part, Internet-accessible.

Before collecting data, the following decisions were made. First, class numbers of a title would be compared with reference to a specific edition. That is, only class numbers of a title with a specific imprint (e.g., New York, MacMillan, 1958) would be compared across various library systems. Second, for each title in the sample of titles, the class number that occurred most frequently among libraries under investigation was defined as the standard class number (SCN). Third, each university library system, whether it has one or more than one branch, was considered to be a single entity. Finally, an imprint-specific class number was considered a match if and only if the title in all branches of a library system exhibited the same one. Even if the class number differed in only one branch, a non-match would be recorded.

The method research included the following steps. A list of 52 large, American university library systems that use LCC as the main scheme for their collections and whose online catalogs were available in the public domain was compiled from the list of the Association of Research

Libraries' member libraries (see appendix A). The current list has 113 university libraries in the United States and Canada; when this project was started in 1996, only 107 library systems were members. From that list of 107 library systems, systems in Canada and those not using LCC were eliminated. In the end, only 52 library systems whose catalogs were available on the Internet formed the library sample. The possibility of finding a given title in a large library system is higher than in a smaller library.¹²

The six-volume *Books for College Libraries* was used as a source for monograph titles to be sampled for this study.¹³ This publication serves as a retrospective book selection tool for undergraduate libraries. The author assumed that a well-stocked academic library system will have at least one-third of the titles listed in it, and that large university libraries will contain a majority of the items listed.

First, a pretest was conducted using 20 randomly selected monographic titles from *Books for College Libraries* by collecting class numbers for 20 edition-specific titles. Due to the lack of literature on this topic, there was no way to predict the probability that a title will have a consistent class number across library systems without a pretest. Such an estimate was needed to formulate the hypothesis for testing in order to answer the first research question.

The probability estimate for a title to have the same class number across library systems (L_j , $j = 1, 2, 52$) was calculated as 0.88 using the formula:

$$\hat{p} = \frac{\sum_{j=1}^{52} \text{Total number of matches in } L_j}{\sum_{j=1}^{52} \text{Total number of titles available in } L_j}$$

Using a conservative value of $p_0 = 0.85$, the following hypothesis for testing was constructed in order to answer the first research question: The probability of a title having the same class number across library systems is greater 0.85.

From this chosen probability estimate $p_0 = 0.85$, the minimum sample size of titles "n" was determined using the formula, $n = z^2 p_0 (1 - p_0) / e^2$, where z is the standard normal variable and e the allowable margin of error. Setting the level of significance $\alpha = 0.05$, $e = 0.05$, $z_{0.05} = 1.645$, $p_0 = 0.85$, $1 - p_0 = 0.15$, yields "n" to be 138. Because the first five volumes of *Books for College Libraries* were used (excluding the sixth volume—an index), this figure of 138 was rounded to 200 titles so that an equal number of 40 titles could be selected from each volume.

In order to test the hypothesis, the required sample of titles was randomly chosen from *Books for College Libraries*. Each volume has more than one section. For example, volume 1 (Humanities) consists of five sections—General Works, Philosophy, Religion, Music, and Fine Arts—covering 4, 64, 89, 51, and 107 pages respectively. In this volume, 40 titles were selected by randomly choosing

1, 8, 11, 6, and 14 titles, respectively, from the previously mentioned five sections based on the relative sizes of their listings. The sample of target titles was adjusted to increase comparability and achieve representation. For instance, one title in the representative sample, *Burton's Narrative of a Pilgrimage* (1969), was available in only one library system; it was not possible to make a comparison of class numbers. It was replaced by its earlier 1964 publication, which was held by several library systems. The year of publication for the sample titles ranged from 1911 to 1987. The sample titles did not include any titles published in the 1990s because the most recent edition of *Books for College Libraries* was published in 1988. To overcome this lack of representation, the author decided to find out whether any of the sample titles had editions published in the 1990s. Twenty-one such titles were found, and the later editions were used for comparison. The holdings of these 21 titles of 1990s editions were compared with those of their earlier editions. For 14 of those titles, earlier and later editions were found to be more or less the same. For example, 41 library systems held the 1975 edition of Becker's *Human Capital*, the original sample title, as opposed to 43 that owned its 1993 edition, leading to the replacement of the older edition by 1993 edition. In other words, a replacement was made based solely on whether a library system has at least one copy of the item and nothing else; the quality of assigned class number was not even an issue at this stage. The year of publication for the 200 titles thus adjusted ranged from 1911 to 1995, each year being represented by at least one title. The class number for each title across each library system was recorded using Microsoft Excel. Then for each imprint-specific title, the SCN was identified. SCN matches were recorded as 1, non-matches was recorded as 0, and the results were tallied in a title database; this listing is available from the author upon request.

Analysis—Hypothesis Testing

Using the data in the title database, the following hypothesis was tested:

The probability p of a title having the same class number across library systems is greater than 0.85:

$$H_0: p = 0.85$$

$$H_1: p > 0.85$$

H_0 can be rejected if the T statistic given by:

$$T = \frac{\hat{p} - 0.85}{\sqrt{(0.85)(1-0.85)/N}} > z_{0.05} \quad (A)$$

where \hat{p} is the probability estimate from a representative sample, z is the standard normal variable, the level of

significance \hat{p} is 0.05, and $z_{0.05} = 1.645$ (from the Normal Distribution Table). Probability estimate \hat{p} was calculated using the values from Title-Database as:

$$\frac{\sum_{j=1}^{200} \text{Total number of library systems having a match for } T_i}{\sum_{j=1}^{200} \text{Total number of library systems holding } T_i}$$

= 6404 / 7303 = 0.91. Inputting $N = 7303$ and \hat{p} in Equation A:

$$T = \frac{0.91 - 0.85}{\sqrt{(0.85)(0.15)/7073}} = 14.13.$$

As $T = 14.13$ is greater than $z_{0.05} = 1.645$, H_0 ($p = 0.85$) is rejected at the 0.05 significance level. Consequently, the probability that title has the same class number across library systems is greater than 0.85, implying that more than 85 percent of library systems will have the same class number for the same title, which answers the first research question.

Analysis of Non-Matches in the Sample Title-Database

Of the 200 titles in the title sample, only 79 displayed a perfect match to the SCNs across different library systems (see title database). The remaining 121 titles exhibited non-matches; the extent of the non-matches across library systems ranged from 2.0 to 71.4 percent, depending on the title. A class number differing from the SCN for a title (i.e., a non-match) can be merely a variant, meaning a number different from the norm or an alternate, or a number that is a legal substitute for the norm.

Variation can occur in any part of an SCN. A class number for a title consists of, at most, three parts. The first part is the alphabetical part, which indicates the main class to which the title belongs; the second part (the numerical part following the alphabetical part) indicates the specific topic of the title under that main class; and the third, a Cutter number, which may or may not be added, is used to bring out an aspect, form, period, place, or subtopic of the subject. An alternate or variant class number, one that does not match a SCN, can display variation in any one of these three parts. For each title, the alternates and their respective SCNs were compared to identify where variation occurred.

Those parts of a title that generally contribute to the design of a class number—such as author; subject headings as generated from the LCSH list; the form of the title, such as biography, bibliography—and the appropriate editions of the Library of Congress schedules were examined in order to find out whether any of these could explain the

presence of alternate or variant class numbers for a title. On the assumption that titles are acquired and processed by libraries soon after their publication using the then-available edition of the schedule, non-matching class numbers were analyzed using valid editions of appropriate schedules. For example, non-matches for Borden's *Robert Laird Borden* (1938) were analyzed using the second edition (1913) of F schedule. The third edition was published only in 1958 and was reprinted in 1965. Such an analysis led to identification of four groups of titles:

- Group I: Titles that received alternate class numbers (37 titles).
- Group II: Titles that were assigned variant class numbers due to latitude in the schedules (34 titles).
- Group III: Titles that were assigned variant class numbers for reasons unknown (42 titles).
- Group IV: Titles that were assigned both alternate and variant class numbers (8 titles).

Group I: Titles with Alternate Class Numbers

The 37 titles with alternate class numbers fall into the 8 categories described below.

- Titles with multiple foci: When a title covers more than one subject, it is given multiple subject headings. From these multiple subject headings, a library can choose as a basis for the class number assignment one not favored by a majority of libraries. Sometimes the subject headings for a title may correspond to the same main class but point to different subtopics under that class. Sometimes they may indicate different classes. For example, in the case of Bowen's *Protestants in a Catholic State: Ireland's Privileged Minority* (1983; 56 percent, 4.19), 22 library systems chose HN400.3.A8 (Social history and conditions, Ireland). An equal number of library systems chose the alternative BX4839 (Religion, Protestantism, Ireland). (Parentheses for each title include date of publication, percentage of library systems with non-matches, and serial number in the title database).
- Titles that belong to monographic series: These titles present a dilemma to the classifier: should the class number reflect the series to which the title belongs, or the contents of that specific title? There seems to be no uniform policy among library systems regarding the assignment of series versus analytical class numbers. Consequently, there is a considerable inconsistency in assigned numbers for titles in series. For instance, the title *Environmental Physiology* (1974; 45.5 percent, 5.28) by Robertshaw is volume 7 of series titled Physiology Series One. Its

SCN, QP1 (Physiology, periodicals, societies, serial collections), classed it with the series, while the predominant alternate number QP135 (Physiology, Body temperature) classed it according to its contents as indicated by the subject heading Body temperature-Regulation.

- Literary works of colonial authors: For works of authors from ex-colonies, the PR (English Literature) schedule provides options for two types of class numbers. They can either be classed as "individual authors flourishing after 1960s," or as "authors from colonies, provinces etc." Naipaul's *A House for Mr. Biswas* (1961; 41.7 percent, 2.22) and Birney's *Spreading Time: Remarks on Canadian Writing and Writers* (1980; 16 percent, 2.24) exemplify this problem. In both cases, the SCN classed them with respect to their geographical locations, while the alternates classed them as individual authors flourishing after 1960.
- English fiction titles (cataloged before 1980): Until July 1, 1980, all works of fiction in English were classed in PZ (Fiction in English) at the Library of Congress. The PZ schedule consists of two parts: PZ1-4 (Fiction in English) and PZ5-90 (Juvenile Belles Lettres). Because such an arrangement separated British and American fiction from the rest of the British and American literature as well as English translations of foreign fiction from their originals, the Library of Congress changed its policy to class American fiction with PS, English fiction with PR, and translations into English with the original national literature. Previously, all juvenile nonfiction was also classed in PZ (Juvenile Belles Lettres). Now all juvenile nonfiction is classed with the appropriate subject areas. For example, the SCN of Opie's *Oxford Dictionary of Nursery Rhymes* (1951; 48.6 percent, 2.14) was PZ8.3 O6 (Juvenile Belles Lettres, Verses for children) while the alternate was PN 6110.C4 (Poetry for children).
- Titles belonging to specific literary form, such as poetry: They can be classed with the form schedule or with individual authors. Example: *The Dangerous Edge* (1975; 4.2 percent, 2.23) by Lambert covering history and criticism of detective and mystery stories had PN3448 D4 (Prose, non-historical novels) as its SCN while the alternate was PS3562 A44, a class number for an individual author from the American Literature schedule.
- Memoirs: Memoirs can be classed either with literature or with the subject area of the memoirs. One title illustrated this: Sassoon's *Memoirs of an Infantry Officer* (1930; 42.3 percent, 3.17) with the subject heading World War I—Personal Narratives, English was assigned D640 (First World War, Personal

narratives and other accounts) while alternates PR6037.A65, PR6037.S11, PR6037.A86, PR6037.A84, and PR 6037.A1 were taken from the individual authors section earlier. The alternates also displayed variations in Cutter numbers for reasons discussed under group II titles.

- Bibliographies: Should a bibliography be classed under Z (Bibliography), or under the class number for the bibliography's subject? There seems to be no uniform answer to this question. Thus a bibliography is assigned a Z-based number by some libraries, while others may prefer to classify it under its subject. There were 3 bibliographic titles with 9.8 percent, 14.9 percent, and 13.7 percent non-matches; in all 3 cases the Z-based class number was the SCN. The alternates in each case were the subject-based numbers. Example: Griffin's *Latin America: A Guide to Historical Literature* (1971; 13.7 percent, 5.26) had as its SCN Z1601 (National bibliography, Latin America) and as alternates F1401.5, F1406, F1408, and F1410 (F1401–F1418: History, Latin America) of the English literature schedule. The alternates also displayed variations in Cutter numbers due to reasons discussed under Group II titles.
- Titles that may fall into more than one category with respect to expression: Initially, the subject specialists who designed the LC schedules made use of seven divisions known as Martel's points: (i) General form divisions, (ii) Theory & philosophy, (iii) History, (iv) Treatises, general works, (v) Law, regulation, state relations, (vi) Study and teaching, and (vii) Special subjects and subdivisions of subjects progressing from the more general to the specific.¹⁴ A given title can fall into more than one division, and it is possible for one library to favor one division over the other. Unlike the DDC scheme, LCC does not have a preference order for assigning class numbers in such cases. So, for edited works such as Grzimek's *Animal Life Encyclopedia* (1972; 19.6 percent, 5.30), which exhibited the highest non-match percentage for this category, the majority of libraries preferred QL3 (Zoology, collected works), while others preferred such class numbers as QL7 (Zoology, Encyclopedia) or QL45 (Zoology, general works, 1760–1969).

Group II: Titles That Were Assigned Variant Class Numbers Due to Either Latitude in the Schedule Itself or Lack of Enumerated Numbers

For bibliographies associated with a specific subject (excluding literature), changes may occur in the Cutter parts of their class numbers due to latitude in the schedule. For bibliographies, under the appropriate subject, an .A–Z range

is provided by LCC scheme for the classifier to generate a Cutter number unique for the biography. It is the responsibility of the classifier to generate the Cutter number unless the subject of the biography is well-known, in which case the LC schedule will provide an enumerated number. Variation in generating Cutter numbers produces a corresponding variation in class numbers. One example is Mandel's *Samuel Gompers, A Biography* (1963; 14.6 percent, 4.25), which was assigned HD8073 G6 by a majority of libraries, where HD8073 refers to Labor, Biography and G6 is the Cutter number for Gompers from the LC Cutter Table. All the alternates for this title occurred in the Cutter area, namely G63, G5, G634, and G586.

In the literature schedules, a provision is made for classifying works by and about individual literary authors by means of individual author numbers. Again it is the classifiers' responsibility to extend the appropriate literature number by providing a Cutter number to uniquely identify the author in question. To this end, the classifiers have to use the elaborate Author Tables. For example, Wykes's *Triad of Genius* (1953; 28.6 percent, 2.28) is a book consisting of the collected works of Edith Sitwell. Its SCN was PR6037 I83. The PR6037 part of the number stands for English literature, 20th century, individual authors starting with S. The Cutter number I83 was derived from the Author Table. Variants in the Cutter number were as follows: I898, I88, I9, S62 (probably from the Cutter-Sanborn table), I85, I8, and It757.

The Literature schedules also provide ranges of numbers to accommodate the individual works of prolific authors. Differences can occur in the selection of numbers within these ranges, thus generating variations in the numerical part of a class number. Example: *Selected Letters of Samuel Richardson* (1965; 41.4 percent, 2.34) for which the classifier is provided with numbers in the range PR3660–3667 for Richardson as author. Here a majority of libraries chose PR3664, while variants included PR3666, PR3667, PR 3666 A51, and PO 3666 (possibly a clerical error).

Numbers for classical philosophy exist in both the B and PA schedules. Even though PA class numbers are used only for original Greek and Roman texts, some libraries used them also for translations, which should be classed under B religion. For example, the SCN for Aristotle's *Metaphysics* (1976; 6.5 percent, 1.29) was derived from the Philosophy schedule, while the variants classed it under Greek literature. The same discrepancy was exemplified by a nonclassical work. For Fowles's *Aristos* (1970; 7.5 percent, 1.28), the SCN classed it with Philosophy, while the variants classed it with English literature.

Four titles demonstrated that schedules J and K are capable of producing alternate numbers. La Nauze's *The Making of the Australian Constitution* (1972; 71.4 percent, 4.27), which exhibited the highest percentage of

non-matches (72 percent) among all titles in the sample, fell into this category. Its SCN was JQ4011 (Constitutional history, Australia). More than 20 library systems had alternate class numbers, none matching any other in K. This was probably due to there being no LC law schedules at the time the book was classified, so the libraries under investigation had to use locally designed K schedules. The same was true of Goldfarb's *Contempt of Court* (1963; 34.1 percent, 4.26), which received both JK and KF class numbers. In the case of such titles, the SCNs and the variants were completely different.

Titles dealing with women's issues also exhibited dual numbers. There were two such occurrences: Gilman's *Women and Economics* (1970; 25 percent, 4.31) and *Women and Development: A Resource Guide* (1984; 19 percent, 4.32). In both cases, the first subject heading was Women-Economic Conditions. For Gilman's work, the SCN was HQ1426 (Woman, Feminism, Reform literature, Later 1860-) while the more specific variant HQ1381 (Women and Economics) was given by 4 libraries. The problem with the second title was unique—it had no CIP record, and the LC-assigned class number on the OCLC record was HQ1240.5 (Feminism, Women and state). The *Books for College Libraries* had HQ1154 (Woman, Feminism, Emancipation of Women, 20th century), which was assigned by seven library systems, but the SCN was HQ1870.9 (Woman, Feminism, Underdeveloped areas).

Sometimes the LC schedules provide Cuttering for subtopics. In such cases, variations may occur in the Cutter part of the class number. The HD (Economic history) schedule has listings of special industries and trade. For some industries Cutter numbers are enumerated, while for others they are to be chosen from a range. In the example of Glasner's *Politics, Prices, and Petroleum* (1985; 3.9 percent, 4.30), the SCN was HD9579 G5, where G5 stands for Gasoline. For the non-matches, the variations occurred in the Cutter part of the class numbers—for example, G61 and G41.

Group III: Titles That Were Assigned Variant Class Numbers for Reasons Unknown

Generalizing about the source of variation in class numbers was impossible with 42 titles. Individual libraries may have their own reasons for providing alternate numbers, but those reasons were not obvious from a mere examination of titles. The variations for these titles occurred in the alphabetical, numerical, or Cutter parts of the class numbers. One example is *Solo Song, 1580–1730* (1973; 25 percent, 1.35), which had as its SCN M1619 M17 even though there was no instruction to Cutter by author in the M schedule. It had, however, 75 percent matches across library systems. Non-matching library systems assigned just M1619 or used a Cutter number variation. The CIP

record used the SCN, which may have served as a source for class number.

Cuttering was used by a majority of libraries in three cases to introduce subtopics not enumerated in the schedules. They were Davis's *From the Dark Tower: The Afro-American Writers, 1900–1960* (1974; 4.3 percent, 2.35), with the SCN PS153 N5, where N5 points to Afro-American writers; *Mencius* (1963; 22.9 percent, 1.30) with the SCN BR128 M33, where M33 refers to *Mencius*; and Platnauer's *Latin Elegiac Verse* (1971; 4.8 percent, 2.36), with the SCN PA2335.E5, where E5 refers to elegiac verse. In these 3 cases, the non-matching variants differed only in the Cutter part of the number.

In some cases, the nonconforming libraries seem to have chosen a class number that was broader in scope than the specific enumerated number. For example, there was a specific class number for Leibniz's *Philosophy of Logic and Language* (1972; 2.1 percent, 1.38), yet 1 library chose a broader number. In this miscellaneous group, there were also 2 titles illustrating that it is possible for editions of a title not to get the same class number. In the case of Kochan's *Jews in Soviet Russia since 1917* (1970; 9.1 percent, 3.28), its SCN DS135.R9 is an enumerated number. DS135 in the DS schedule (1959) refers to Jews outside Palestine, and the Cutter R9 refers to Russia, general works, 1917-. The variant was DS135.R92, an enumerated number in a later edition of the DS schedule (1987). Exemplifying the same situation was the title *Muslim Creed* by Wensinck (1938; 18.2 percent, 1.31). Its SCN was BP161 (General works on Islam, 1801–1950), the class number also assigned by the Library of Congress to this particular edition. Two library systems assigned BP166.1 (Islam, History of Theology, General works). A check of the BP schedules in use at the time revealed that BP166.1 was not listed in the appropriate edition (1950). However, during the analysis of editions of this title, it was found that another edition, published in 1968, was given BP166.1 by the Library of Congress, which most libraries also used. The use of BP166.1 for the 1938 edition probably can be attributed to reclassification of the older edition to match the class number of the recent edition by the two library systems in question, while others chose to maintain the old number for the earlier edition and the new one for the current edition. In this case the difference between the variant and the SCN was in the numerical part of the class number. There are times when a number has been discarded in the LC schedules, but some libraries continued to use the old number in favor of the new. Two titles represented this case. An example is Davis's *Systematic Embryology of the Angiosperms* (1966; 26.3 percent, 5.37), for which the SCN was QK 643 A5, while the variant was the discarded number QK 693, which also happened to be the Library of Congress-assigned number.

Sometimes libraries chose a specific number even though there was an explicit “see” reference under that number directing the use of another number. In the case of Frank’s *Historic Pottery of the Pueblos, 1600–1880* (1974; 2.3 percent, 3.21), a library chose NK4017 (Ceramics) even though there was an instruction “for American Indians use E99” underneath that entry. In this case the SCN and the sole non-matching number were completely different. Sometimes, non-matches appear to be wrong. For Walker’s *The Decline of Hell* (1965; 5.1 percent, 1.33), the SCN was BT836.2 (Doctrinal theology, General works, 1951–) but 2 library systems chose BT836 (Doctrinal theology, General works, 1801–1950). So also in the case of Born’s *Principles of Optics* (1980; 9.5 percent, 5.40), the SCN and the variant class numbers were respectively QC355.2 (Optics, general works, 1970–) and QC355 (Optics, General works, 1801–1969).

Group IV: Titles That Received Both Alternate and Variant Class Numbers

Eight titles displayed a combination of sources for variation. Their non-matches ranged from 4.3 percent to 54.3 percent. The following are two examples.

- Borde’s *Sir John Beverly Robinson: Bone and Sinew of the Compact* (1984; 54.3 percent, 4.23), with 54.3 percent non-matches, had two subject headings: Robinson, John Beverly and Judges—Ontario-Biography. Its SCN was KE 406 R63 (Law of Canada, Biography). Its variant was KE406 R62, a different Cutter. Its alternates were F1058 R6 (History, Ontario, Biographies) and KE8248 R66 and KE8248 R6, where KE8248 (Supreme Court of Canada, Judges, Biography) represents the number for second subject heading.
- *Victorian Science and Victorian Values: Literary Perspectives* (1981; 33.3 percent, 5.33). This is a title in the series *Annals of New Academy of Sciences*. Its SCN was Q11 (Science, Societies, United States), corresponding to series. Its alternates were Q175.52 (Science, Philosophy) and PR468.S34 (History of English Literature, special topics), reflecting its two subject headings.

Conclusion

The finding that a title is likely to have the same class number in more than 85 percent of the library systems holding it was surprising. While one expected that the LCC scheme, being an enumerative classification scheme, would yield considerable consistency among assigned class num-

bers, such high degree was nevertheless impressive given the study’s stringent condition imposed for matching class numbers; namely, for a match to occur all the class numbers for a title in a library system had to be the same. The question that then arises is how best to bring the remaining 15 percent of nonconforming library systems into the fold, at least online, in order to make developing a universal search language that can automatically search all catalogs using one class number for a title feasible.

The sample included titles that must have been cataloged in the precooperative cataloging era. The increasing trend of accepting call numbers as is from the copy cataloging record for local use by libraries will enhance the level of consistency in class number for a title across library systems. Moreover, the Library of Congress is replacing ranges with enumerated class numbers in the online version of the scheme as recommended by Williamson, leading to increasing consistency in the Cutter number part of the class number.¹⁵ Thus copy cataloging and replacement of ranges with enumerated numbers will help to reduce the size of variants discussed under group II, III, and IV titles in new records.

Another equally important finding of this study is that a fair number of titles were correctly assigned alternate class numbers. These include those with multiple foci, titles in series, bibliographies, memoirs, titles by colonial authors, English fiction titles, and works that belong to more than one form of expression in particular, discussed under group I titles. An individual library may like to continue assigning alternate class numbers in order to integrate older materials with new ones on the same topic locally. However, such a library will encounter problems if it has a blanket policy of accepting class numbers found in the copy cataloging record and the copy cataloging record has no alternate numbers. For example, the Library of Congress provides both Q11 and PR468 in its catalog record for the 1981 edition of *Victorian Science and Victorian Values* in the sample, but its record for the 1985 edition has only PR468. Eight library systems that had the 1981 edition under Q11 used PR468 for their 1985 volume. The dispersion of even successive editions of the title by class number violates the fundamental principle of classification, namely collocation, and thus is troubling. Ideally, cataloging copy would provide both standard and alternate class numbers for those materials, and online systems would find a way to link those class numbers.

In the card catalog era, when class numbers served primarily as location devices, libraries were forced to choose only 1 class number. Fortunately, online systems have the potential to handle multiple access points, including multiple class numbers. Internet resource databases provide access to thorough multiple class numbers for the same source, as has been recommended in the *Final Report of the Subcommittee on Metadata and Subject Analysis*

of the Subject Analysis Committee in the Association for Library Collections & Technical Services Cataloging and Classification Section.¹⁶ It would be useful to exploit this potential by providing at least all LCC call numbers in the MARC record for a title for the following reasons:

- First, a preferred call number, together with its alternates, more fully represents a title than just the preferred call number alone. Just as a MARC record for a title carries series information and multiple LCSH describing its contents fully, so also it might carry assigned multiple call numbers in order to increase the availability of subject information through call numbers. Such a record can be easily integrated with digital resource records with multiple class numbers as access points.
- Second, thanks to the activation of many types of search keys—such as publication year—in addition to traditional ones—such as author—a user can browse the collections in a library through many types of virtual displays with the stroke of a key. For instance, a multi-foci title can be assigned multiple call numbers, one of which can be designated as a number for fixing the physical location of that title in a library, but all of which can be searched online. Liberating a library's collection from the constraints of linear displays on shelves and in the card catalog would enhance retrieval by providing the user with multiple subject displays. Assignment of multiple class numbers for digital resources is now a reality. Multiple class numbers at the searching phase for nondigital resources will enable better integration of digital and nondigital resources online.
- Third, an exhaustive listing of class numbers for a title—and, by extension, for related titles—would give data to researchers interested in subject analysis. A cluster of call numbers connected by *see also* entries, similar to clusters of subject headings in authority lists of subject headings, leading from one class number to others connected with it in some way, can be a source of valuable information for researchers interested in subject structures.

In short, where deviations from a common number are warranted, libraries could assign two class numbers—the Library of Congress–assigned class or call number(s) for searching and a local class or call number as a shelf collocation device. It would not be difficult for a library, while downloading a record from OCLC, to keep the Library of Congress–assigned class number associated with that record for searching, even if a decision is made to use another number for the shelf listing. The precedence for such a practice already exists. OCLC member libraries

can generate subject headings for local use in generating entries for their online catalogs. Enabling class number searching in sophisticated systems and devising ways to reduce inconsistency will give users enhanced and more precise access to information.

References

1. Robert Freeman and Pauline Atherton, *AUDACIOUS—An Experiment with an Online Interactive Reference Retrieval System Using the Universal Decimal Classification as the Index Language in the Field of Nuclear Science* (Washington, D.C.: American Institute of Physics, 1968).
2. Lois Mai Chan, "Glossary," in *Cataloging and Classification: An Introduction*, Lois Mai Chan (New York: McGraw Hill, 1994), 480–81, 486.
3. Gerard McKiernan, "Beyond Bookmarks: A Review of Frameworks, Features, and Functionalities of Schemes for Organizing the Web," *Internet Reference Services Quarterly* 3, no. 1 (1998): 69–82; Gerald McKiernan, "Beyond Bookmarks: Schemes for Organizing the Web" (Mar. 22, 2001), www.public.iastate.edu/~CYBERSTACKS/CTW.htm (accessed Apr. 20, 2005).
4. Elaine Svenonius, "Use of Classification in Online Retrieval," *Library Resources & Technical Services* 27, no. 1 (Jan. 1983): 76–80; Janet Swan Hill, "Online Classification Number Access: Some Practical Considerations," *Journal of Academic Librarianship* 10, no. 1 (1984): 17–22; Janet Swan Hill, "Things Are Taking a Little Longer than That: A Response to DDC in the Online Environment," *Cataloging & Classification Quarterly* 11, no. 1 (1990): 59–69; Nancy Williamson, "The Role of Classification in Online Systems," *Cataloging & Classification Quarterly* 10, no. 2/3 (1989): 95–103; Lois Mai Chan, "Classification, Present and Future," *Cataloging & Classification Quarterly* 21, no. 2 (1995): 5–17; Lois Mai Chan, "Exploiting LCSH, LCC and DDC to Retrieve Networked Resources: Issues and Challenges," Bicentennial Conference on Bibliographic Control for New Millennium, Nov. 15–17, 2000, <http://lcweb.loc.gov/catdir/bibcontrol/chan-paper.html> (accessed Apr. 20, 2005); Jens-Eric Mai, "The Future of General Classification," *Cataloging & Classification Quarterly* 37, no. 1/2 (2003): 3–12.
5. Richard Hyman, *Shelf Classification Research: Past, Present, Future?* (Urbana-Champaign, Ill.: Univ. of Illinois, Graduate School of Library Science, 1980), 15–16.
6. Arlene G. Taylor, *Cataloging with Copy: A Decision-maker's Handbook* (Englewood, Colo.: Libraries Unlimited, 1988).
7. Ibid.
8. Jennifer Younger and Elizabeth Nichols, *Library Materials Classification Decision Making Survey: A Study for the ALA/RTSD Classification Institutes* (Chicago: ALA/RTSD Classification Institutes, 1986).
9. Edward O'Neill and Patrick McCain, "Copy Cataloging Practices: Use of Call Numbers by Dewey Libraries," Annual Review of OCLC Research (1995), <http://digitalarchive.oclc.org/da/ViewObject.jsp?fileid=0000002651:000000058714&reqid=19460> (accessed Dec. 29, 2005).
10. Ibid.

11. Susan A. Massey and Michael Malinconico, "Cutting Cataloging Costs: Accepting LC Classification Call Numbers from OCLC Cataloging Copy," *Library Resources & Technical Services* 41, no. 1 (1997): 29–38.
12. Association of Research Libraries, "Library Institution Codes 2002: University Libraries," www.arl.org/stats/arlstat/instno_inam.html#univ (accessed Apr. 20, 2005).
13. *Books for College Libraries: A Core Collection of 50,000 Titles* (Chicago: ALA, 1998).
14. Lois Mai Chan, *Imbroth's Guide to Library of Congress Classification* (Littleton, Colo.: Libraries Unlimited, 1980).
15. Nancy Williamson, *The Library of Congress Classification: A Content Analysis of the Schedules in Preparation for Their Conversion into Machine Readable Form* (Washington, D.C.: Library of Congress, Cataloging Distribution Service, 1995).
16. Association for Library Collections & Technical Services, Cataloging and Classification Section, Subject Analysis Committee, *Subcommittee on Metadata and Subject Analysis, Final Report*, ALA Annual Conference 2001, <http://webserve.govst.edu/users/gddcasey/sac/msafinalreport.html> (accessed Apr. 20, 2005).

Appendix: Library Systems Sampled

Arizona State University	Rutgers, the State University of New Jersey	University of Iowa
Boston University	Stanford University	University of Kentucky
Brigham Young University	State University of New York, Stony Brook	University of Maryland
Brown University	Texas A & M University	University of Massachusetts
Columbia University	University of California, Berkeley	University of Michigan
Cornell University	University of California, Davis	University of Nebraska
Dartmouth College	University of California, Irvine	University of North Carolina
Emory University	University of California, Los Angeles	University of Notre Dame
Florida State University	University of California, Riverside	University of Oklahoma
Harvard University	University of California, Santa Barbara	University of Oregon
Indiana University	University of California, San Diego	University of Pennsylvania
Johns Hopkins University	University of Delaware	University of Pittsburgh
Kent State University	University of Georgia	University of Utah
Michigan State University	University of Hawaii	University of Wisconsin–Madison
New York University	University of Houston	Vanderbilt University
North Carolina State University	University of Illinois, Chicago	Washington State University
Pennsylvania State University		Washington University, St. Louis
Rice University		Wayne State University

Index to Advertisers

Library of Congress.....	cover 2
LITA	136
Archival Products.....	148
Library Technologies.....	cover 3

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

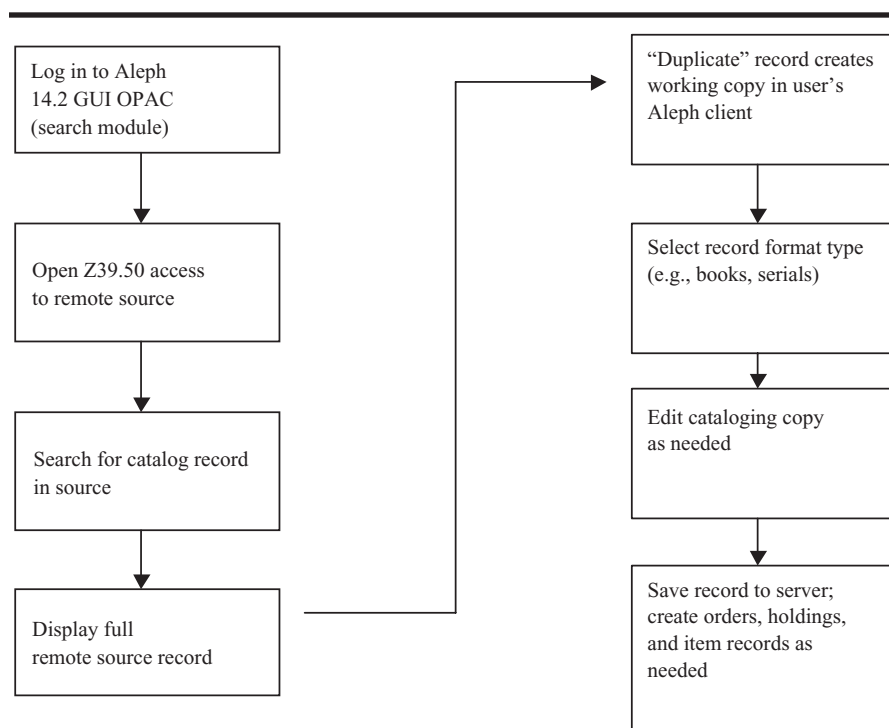


Figure 1. Workflow chart for importing bibliographic records in Aleph500 14.2 through Z39.50

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, UML's 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

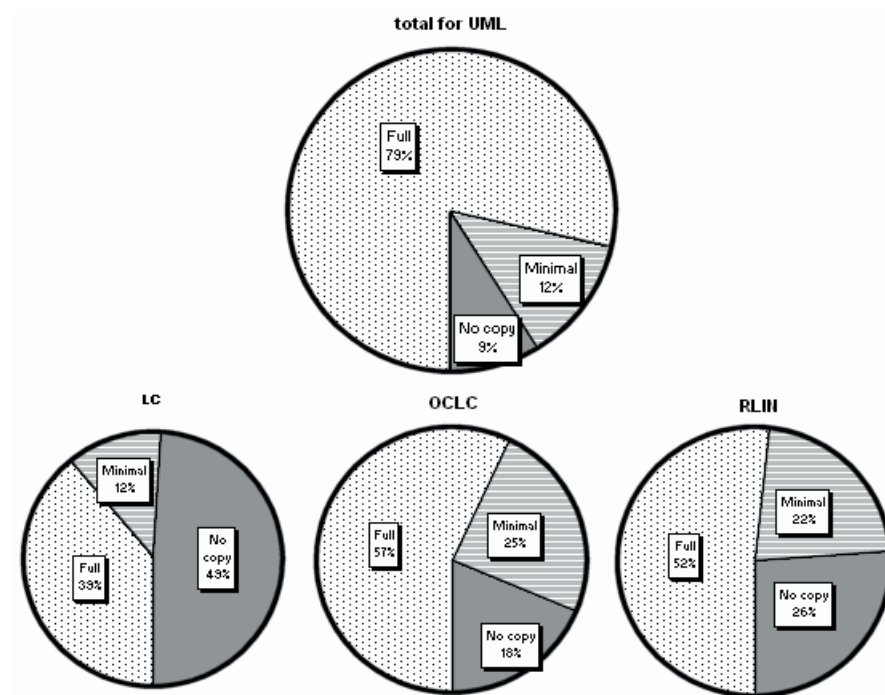


Figure 2. Hit rates by utility and level of copy found

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	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. 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.
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).

Notes on Operations

If You Buy It, Will They Use It?

A Case Study on the Use of Classification Web

Anna M. Ferris

This paper presents a study conducted at the University of Colorado at Boulder (CU-Boulder) to assess the extent to which its catalogers were using Classification Web (Class Web), the subscription-based, online cataloging documentation resource provided by the Library of Congress. In addition, this paper will explore assumptions made by management regarding CU-Boulder catalogers' use of the product, possible reasons for the lower-than-expected use, and recommendations for promoting a more efficient and cost-effective use of Class Web at other institutions similar to CU-Boulder.

Catalogers at the University Libraries of the University of Colorado at Boulder (CU-Boulder) began using Classification Web (Class Web) in June 2002, shortly after it was introduced by the Cataloging Distribution Service (CDS) of the Library of Congress (LC).¹ At that time, Class Web was publicized as the first LC resource to offer cataloging documentation via the Web. Today, it is used by working catalogers throughout the world to formulate classification numbers and subject headings according to the standards and rules published in the Library of Congress Classification (LCC) and the Library of Congress Subject Headings (LCSH).²

At CU-Boulder, catalogers have come to depend on Class Web for various reasons. The most important is the convenience it offers catalogers, who can now access both LCC and LCSH from their personal workstations. Other noteworthy advantages to using the tool include:

- a correlation feature that links both resources;
- automatic calculation of classification numbers;
- a correlation (added July 1, 2004) between Dewey Decimal

Classification numbers and their corresponding LCC numbers and LCSH entries;

- files that are updated by LC on a weekly basis;
- links to other libraries' catalogs (such as LC's Online Public Access Catalog), including the ability to customize a link to the user's own catalog; and
- software features that allow an individual user to limit searches to specific portions of files that are most pertinent to his or her needs.

The first desktop resources for catalogers offered by CDS, Cataloger's Desktop (Desktop) and Classification Plus (Class Plus), were introduced at CU-Boulder in 2001, although they entered the market in 1994 and 1996 respectively. In April 2002, CDS announced that it would be discontinuing Class Plus in favor of Class Web, a new Web interface for accessing the same LC classification and subject headings data. Based on product descriptions at the time, management at CU-Boulder assumed that the 26 catalogers throughout the University Libraries would be able to make a smooth transition from Class Plus to Class Web. Accordingly, the heads of

Anna M. Ferris (anna.ferris@colorado.edu) is Catalog Librarian and Assistant Professor at University Libraries, University of Colorado at Boulder.

The author is indebted to Cheryl C. Cook of the Library of Congress, Janet Swan Hill of CU-Boulder, and J. G. Concepcion for their comments and suggestions.

CU-Boulder's cataloging, acquisition, and systems departments decided to switch the Class Plus subscription to Class Web and to keep CU-Boulder's user agreement at 24 concurrent users at a total cost of \$1,500 annually.

Because Class Web was viewed as simply an enhanced version of Class Plus, and many catalogers were already knowledgeable about how to use Class Plus, department heads did not think that offering specialized training for Class Web was important. Catalogers were simply referred to the tutorial built into Class Web to help them make the transition between the two products. The decisions made during the switchover to Class Web regarding the number of concurrent users, the type of transition, and training were based on assumptions that were thought to be valid for the specific needs of the cataloging department at that particular moment. The validity of these assumptions, a users' survey, and a statistical analysis of Class Web usage are examined in this paper.

Literature Review

The cataloging literature includes several papers that address the importance of providing catalogers with access to cataloging resources in electronic format. Hine investigated the impact that early automation had on increasing productivity and quality of work and stressed the need for further research on the subject of automated workstations for professional catalogers.³ Brisson took up this charge, and traced the historical development and foundations of the cataloger's personal workstation, citing Desktop as "only the beginning" of the types of tools being made available for catalogers.⁴ The impact of automation itself on routine cataloging operations and on working catalogers, in particular, was the subject of papers dealing with the process of change from a perspective of human

reactions (Fiste and Thornton) and behavioral psychology (Cooper).⁵

Other papers refer specifically to the LC cataloging tools. Basic reviews have been written about Desktop (Leazer) and Class Web (Creamer, Selden).⁶ Johnson evaluated online resources available to serials catalogers and briefly touched on the applicability of Desktop for their purposes.⁷ Leroy and Thomas provided a similar review of Web-based resources for catalogers in general, and included short summaries on Desktop and Class Web.⁸ Simpson and Williams went beyond the cursory overviews and provided an in-depth analysis of the impact that Desktop had on cataloging operations at 159 academic and public libraries.⁹

No comparable in-depth research has been produced to date on the subject of Class Web. This paper seeks partly to rectify this situation by providing a detailed assessment of how Class Web is being used by catalogers at CU-Boulder.

Exploring Usage of Class Web at CU-Boulder

Although no formal evaluations were conducted, a year and a half after switching to Class Web, catalogers at CU-Boulder appeared satisfied with the convenience provided by this new, Web-based product. As Class Web had apparently become a valuable resource for CU-Boulder's catalogers, and because they appeared to have adapted to it with little difficulty, the author speculated that the same product could also prove to be a useful tool for noncatalogers. Would it be helpful, for example, for reference personnel to have direct access to LCSH at the reference desk? Before developing a program to introduce Class Web to the more than 20 members of the reference department, the author decided to test her assumptions about CU-Boulder's existing use

of Class Web by gathering data. These statistics would establish a baseline level of use by CU-Boulder catalogers and could then be compared to later usage figures once reference personnel began using Class Web for their own purposes. CDS expressed interest in the study and agreed to supply summary data reports.

Upon request (on a case-by-case basis), CDS is able to produce a summary data report that provides:

- date;
- number of maximum, concurrent users logged on to Class Web for that date at a particular time;
- count of how many records were viewed each time the next or previous buttons were pressed during a twenty-four-hour period per individual cataloger; and
- count of the total number of MARC records displayed per user.

The author requested reports for four consecutive months, from November 2003 through February 2004, which recorded the number of maximum concurrent users logged on to Class Web during that time.

The summary reports, received from CDS on February 23, 2004, showed a lower-than-expected use of Class Web—out of a total of 26 catalogers who had access to Class Web, only 3 or 4 were logged on at a time. Figures 1–4 illustrate Class Web use for each of the four months, showing the maximum number of simultaneous users that were reported for each day. Only days for which there was use are shown. Average use per day was also calculated.

To gain an understanding of the meaning of these baseline statistics, each cataloger was surveyed regarding Class Web use. The survey, appendix A, was sent to each cataloger by e-mail on April 12, 2004. (The

survey also included questions about how catalogers use of Class Web and other resources to assign classification numbers and subject headings; these data are not reported in this paper and the questions are not present in appendix A.)

Survey Results

Of the 26 surveys that were sent, 23 (88 percent) were returned. Table 1 summarizes the findings regarding general Class Web and Class Plus use.

The data reveal two significant points. First, while Class Web was available to all 26 catalogers, and management had presumed that Class Web was being used by them all, 6 copy catalogers (26 percent) did not use the product, and 5 of these copy catalogers (22 percent) had not used Class Plus before the switch. A distinction should be made between the terms *use*, *non-use*, and *non-adoption*. Use refers to the act of consciously choosing to consult Class Web in order to search for information pertaining to LCC or LCSH; non-use refers to the interval during which a cataloger who usually consults Class Web is not using the product; non-adoption refers to the conscious choice by a cataloger not to consult the product. The fact that one copy cataloger used Class Plus but chose not to adopt Class Web is of questionable significance because this cataloger reported extremely low use of Class Plus—less than once a week. (See frequency of use data in table 2.)

Second, the distribution between copy and original catalogers shows that the percentage of Class Web users among the original catalogers was higher (100 percent) than among the copy catalogers (60 percent). This high percentage is almost certainly attributable to the fact that CU-Boulder's original catalogers are active participants in the Program

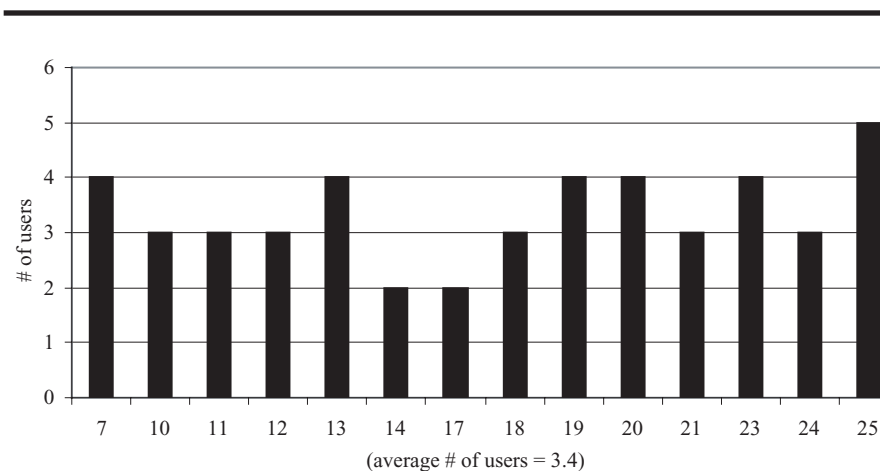


Figure 1. November 2003 Class Web use (14 days)

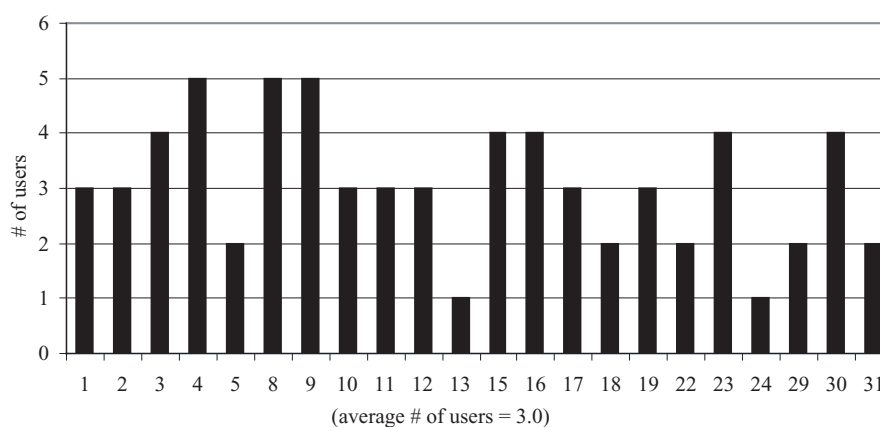


Figure 2. December 2003 Class Web use (22 days)

Table 1. General usage

Respondents	Copy catalogers		Original catalogers		Total	
	%	No.	%	No.	%	No.
Use Class Web	39	9	35	8	74	17
Do not use Class Web	26	6	0	0	26	6
Use Class Plus	43	10	35	8	78	18
Do not use Class Plus	22	5	0	0	22	5
Total respondents	65	15	35	8	100	23

for Cooperative Cataloging and are therefore necessarily adherents of the cataloging standards and guidelines promoted by LC.

The frequency with which catalogers logged on to both Class Web and Class Plus was also explored in the survey. Table 2 provides a

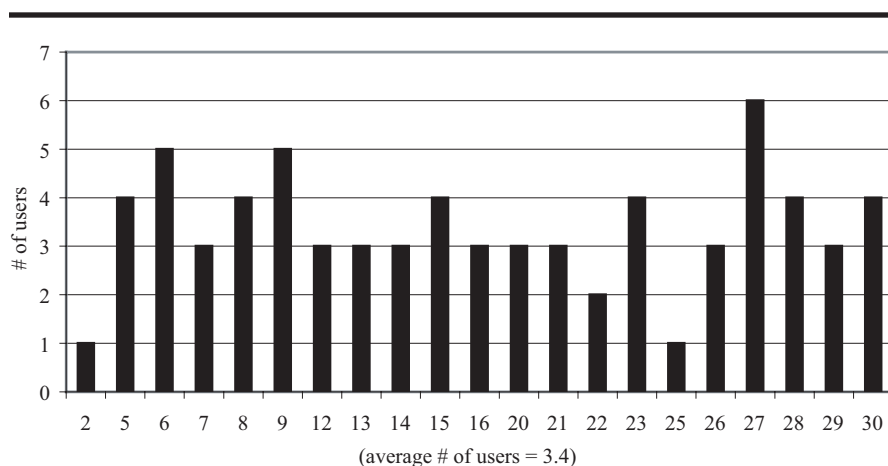


Figure 3. January 2004 Class Web use (21 days)

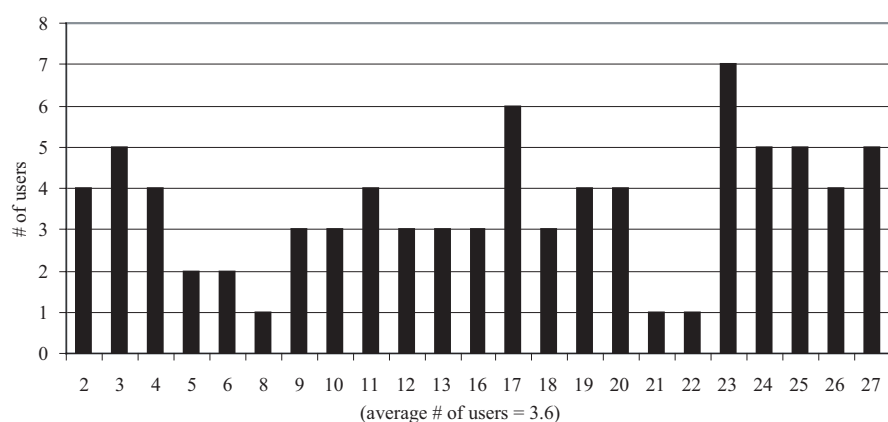


Figure 4. February 2004 Class Web use (23 days)

breakdown of how often users logged on to each product.

Possible Explanations for Low Class Web Usage at CU-Boulder

While there are a number of reasons that could explain the lower-than-expected usage statistics at CU-Boulder, low use or non-adoption do not appear to be the result of any deficiencies or issues of quality connected to the LC product itself. Other factors that might contribute to lower-than-expected usage may include:

- No formal training was conducted for Class Web as had been done for Class Plus.
- Copy catalogers do not need to use Class Web as frequently as original catalogers because call numbers and subject headings are already provided in most LC and member-copy records.
- Other methods of verifying call number and subject heading data are available.
- Catalogers may be experiencing computer access problems.
- Original catalogers may be away from cataloging tasks frequently enough (at confer-

ences, on research leave, and so on) that this may have an impact on the frequency with which they access Class Web.

- The structure of the tool may lead catalogers to log on and off in reaction to a particular query (or set of queries) rather than to leave a session running.
- Some catalogers may be typical *late adopters*; that is, they are comfortable with their current routines and do not immediately appreciate the advantage of adopting a new technology. Other catalogers may be selectively technophobic—reluctant to shift over to an electronic resource when the paper version remains readily available. This seems a likely explanation for the non-adoption of either Class Plus or Class Web.

CU-Boulder's case is likely not unique. In any library, the level of Class Web use may be influenced by some combination of the previously mentioned factors. While many of these factors deserve further investigation, the author chose to concentrate on one in particular—lack of training—as that was a factor that could be corrected. As a result, a formal training program was developed in the hope that it would increase the use of Class Web.

Effects of Formal Training for Class Web

To test whether the low use was attributable to a lack of training, a hands-on introduction to Class Web was presented on April 13, 2004, at a meeting of all catalogers. As a result of preparing for and giving the demonstration, several things became clear. The two products (Class Web and Class Plus) were significantly different. Despite offering the same content as Class Plus, Class Web's browser interface

and other enhancements made it a superior product. Although they were encouraged to do so, many catalogers did not make use of the tutorial or the trial account to familiarize themselves with the updated product. Catalogers appreciated having someone demonstrate the product to them. Logging on and off was not an issue—the product does log users off after thirty minutes of inactivity.

In order to assess whether formal training had had a noticeable impact on use, the author requested two additional summary data reports from CDS for the months of April and November 2004. (Because November 2003 was already available, at the time of this writing, November 2004 would be the first summary report to enable a comparison across years.) The charts corresponding to the April and November 2004 reports are shown in figures 5 and 6.

Data Analysis

Microsoft Excel 2003 was used for the statistical analysis of the summary data reports. The descriptive statistics following are presented in the chronological order in which the reports were requested from CDS: (1) the initial batch of reports from November 2003 through February 2004; (2) April 2004, for the before and after training figures; and (3) November 2004, the first report providing annual totals. Appendix B provides monthly lists of concurrent users taken from these reports.

November 2003 through February 2004

Each data point corresponds to the number of maximum users logged on to Class Web during 14 days in November, 22 days in December, 21 days in January, and 23 days in February (see figures 1–4). Shaded cells are used in appendix B to show weekends and holidays; weekends

Table 2. Frequency of use

Frequency of use	Copy catalogers		Original catalogers	
	Class Web	Class Plus	Class Web	Class Plus
More than once a day	6	5	3	3
Once a day	0	0	1	0
2–3 times a day	1	0	4	5
Less than once a week	2	5	0	0

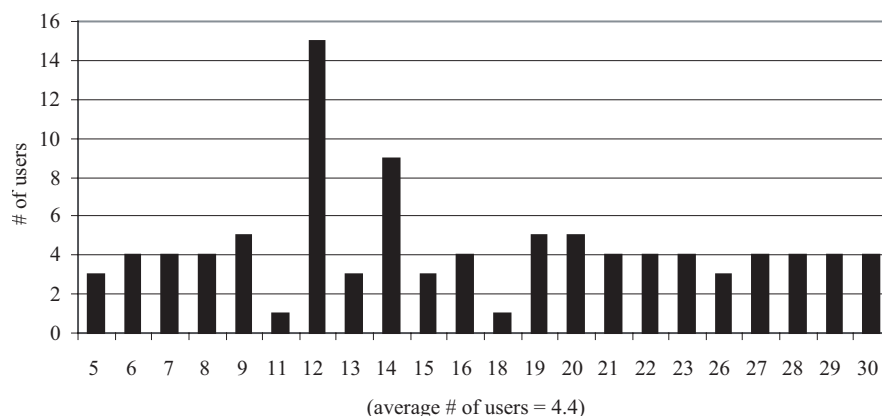


Figure 5. April 2004 Class Web use (22 days)

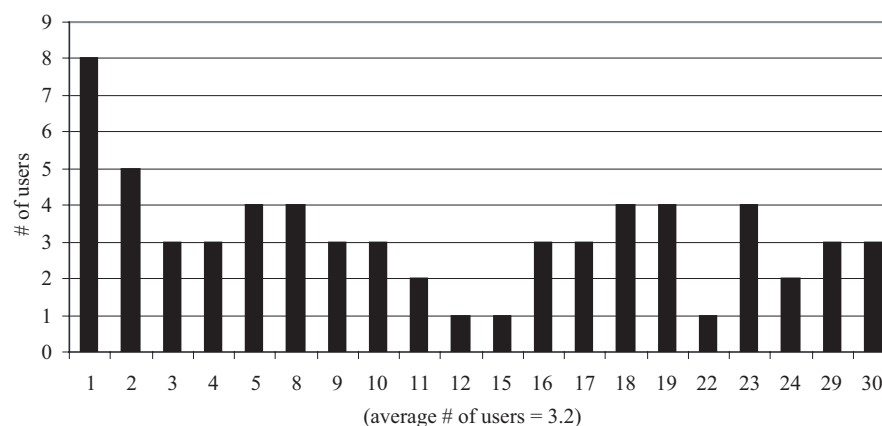


Figure 6. November 2004 Class Web use (20 days)

account for half of the 12 days that only 1 user was logged on to the system at a time. The central tendency in the distribution of data points, and as seen in the four mean totals, is spread between 3 and 4 users per day.

April 2004: Before and after Training

The next set of figures, taken from the summary data report for April 2004 (see figure 5), includes the training date (marked with an asterisk in appendix B). Background use

remained between 3 and 5 users, but significantly higher use was reported on the days immediately preceding and following April 13, the day training occurred. Fifteen users logged on to the system the day before training, and 9 users logged on the day after training. The high use on these two days can be ascribed to the fact that catalogers were taking a closer look at the product. Despite these two days of increased use, the mean total of users remains at a low level (4.409), and the standard deviation is 2.823 from the mean. When the anomalous numbers (15 and 9) are removed from the calculations, the mean total for April goes down to 3.65, and the standard deviation drops to 1.089—closer to the pattern of usage seen in the November 2003 through February 2004 totals.

November 2004

As mentioned earlier, the first set of comparative across-year statistics received from CDS were for the month of November 2004 (see figure 6). These statistics show a decrease in the average number of users (3.2) in 2004 as compared to the average number of users (3.4) during the same month the previous year. Even with six more working days in 2004, the average number of users is still lower than in 2003. Given the unforeseen results of this comparison between the 2003 and 2004 mean totals of Class Web users, a t-test analysis was conducted to test the significance of the difference between the two totals.

T-Test Results

The main reason for performing a t-test (in this case, a paired t-test, as the same subjects were involved in both monthly reports) was to test the likelihood of whether the author's hypothesis, relating low Class Web usage to a lack of training, was a valid one. The supposition was that once catalogers were offered formal train-

ing, use of the product (and consequently the statistics) would increase.

The null hypothesis (difference is equal to 0) presumed little or no change between the levels of use before and after training. The alternative hypothesis (difference is not equal to 0) presumed a significant change in the level of use after training. The determination of a statistically significant difference between the two means is reported as a p-value. Using the mean totals of 3.4 (November 2003) and 3.2 (November 2004), the calculation resulted in a p-value of .709672. Typically, if the p-value is greater than 0.05, the conclusion is that no significant difference exists. Inasmuch as .709672 is greater than 0.05, the null hypothesis must be accepted—no statistically significant difference exists between the mean totals for November 2003 and 2004. In other words, there is insufficient evidence to conclude that formal training helped increase the use of Class Web, given that the pattern of use after training was comparatively unchanged from what the pattern had been prior to the training.

Discussion

Some important issues were uncovered by exploring Class Web use through an in-depth analysis of summary data reports and a usage survey for catalogers. First, Class Web use (an average of 3 to 5 maximum users at any one time) at CU-Boulder was consistently lower than expected. This was perceived as low because access to the resource had been set to accommodate as many as 24 concurrent users. Second, formal training did not have a significant effect on whether or not catalogers chose to use the product.

Two additional issues were highlighted by the usage survey. The first is that managers should not assume that every cataloger will need or choose to use Class Web. At CU-Boulder,

26 percent of the catalogers surveyed (6 out of 23) did not use the product. Additionally, the extent to which catalogers use Class Web may depend upon the job responsibilities held by those catalogers; that is, original catalogers may need to use Class Web more extensively than copy catalogers. Anecdotal evidence suggests that senior catalogers at CU-Boulder made more use of paper documentation in the past than they make use of electronic resources today, but this difference appears to be because their current job responsibilities require them to focus more on managerial duties and less on cataloging or authority work functions than previously had been the case. Many observed that catalogers would use the electronic resources as much as they had used the paper cataloging tools (and perhaps more), if their job responsibilities returned to what they had been in the past.

While the results of this study were not what management at CU-Boulder had expected to find in regards to user numbers and the transition from one product to the other, its significance may well be appreciated by other institutions that are currently subscribed—or are planning to subscribe—to Class Web. The author hopes that addressing these issues will lead to a more efficient and cost-effective use of this product.

Conclusion

The case study presented in this paper has implications for use of Class Web at CU-Boulder as well as at other institutions. At CU-Boulder, the cataloging department will be reassessing its user agreement for 24 concurrent users. A more reasonable access level will need to be determined based on the data collected. A complicating factor will be whether Class Web is made available to noncataloging personnel. If that is the case, reducing access to 3 or 4 concurrent users would not be wise. Formal

training, accompanied by adequate documentation, will continue to be provided for new catalogers. Despite the formal training's lack of impact on the level of Class Web use, survey results showed that catalogers did appreciate the training, and that it made a difference in the effectiveness of catalogers' use of the tool. The benefits of using Class Web will be reinforced from time to time by reporting on special situations or examples in which the product can be especially helpful.

The original impetus for this study was to explore the possibility of extending the utility of Class Web beyond catalogers to noncataloging personnel. A study of current use among catalogers was considered a necessary first step in determining training approaches and estimating possible levels of use by public services personnel. While this study may suggest, based on the moderate use of Class Web made by catalogers, that expectations for adoption outside cataloging should be modest, it revealed no obstacles to providing appropriate training and introducing Class Web to other units within the library. Accordingly, a formal presentation is being developed for personnel outside of technical services, such as reference staff and librarians, bibliographers, and library workers at branch libraries, in order to demonstrate the ways in which Class Web can be beneficial to them. A follow-up study of potential and actual use is also being designed.

Broader implications exist for institutions similar to CU-Boulder. The following recommendations will address some of these issues. First-time subscribers should make sure that appropriate concurrent use is ordered. As this study shows, matching concurrent user access to the number of potential Class Web users at an institution is not necessary. Even at LC, the ratio of Class Web users to concurrent user access is 4:1 (roughly 644 potential users to 161 concurrent users) according to Cheryl C.

Cook, Class Web product manager at the time of this writing.¹⁰ First-time subscribers should consider ordering minimal access at first, and then adding more access as the need arises. At institutions that are already subscribed to Class Web, management should be mindful of how the product is received and used by cataloging personnel—or, perhaps, not used. If possible, use should be monitored. This is easier to do with Web-based products from which statistics are readily extracted. Statistics are not as easily obtained from CD-ROM products unless specific parameters are set by systems personnel beforehand.

Areas meriting further investigation were revealed through the course of this analysis. How do catalogers use Class Web? What motivates them to turn to Class Web rather than to another resource? Despite prompting, why do catalogers tend not to make use of tutorials or trial accounts when they are available? A follow-up survey of nonusers might help uncover the reasons behind non-adoption decisions and low usage in general. Lastly, a detailed analysis of *Cataloger's Desktop* use would be a welcome addition to the literature, particularly as the product has recently undergone a similar shift from CD-ROM to Web format.

This study was prompted by an interest in expanding the use of Class Web to other noncataloging units. Many tools developed exclusively for catalogers have come to be used outside the cataloging department (for example, the LCSH printed editions). Some have been transformed for wider use (the OCLC database, once seen only by catalogers, is today known as WorldCat and searched directly by the public). When cataloging tools were in paper format and expensive, consulting them outside of technical services units presented a problem and limited their usefulness. Now that these cataloging tools are being offered online, access to them is limited only by the availability of a computer, a licensing

agreement, sufficient concurrent user access, or simply the awareness that such tools exist.

Class Web is more than a tool specifically designed for and widely used by catalogers. It represents the innovations that are having an impact on library personnel everywhere. Federated searching, aggregator databases, and other digital initiatives, such as electronic resource management, are only a few such innovations. During this time of technological advancements, cataloging departments should make the most efficient use of such tools as Class Web, not only by putting it to use within their own unit, but also by promoting its use to other noncataloging units, and thereby furthering the collaboration that already exists between technical services personnel and their colleagues in public services.

References

1. Library of Congress, "Cataloging Distribution Service: Bibliographic Products & Services from the Library of Congress," www.loc.gov/cds/classweb (accessed July 5, 2005).
2. *Library of Congress Classification Schedules (A-Z)*, (Washington, D.C.: Library of Congress, Cataloging Distribution Service, 1996–2005); *Library of Congress Subject Headings*, 28th ed. (Washington, D.C.: Library of Congress, Cataloging Distribution Service, 2005).
3. Betsy N. Hine, "Automated Workstations for Professional Catalogers: A Survey of 100 Non-ARL Academic Libraries," *Library Resources & Technical Services* 36, no. 1 (1991): 96–104.
4. Roger Brisson, "The Cataloger's Workstation and the Continuing Transformation of Cataloging: Part II," *Cataloging & Classification Quarterly* 20, no. 2 (1995): 99.
5. David A. Fiste and Christopher P. Thornton, "Bibliographical Services of the Future," *Technical Services Quarterly* 10, no. 3 (1993): 27–43; Eric A. Cooper, "Managing Change to Enhance Technological Orientation and Knowledge among Library Staff,"

- The Electronic Library* 16, no. 4 (Aug. 1998): 247–50.
6. Gregory H. Leazer, "Cataloger's Desktop" [book review], *Library Resources & Technical Services* 41, no. 4 (2004): 347–48; Lesley Creamer, "Review of Classification Web," *Catalogue & Index* no. 149 (Aug. 2003): 7–8; Karen Selden, "Library of Congress Classification Web," *Technical Services Law Librarian* 26, no. 3/4 (Mar./June 2001), www.aallnet.org/sis/tssis/tssl/26-0304/26-0304.htm (accessed July 5, 2005).
 7. Kay G. Johnson, "A Select Survey of AACR2 Tools for Serials Cataloging," *The Serials Librarian* 32, no. 1/2 (1997): 39–52.
 8. Sarah Yoder Leroy and Suzanne Leffard Thomas, "Impact of Web Access on Cataloging," *Cataloging & Classification Quarterly* 38, no. 2 (2004): 7–16.
 9. Betsy Simpson and Priscilla Williams, "The Cataloger's Workstation Revisited: Utilizing Cataloger's Desktop," *Cataloging & Classification Quarterly* 33, no. 1 (2001): 21–32.
 10. Cheryl C. Cook, e-mail message to author, Jan. 31, 2005.

Appendix A. Classification Web Survey for Catalogers

Please check one: ☐ I am a copy cataloger ☐ I am an original cataloger

1. How often do you use Classification Web (Class Web)?

☐ More than once a day ☐ 2–3 times a week ☐ Less than once a week ☐ Once a day
☐ I do not use Classification Web

2. How often did you use Classification Plus (via Cataloger's Desktop) before the transfer to Class Web?

☐ More than once a day ☐ 2–3 times a week ☐ Less than once a week ☐ Once a day
☐ I do not use Classification Plus

Please forward the completed survey to Anna Ferris. Thank you.



2006 LITA PRECONFERENCES Friday, June 23, 2006 New Orleans

Contracting for Content in a Digital World

Join eminent panelists in exploring the rights of content owners vs. users; licensing standardization; pricing; and more. Discuss digital content contracted both by and from libraries; and consider the forces at work in Congress, the courts, and international bodies that influence contracts for digital content.

Moderator: Sybil Boutillier, San Francisco Public Library; *Panel:* Ann Okerson, Yale University Library; Helen Wilbur, Thomson/Gale; David Ferriero, New York Public Library; Alica Wise, Publishers Licensing Society; Miriam Nisbet, ALA-Washington Office

Greenstone Implementation and Customization

Greenstone Digital library software is a free, open source system that enables libraries to create, present, and maintain their digital collections. Learn how to get your digital library up and running and ways to customize Greenstone so that it best suits your needs and vision for your collection.

Speakers: Ian Witten, University of Waikato; Allison Zhang and Don Gourley, Washington Research Library Consortium; Tod Olson, University of Chicago Library; Mark Sullivan, University of Florida Library

LITA accepts on site registrations at the preconferences.

For further information, contact the LITA Office at 1-800-545-2433 x4269

Appendix B. Daily Usage Data from Monthly Summary Reports

Date	Nov. 03	Dec. 03	Jan. 04	Feb. 04	Apr. 04	Nov. 04
1		3				8
2		3	1	4		5
3		4		5		3
4		5		4		3
5		2	4	2	3	4
6			5	2	4	
7	4		3		4	
8		5	4	1	4	4
9		5	5	3	5	3
10	3	3		3		3
11	3	3		4	1	2
12	3	3	3	3	15	1
13	4	1	3	3	3*	
14	2		3		9	
15		4	4		3	1
16		4	3	3	4	3
17	2	3		6		3
18	3	2		3	1	4
19	4	3		4	5	4
20	4		3	4	5	
21	3		3	1	4	
22		2	2	1	4	1
23		3	4	7	4	4
24	4	1		5		2
25	3		1	5		
26	5		3	4	3	
27			6	5	4	
28			4		4	
29		2	3		4	3
30		4	4		4	3
31		2				
Mean	3.357	3.045	3.381	3.565	4.409	3.2
StDv	.841	1.174	1.203	1.561	2.823	1.576

Note: Shaded boxes indicate weekends and holidays; *indicates the training day.

Book Reviews

Edward Swanson

The Complete Guide to Acquisitions Management. By Frances C. Wilkinson and Linda K. Lewis. Westport, Conn.: Libraries Unlimited, 2003. 312p. \$45 softbound (ISBN 1-56308-892-4); \$60 hardbound (ISBN 1-56308-890-9).

Ross Atkinson writes, "The place and image of the library in the institution is mirrored in the position and perception of the acquisitions operation in the academic library."¹ Acquiring and managing access to information resources remains a core function within all types of libraries. However, the library acquisition function has changed during the past decade. The exponential growth of electronic resources combined with budgetary constraints and rising information costs have resulted in many challenges for libraries. It is important that librarians and support staff responsible for library acquisitions have the skills and knowledge necessary to successfully manage the acquisition and delivery of information resources.

The Complete Guide to Acquisitions Management was written to fill a gap in library education. The authors conducted a survey of American Library Association-accredited graduate library programs in 1999 and summarized the outcomes in the introduction. Results of the survey (64 percent response rate) indicated that only 13.9 percent of graduate library programs teach a separate course on acquisitions management, but 93.5 percent do include it as a component in other courses. The majority of respondents (79.3 percent) supported the publication of a new textbook on acquisitions management. The authors' goal was to publish an up-to-date resource to support training of staff new to library acquisitions and "provide the reader with both procedural and philosophical approaches to acquisitions"(xi).

Authors Wilkinson and Lewis are well qualified to write a book on acquisitions management based on their own relevant job experience at the University of New Mexico and numerous publications on library acquisition related topics. The authors effectively synthesize information gathered from numerous resources into a well-organized volume that succinctly covers a broad range of topics. However, this book does not provide the same level of detail and discussion that appeared in the highly regarded second edition of *Understanding the Business of Library Acquisitions*, edited by Karen Schmidt.² What the book does offer is a good introduction to acquisitions management, educating the reader on core acquisition functions. References and links throughout the book lead the reader to resources that will most often provide more in-depth discussion of topics.

The book is organized into twelve chapters that address the following topics: organization of acquisition departments; acquisitions systems; publishing industry; vendor selection and evaluation; acquiring books and media; serials; electronic resources; out-of-print and antiquarian materials; gifts and exchange programs; bindery operations; outsourcing acquisitions; and professional ethics. An appendix follows that includes useful references and links to automated acquisitions systems, relevant conferences, publisher and vendor sites, electronic discussion lists and Web sites, acquisitions-related journals, organizations, and reference tools. A glossary of more than two hundred terms (some with links to additional resources) is included. An index to the book's content completes the volume.

Chapters 4 through 9 offer the most immediate value to staff new to library acquisitions. Chapter 4 provides excellent discussions on vendor types, vendor selection criteria, the basics of the request for proposal (RFP) process, and vendor evaluation criteria; the RFP discussion is especially useful for anyone not experienced with that process. Chapters 5 through 9 address the specifics of acquiring different formats (such as books) and the management of gifts and exchange programs. Of particular interest is chapter 7, "Acquiring Electronic Resources." During the past decade, electronic resources have increased in both number and acceptance by library users. A particular challenge for libraries has been how to effectively integrate management of electronic resources into existing organizational structures. Chapter 7 provides the reader with a good introductory treatment of electronic resource management. New developments in the area of electronic resource management since publication of this book, such as the Digital Library Federation Electronic Resource Management Initiative, will require the reader to supplement their knowledge with more current information. However, if the reader refers to the appendix or glossary, there are references and links to useful resources (for example, the entry under Digital Library Federation in the glossary includes a link to the federation's Web site).

Chapter 11, "Outsourcing Acquisitions," provides a balanced discussion of when to outsource, what to outsource, and the advantages and disadvantages of outsourcing. As the authors point out, "Outsourcing should be a management tool, not a goal" (216). Chapter 12, "Professional Ethics," provides a good introduction to the issues, especially ethics in acquisitions. This is a topic that is often inadequately addressed in similar texts. The appendix and glossary add

value to the book by providing the reader with access to a wide range of useful resources.

The Complete Guide to Acquisitions Management does achieve the authors' goal of providing an up-to-date, comprehensive guide to acquisitions management for staff new to the field. Published in 2003, the content remains timely. As previously noted, the book does not provide the detailed level of discussion that more experienced acquisitions and electronic resource librarians require that appeared in *Understanding the Business of Library Acquisitions*. However, this book does provide a good introduction to acquisitions management and is recommended for graduate library school collections and library technical service or acquisitions department reference collections.—Robert Alan (roal@psulias.psu.edu), Pennsylvania State University, University Park

References

1. Ross Atkinson, "The Acquisitions Librarian As Change Agent in the Transition to the Electronic Library," *Library Resources & Technical Services* 36, no. 1 (Jan. 1992): 7.
2. Karen Schmidt, ed., *Understanding the Business of Library Acquisitions*, 2nd ed. (Chicago: ALA, 1999).

Managing Information Technology: A Handbook for Systems Librarians. By Patricia Ingersoll and John Culshaw. Westport, Conn.: Libraries Unlimited, 2004. 199p. \$45 hardbound (ISBN 0-313-32476-X).

This handbook is a much-needed text regarding systems librarianship and covers topics not encompassed by or in need of updating from earlier works by Thomas C. Wilson and Rachel Singer Gordon.¹ Many systems librarians today are still self-taught, despite the widening of library school curricula to include more information technology basics, and this handbook is a must for the desks of those librarians. Although the Wilson text contains a more thorough (though rapidly becoming dated) treatment of the history of computers in libraries and the philosophy of systems librarianship, Ingersoll and Culshaw provide more than adequate introductory material and address a wide assortment of subject areas relating to information technology. Including many practical tools and lots of tips and advice, their work amounts to a compilation of the many and sundry things they have learned and developed during their collective thirty years of on-the-job learning in an academic library setting.

The first chapter of the work is devoted to planning. With technology playing the major role it does in today's libraries, the importance of the systems librarian's involvement from the outset in any planning process is stressed. The authors appropriately emphasize that successful results are particularly dependent on systems staff involvement in any planning related to integrated library system (ILS)

migration. With the flux in the technology marketplace and the accompanying demise of so many ILS vendors and systems, system migration has become a high priority for many systems departments, and those not affected now can expect to have to deal with the problem somewhere down the road. Ingersoll and Culshaw therefore are right to consider the planning for migration as a subject to be handled in some detail, as they do.

The authors emphasize that reporting lines and staffing should respond in a dynamic way to environmental changes. They note that close alliances between the library and the computer center are likely on today's college and university campuses; if a formal reporting connection does not exist between the library and the computer center, one can generally expect that there is a strong informal relationship, and they stress communication with the intent of shifting "the emphasis from control of information to sharing it" (39). Good advice.

A systems librarian must be an excellent communicator within his or her own organization and also in interacting with peers outside their library as well as with others who may not be so knowledgeable about the ins and outs of technology. Ingersoll and Culshaw stress that the use of technical jargon and acronyms can often inhibit communication with those outside the systems department, in much the same way as the use of library jargon and acronyms can sometimes inhibit communication between librarians and nonlibrarians.

The wide variety of topics covered in the work is both a strength and a weakness. On the plus side, this work will surely open the eyes of some students and even some systems librarians to the ever-expanding role that they are being asked to play. On the other hand, some of the important topics covered, such as statistical reporting, link checking, wireless computing, and other service-related topics, are covered in a summary style, taking up only a few pages. The reader will often wish that these and a number of other topics were covered in more detail.

For a book published in 2004, the dates of the literature cited seem a bit problematic. The authors quote abundantly from literature published in the mid-1990s or even earlier. Given the subject nature of the work and the fluid nature of the systems field, the reader is often left wondering about the relative absence of systems librarianship literature from 2000 and beyond. Although such literature is not plentiful, it is always difficult for the skeptical reader to give credence to the continuing validity of materials more than ten years old in this rapidly changing field. The authors might have found it fruitful and useful to their readers to do a better job of explaining the relevance of the materials cited, even if only as an attempt to provide a historical perspective.

At 132 pages, with an additional 49 pages devoted to resource materials containing a sample local area network

(LAN) account policy, a sample acceptable-use policy, and a library technology plan, the handbook will serve its readers well as a way to get their arms around a wide-ranging subject. Other resource material provided includes ILS vendor information, a job evaluation form, a sample of an online help page, a professional reading list, and a good bit more. These resource materials will no doubt be helpful to systems librarians at many levels and in many situations.

Although library and information science education schools should now be supplying future librarians with systems librarianship specializations, most practicing systems librarians today got into their roles mostly because they are the type of people who feel comfortable with and are undaunted by computers, hardware, and software. Especially for this virtually self-taught type of systems librarian, the Ingersoll and Culshaw text should be a most welcome addition to their collection.—Vicki L. Gregory (*Gregory@shell.cas.usf.edu*), University of South Florida, Tampa

Reference

1. Rachel Singer Gordon, *Accidental Systems Librarian* (Medford, N.J.: Information Today, 2003); Thomas C. Wilson, *The Systems Librarian: Defining Roles, Defining Skills* (Chicago: ALA, 1998).

Binding and Care of Printed Music. Music Library Association Basic Manual Series, no. 2. By Alice Carli. Lanham, Md.: Scarecrow, 2003. 179p. \$52.95 cloth (ISBN 0-8108-4651-9); \$29.95 paper (ISBN 0-8108-4652-7).

This straightforward manual on binding music scores and maintaining music collections is the second volume in the Music Library Association's Basic Manual series. It is aimed principally at "music librarians with little conservation background and to library conservators with little music background" (vii). The author notes that there are many other resources that cover the binding and conservation of printed books, and that this manual pertains specifically to the use of those techniques for binding, conserving, and preserving printed music. What sets it apart is the author's knowledge of how printed music is used, because such knowledge is essential to ensuring that the music will be useable after it is bound. Music librarians and music collection managers are well-aware that printed music bindings must meet study and performance requirements—scores should lie flat on a music stand, and parts should not be jammed into pockets that are too tight or improperly constructed. Thus, the author addresses the particular stresses to which printed music is liable that necessitate bindings that allow for openability, secure page attachment, and storage of loose parts.

There are several options for music binding, as there are for binding books and journals. Carli describes each type of binding and comments upon its suitability for reading

or performance. She then gives clear, detailed, and precise instructions for making each kind of binding. There is sound advice throughout for those who make overarching decisions about repairing and replacing music materials. The instructions for all aspects of the binding process are succinct, free of jargon, and easy to follow. There is an extensive glossary covering terms that are specific to binding, conservation, and preservation.

The verbal instructions are helped immensely by the accompanying simple line drawings, making it fairly easy to carry out the step-by-step instructions for the binding and conservation tasks described. Additionally, even if one does not carry out the instructions oneself, their usefulness lies in the detailed explanations of how bindings are constructed, when one type of binding is to be preferred over another, and what is to be avoided all together. It is a valuable manual for training bindery workers, and would also be a worthwhile textbook in a music librarianship class. People who are working with commercial binders will also find these explanations helpful as they write their binding instructions.

The chapter devoted to working with commercial binders is a compendium of advice on workflow, shipping preparation, and record keeping. The explicit instructions, such as those for preparing a shipment for the bindery and receiving a shipment back into the collection, are invaluable to anyone who is new to the job and does not even know what questions to ask. They are also of value for reevaluating existing procedures. The author suggests that most of the work should be done by students under the supervision of a staff member who is well-versed in bindery procedures. Of course, the amount of preparation done at any library depends upon the level of staffing. One suggestion that the author does not mention, but that does make a real difference in one's approach to working with a binder, is to tour the bindery and talk to the people there about their procedures.

Carli's view that there is a decline in the quality found among commercial binders, due to low pay and repetitive work, seemed a tad harsh, considering the pride that most binders take in the quality of their work and their compliance with the *Library Binding Institute Standard for Library Binding*.¹ However, her point that it is the responsibility of the library to correctly prepare the material and provide clear and complete instructions is well-taken.

Most library managers find it difficult to carry out conservation and preservation plans for their existing collections in the face of the need to acquire and process new materials. Scores published before 1990, other than monumental pieces, tended to be printed on low-quality paper that has become brittle or disintegrated over time. Finding and repairing or replacing brittle or damaged scores is an undertaking that is beyond most libraries. Typically, libraries identify the damaged scores when they are returned to

circulation, which means that only those that are part of the standard repertory or are popular are treated. Carli suggests that a more proactive method of defining a group of materials to be tested for repair or replacement can be the basis for a grant proposal. Her discussion of the deacidification process is enlightening. It is clear that this is a process that is unlikely to be undertaken in any but the largest facilities, but any librarian could use what Carli has written as the basis for an informed grant proposal or a request for proposals from a conservation firm.

Part of the process of determining how a piece will be treated is the development of a preservation policy based on the library's mission, with acknowledgment of local budget restrictions. The policy should be based on the collection manager's knowledge of the needs of the collection and the possibilities of finding replacement materials; the collection manager then brings the specialized knowledge of the various options for treatment. As in the sections on binding new materials, the chapter on repair and replacement is full of definitions and explanations of standard practices. For example, one might not think of microfilming printed music, but this is the most reliable preservation method, and it is a format that can be printed or digitally scanned with equal success.

The appendixes are welcome additions to the main text. The list of supplies and the drawing of tools and a properly laid out workstation are helpful for those not familiar with bookbinding. The sample conservation and preservation policies are useful for people who have little background in this area and are based on the author's expertise in the various techniques and options for preservation and conservation.

This book gives useful advice on the practices and procedures of the care and handling of printed music to managers of printed music collections, whether or not they do their own binding and conservation. Most of it is aimed at institutions that do their own conservation and preservation work, but it also offers sound advice for those that do not. The author's comments on the hazards of book repair include the danger of repetitive stress injuries and the need to have a work area that is designed with adequate ventilation to expel dangerous fumes. She even makes suggestions for varying the work to prevent workers from getting bored with the work and becoming sloppy. The detailed instructions and illustrations in each chapter are instructive for music librarians and collection managers who make binding decisions and have to communicate with commercial binders. The author's commonsense approach and thorough treatment of the topic make this a fine and unique reference. However, the book would be greatly enriched by an index, its lack being the book's only real fault.—*Christina Bellinger* (christina.bellinger@unh.edu), *University of New Hampshire Library, Durham*

Reference

1. Paul A. Parisi and Jan Merrill-Oldham, eds., *Library Binding Institute Standard for Library Binding*, 8th ed. (Rochester, N.Y.: Library Binding Institute, 1986).

Cataloger's Judgment: Music Cataloging Questions and Answers from the Music OCLC Users Group Newsletter. By Jay Weitz; arranged and ed. by Matthew Sheehy; foreword by H. Stephen Wright. Westport, Conn.: Libraries Unlimited, 2004. 265p. \$46.95 cloth (ISBN 1-59158-052-8).

Since May 1989, Jay Weitz, a consulting database specialist at the OCLC Online Computer Library Center (OCLC) who serves as the OCLC liaison to the Music OCLC Users Group (MOUG), and is a former music cataloger himself, has been responsible for the question-and-answer (Q&A) column in the *MOUG Newsletter*. At the instigation of H. Stephen Wright, former MOUG chair, and Martin Dillon, Libraries Unlimited's acquisitions director, the complete text of all questions and all answers, from the first appearance of the Q&A column in May 1989 to the issue number 81, dated September 2002, of the *MOUG Newsletter*, was collected, indexed, and edited by Matthew Sheehy, and published in this monograph.

In the preface to her *Bibliographic Relationships in Music Catalogs*, Sherry Vellucci reports that music catalogers often hear the phrase "If you can catalog music, you can catalog anything!" "uttered with great respect by librarians whose primary responsibility is monographic print cataloging," and it is true that music cataloging is a tricky activity, requiring a vast array of skills.¹ Despite "Weitz's First Law: Don't agonize" (45) it sometimes proves to be of utmost difficulty not to agonize. In that regard, the Q&A column in the *MOUG Newsletter* is certainly an extremely helpful tool. The information provided there is always clear, accurate, and written in a pleasant way—as H. Stephen Wright says in his foreword, Jay Weitz's answers are "not in the stilted, officious voice of AACR2 or Library of Congress rule interpretations, but in a friendly, casual tone that suggested that he was *one of us*" (xii). Some of the catalogers who sent him their questions took that opportunity to confess to him that they cut all Q&A columns out of the *MOUG Newsletter* and keep them on their desk as a collection of particularly precious clippings. Such a testimony tends to show that this publication is more than welcome and meets an actual need.

Having said that, I do have a few reservations. To be fair, I have to admit that both the author and the editor are, themselves, aware of most of them and responded to them in advance in their introduction. The choice of comprehensiveness—almost none of the Q&As, even among the oldest ones, was discarded, "except for a few stray ones that were hopelessly misinformed" (xviii)—results in

"countless references to outmoded technology; superseded documentation and page numbers; old forms of headings; dead URLs; bibliographic and authority records that have long since changed; ancient rules; and obsolete practices" (xviii–xix). The categories under which Q&As are filed do not always seem perfectly satisfactory, but admittedly "there was no intuitively obvious logical organizing principle and . . . no matter what choices [the author and the editor] made, there was no pleasing everyone" (xviii). Some of the Q&As do not really address cataloging issues in the generic sense of the term, but rather are down-to-earth advice about how to use the OCLC system that strives to remain as orthodox as possible while coping with the technical limitations of that specific system (e.g., questions 1.1, 1.10, 3.19, etc.)—a concern quite understandable and legitimate in the context of the *MOUG Newsletter*, but that sounds a bit out of scope in the context of a self-contained monograph. Such Q&As are of little help for catalogers who catalog on a different system; perhaps these should have been dropped for the monograph publication.

Wright reports in his foreword that MOUG considered for a while "placing all the columns on a Web site" (xiii). This would certainly have been a better mode of publication than the printed format, as it would have enabled the constant updating of the valuable information provided in these columns. One can hope that this book will soon be transferred to the Web, its appropriate place.

Updating is a real issue for publications of this kind. Will most of the Q&As be still relevant, once the new cataloging code, *Resource Description and Access* (RDA), is in place? If Metadata Object Description Schema (MODS) or some other XML-based format happens to supersede MARC21 in the future, three-quarters of this book will become obsolete at once, for the real topic addressed in this book is not so much cataloging as it is the MARC21 format itself—and its inadequacy. This book can also be read as a kind of compendium of everything that is wrong with MARC21 and could have been titled *How to Do One's Best with Inadequate Tools*. For example, many pages (128–145) are dedicated to field 246, Jay Weitz's personal nightmare. The very existence of that field does not seem more justified after reading all those pages than before. The questions relating to field 041 (pages 150–56), dedicated to language codes, show that this field was primarily designed for printed textual materials and does not suit quite well for sound recordings. Weitz admits, "Regarding the 041 subfield \$b question, I have taken MARC21 at its word, though I'm not sure my interpretation is correct" (151).

Similarly, even more than the "cataloger's judgment," what this book tends to highlight is the inappropriateness of many current cataloging rules. Weitz has too much genuine respect for the rules to make the point, but many of the problems he had to solve in his Q&A column would

vanish into thin air if the rules were more consistent. Of course, as Weitz puts it, "real-world instances, in spite of our never-ending efforts to codify practices, will always defy those efforts" (xix), but some ambiguities could be avoided. For example, there is a basic ambiguity about what it is we catalog at all when we catalog a sound recording: is it just the content infixed on the carrier (the recorded sound), the carrier itself (e.g., the CD), or the whole package that the publisher intends as a *product* (the CD plus its container, liner notes, so-called accompanying material in our traditional but erroneous terminology, and so on)? This basic ambiguity results in a difficulty that many sound recording catalogers encounter: How to assign a correct publication date to a sound recording (pages 23–26 and elsewhere). There is a confusion between dates actually pertaining to the publication described, dates pertaining to bibliographic antecedents, and dates pertaining to the history of the recorded sound itself. Hopefully, RDA will solve such problems.—Patrick Le Boeuf (patrick.le-boeuf@bnf.fr), *Bibliothèque nationale de France, Paris*

Reference

1. Sherry L. Vellucci, *Bibliographic Relationships in Music Catalogs* (Lanham, Md.: Scarecrow, 1997), xvii.

Copyright in Cyberspace 2: Questions and Answers for Librarians. By Gretchen McCord Hoffman. New York: Neal-Schuman, 2005. \$75 paper (ISBN 1-55570-517-0).

Copyright for Teachers and Librarians. By Rebecca P. Butler. New York: Neal-Schuman, 2004. \$59.95 paper (ISBN 1-55570-500-6).

Books appearing more or less simultaneously on essentially the same subject often evoke from contemporary reviewers some reference to "The Blind Men and the Elephant" poem based upon the Indian legend about a number of sightless men who, upon encountering it for the first time, touch and feel different parts of the great beast and draw markedly different conclusions as to its nature.¹ Both of these two books, by contrast, draw much the same conclusions, although they arrive at them by way of somewhat differing approaches. Both are comprehensive works, covering much of the wide range of subjects pertaining to copyright law that are implicit in library operations, whether from the technical services or readers' services sides.

Hoffman, described in the foreword to her book as that rarity—a librarian who went to law school rather than the perhaps more typical lawyer who went to library school—provides the more legalistically oriented of the two works. *Copyright in Cyberspace 2* contains numerous citations to and lengthy excerpts from the applicable statutory law and court cases, and is an updated second edition of Hoffman's 2001 work, which does a good job of bringing up to the present a number of the concepts and problems discussed in the

first edition. Part I improves and brings up to date the excellent short history of the copyright laws and basic applicable concepts contained in the first edition, and is especially useful in regard to the key fair use principle.

Parts II and III undertake to apply the principles of copyright law to specific library situations and emphasize the potential copyright pitfalls librarians now typically face on almost a day-to-day basis. Effectively emphasizing that there do not always exist *bright-line* answers to every question, Hoffman rightly urges librarians to become more involved in developing the public policy issues at work in the copyright arena, especially when it comes to matters involving revising laws to reflect the digital and electronic realities presented in libraries today.

Hoffman utilizes a sidebar Q&A supplementation to concurrent textual discussions in each chapter and to the general Q&A format used in each chapter (a sort of question-within-questions approach) in order to highlight important issues, an approach that is likely to be effective for the casual user treating the work more as a reference tome than a treatise on its subject.

To work best as a reference book, however, a Q&A format demands a comprehensive index, and it is here that the work falls down a bit, as did the first edition. The table of contents might have been of some significant help in this regard if it had included all the questions for which answers are provided in the text (as the first edition did), but that degree of comprehensiveness has been eliminated this time out. A major portion of the work remains devoted, as was the case with the first edition, to lengthy exposition of primary legal sources that, given today's ready Web access to such public domain sources, makes for a longer book than seems justified.

Overall, though, *Copyright in Cyberspace 2* continues the first edition's valuable contribution to the literature in the field, and combines a librarian's insight with a lawyer's knowledge to provide a useful addition to the shelves of librarians for whom the work is intended. Hoffman also commendably eschews the lawyer's tendency to overexplain or pontificate, and her accessible writing style results in a readable book on a difficult and technical subject.

Rebecca Butler is not a lawyer by trade, but *Copyright for Teachers and Librarians* does not suffer from either a lack of authoritativeness or an overly academic approach. Indeed, it is obvious that it was written to be read rather than referred to. Butler has applied and adapted her work in developing copyright workshops, classes, and presentations to give the profession a useful primer on copyright that is accessible and likely to be useful on a daily basis. Working from essential concepts in part I, she brings the reader smoothly through clear discussion of fair use and licensing concepts to specific applications in various situations in part II. The table of contents and list of figures and charts

are very good, and they resolve many of the problems that (again) a somewhat foreshortened index might have otherwise suggested.

When dealing with specific situations and questions, in most respects Butler arrives at the same ultimate conclusions as Hoffman, but when, as is frequently the case, there is no clear answer, Butler tends to opt for the conservative view more often than Hoffman, usually recommending against copying, using, and so on, in the close cases. She provides her recommendations primarily through a series of flowcharts appearing frequently in each chapter of part II, with yes/no decision trees down the paths of which librarians may navigate. She supplements these decision tools with relevant contextual discussion, making the flowcharts much more useful to the discerning reader.

Whether the copyright law, or any other law within the Anglo-American legal tradition, based as it is on court decisions and interpretations as opposed to statutory prescriptive absolutes, necessarily lends itself always to yes or no analysis is beside the point so long as the author's frequent caveats are observed. And, as previously mentioned, because Butler almost ineluctably leans to the conservative approach, librarians following their decision trees will rarely find themselves in a compromised or embarrassing position regarding use of the library's copyrighted materials.

Copyright for Teachers and Librarians would perhaps benefit from some explication of possible and appropriate solutions to those copyright law dilemmas that lead librarians to take (and Butler often to recommend) a conservative approach. Certainly the library profession needs to be in the forefront of providing access to materials, thoughts, and ideas, not serving as surrogate copyright policemen intent on ensuring that materials are squirreled away for the use only of those who can pay each time for their use. Some sort of afterthoughts chapter, which would not disturb the otherwise helpful natural flow of the book, might have been useful in this regard. One would expect Butler's views in this regard to be incisive.

So, is one view of the copyright elephant any more accurate than the other as far as the librarian is concerned? Happily, these books complement each other in many ways. For those looking for a straightforward, handbook type of approach to use in applying the mysterious ways of copyright law to specific, everyday library situations, *Copyright for Teachers and Librarians* might be the better choice. But for more detail in respect to the ins and outs, and the benefits that can be derived from "further review" through examination in some detail of the statutory provisions and court decisions regarding copyright law, there is much to recommend in *Copyright in Cyberspace 2*. The reader of both will find the elephant most clearly in focus.—Vicki L. Gregory (Gregory@shell.cas.usf.edu), University of South Florida, Tampa

Reference

1. John Godfrey Saxe, "The Blind Men and the Elephant," www.wordfocus.com/word-act-blindmen.html (accessed Jan. 16, 2006).

Putting XML to Work in the Library: Tools for Improving Access and Management. By Dick R. Miller and Kevin S. Clarke. Chicago: ALA, 2004. 205p. \$45 (\$40.50 ALA members) paper (ISBN 0-8389-0863-2).

Extensible Mark-up Language (XML) has been at the heart of many discussions for ten years and was sometimes introduced as a kind of miraculous panacea in outbursts of enthusiasm verging on the irrational. XML was not primarily designed for librarians, but quite early some librarians—especially those who had been considering replacing MARC formats with SGML-based formats—saw the potential it has for the profession. Some commentators, however, also expressed reservations about XML's ability to deal with huge amounts of bibliographic data. The authors are both aware of XML's limitations and convinced that XML can do a lot for us: "While XML cannot solve all of our problems, it does offer foundational tools to help transform the way libraries do business" (36). Indeed, it is almost a matter of survival, as the new environment—the Web environment—has profoundly transformed libraries' role and place within the society: "'Library information,' especially that in time-honored MARC formats and in proprietary integrated library system formats, has been segregated too long from mainstream Web resources. Having an online library catalog isn't good enough anymore" (37) and "Conditions for libraries have changed! . . . With instant information everywhere, libraries need to reassess their role and focus on strategies for thriving under the new circumstances" (96). Obsolescence and growing isolation are the major threats impending on libraries and librarians. The authors regard XML as a way to escape both.

This book (written in 2002, published in 2004, and reviewed in 2005 for this 2006 issue of *LRTS*—it is important to keep that time aspect in mind) can be regarded as comprising two distinct sections. The first one (chapters 1 and 2) is a presentation of XML itself and XML-related technologies (validation tools, linking tools, display tools, and so on), a kind of XML manual for librarians who are not acquainted with the mark-up language. The second one (chapters 3 to 5) is more library-specific and exposes how to develop—and put into practice—an XML-based metadata schema, with the potential to solve the many flaws that cataloging rules and MARC formats are fraught with.

Librarians who seek guidance for the development of their first XML Document Type Definition (DTD) or XML schema will find a step-by-step methodology that will prove extremely helpful on pages 94–96. Perhaps more importantly, this section offers the authors an opportunity to express what they think of MARC21 and the Anglo-American

Cataloguing Rules—and they do not think good things, to say the least. They develop seven arguments in favor of XML's superiority over MARC: coded values in fixed-length fields could be replaced with the flexible authority control enabled by XML; the inconsistent way dates are expressed in MARC could be unified in XML; redundancy and inconsistency in entering similar types of information could be avoided; MARC does not clearly separate information elements and information about them, which XML would make possible; relationships could be expressed in a unified way; MARC's complexity could be replaced with a core XML schema to which specialized information elements could be added for certain types of materials (e.g., music, maps); and the MARC-8 character encoding system could be replaced with Unicode. According to the authors, XML could also help solve some of the problems posed by AACR: no clear identification of works; no consistent treatment of relationships; too much emphasis on transcription and description; use of mixed-language headings, which impedes internationalization; inconsistent treatment of initial articles in titles, and so on.

The authors then introduce the XML-based meta-data structure that they have developed at the Lane Medical Library, the XML Organic Bibliographic Information Schema (XOBIS)—in my opinion the biggest revolution in the cataloging world since Cutter's time. XOBIS blurs the traditional—and cumbersome—border between bibliographic and authority records. It enables consistent treatment of bibliographic relationships and controlled use of qualifiers within headings. The authors insist, however, that XOBIS is "experimental," and that "it should not be interpreted as minimizing the problems such an undertaking [i.e., the replacement of MARC with a Web-oriented schema] would entail" (144). It would be fascinating to see what a large-scale bibliographic database in XOBIS might look like.

Unfortunately Chapter 4, which is devoted to the software and practical tools that would enable daring librarians to "put XML to work" in their library, is a bit disappointing. Not that it is not helpful, but it focuses almost exclusively on open source software that perform the following functions: edit XML documents; transform XML into other formats; display; store and index; or any combination of the above. Of course, such a publication could not and should not turn to a collection of advertisements for commercial systems and the vendors who supply them, but it is not always possible to find the qualified staff able to adapt open source software for a library's specific needs, and one has to be completely informed to make the good decision and the good choice.

The fifth and final chapter, devoted to XML's potential for the future, shows a number of the Lane Medical Library's achievements that were made possible thanks to XML. Among other realizations, they maintain an online serials list from their catalog, they use XML "to assist with

the retrieval of information that sometimes gets buried in, or omitted from, the MARC format" (183), and they locate journal articles and so on.

One of the many lessons that this highly valuable book has to offer is that lazy solutions consisting in just transferring MARC fields into XML tags are not the best ones and do not put librarians in a position to envision the future with serenity. "Rather than just attempting to translate existing knowledge structures directly into XML, we have a strategic opportunity to redefine these structures in order to support future information systems. . . . We advocate the need for fundamental changes in order to achieve a viable replacement schema" (92, 93). Is the profession ready to follow pioneers and to invest time, thinking, and money in the XML revolution?—Patrick Le Boeuf (*patrick.le-boeuf@bnf.fr*), *Bibliothèque nationale de France, Paris*

Who Wants Yesterday's Papers?: Essays on the Research Value of Printed Materials in the Digital Age. Yvonne Carignan et al., eds. Lanham, Md.: Scarecrow, 2005. 224p. \$48 paper (ISBN 0-8108-5119-9).

Have you ever noticed that when an issue comes to one's attention, permutations of it seem to arise at every turn? Recently, while listening during my daily commute to a CD educational series on classical music, I heard the instructor bemoan the belief of some scholars that fully 40 percent of the musical works of Johann Sebastian Bach—that would be in the range of 400 works—has been lost, and that the paper they were written on was likely used to wrap cheese or to provide insulation for the roots of plants and trees.¹ During the same period, in preparation for a family move, I sorted through the entire kindergarten through twelfth-grade academic output of my two offspring and decided what among the hundreds of pages of math worksheets and other busy-work could be kept and what discarded. As I composed this review, I reflected that loss and preservation are hardly new or unusual issues, but ones that have long pervaded both the public and private spheres.

In 2001 Nicholson Baker, through the publication of his provocative *Double Fold: Libraries and the Assault on Paper*, created a flurry of attention on libraries' decisions to replace fragile or deteriorating collections of newspapers with microfilm.² Although *Double Fold* and its author have been widely criticized for arguably oversimplifying or obscuring the issues, they have served a valuable purpose by fostering a great deal of professional and public discussion of the issues surrounding the preservation of original materials in an increasingly digital age. Among the responses to Baker was a symposium organized by University of Maryland librarians and graduate students from the university's college of information studies.

While *Double Fold* focused on newspapers (and library card catalogs), the symposium's organizers had a far more

ambitious objective, to address "the whole question of what original materials should be saved more broadly" (ix). The organizers' strategy, culminating in the publication of *Who Wants Yesterday's Papers?*³ was to promote a dialog between researchers from a range of academic disciplines—including humanities and the social and physical sciences—and the librarians and preservationists who face the daunting task of short-term decision-making and long-term planning in these areas. The book contains both the symposium presentations and further essays added to provide a more complete overview of various aspects of the issue. The result provides a useful introduction to the complexities of the topic of the greatly differing research needs of varying disciplines. However, it also accomplishes much more by providing both a historical perspective on library preservation (primarily in part one, "The Race against Time") and an introduction to the technical and other problems associated with digital preservation. Many of the essays are extensively footnoted, and the book also contains an annotated bibliography of sources for further research. Although there are current issues not explicitly dealt with due to the fact that the symposium took place in 2002, such as developments in the area of government documents and depository libraries and the Google digital initiatives, *Who Wants Yesterday's Papers?* provides an excellent springboard for further exploration of the topic.

Part two, titled "Digital Demand vs. Paper Pleas," explores the importance of original paper documents from the perspectives of University of Maryland professors in various disciplines. These scholars were asked what types of materials they used in their research; how important paper materials were, as opposed to microfilmed or digital materials; and whether and how their reliance on original materials had recently changed or would change in the foreseeable future. Science historian Stephen G. Brush's argument that old science textbooks, far from being outdated and useless, are crucially important takes on particular forces, considering the ongoing debate over the teaching of evolution and intelligent design; it's difficult to imagine a better illustration of how textbooks as "social artifacts" (40) provide insight into the values and convictions of their respective eras. By contrast, physicist Jordan Goodman and archivist Kara M. McClurken explore the advent of online scholarly publishing, extolling its potential for rapid publication of scientific research results and enhanced methods of peer review, while touching upon copyright and preservation issues. They also address the problems reliance upon digital information pose for the researcher, describing how historian Michael Bellesiles was stripped of his academic position and awards after being unable to reproduce some of the research data used in writing his book *Arming America: The Origins of a National Gun Culture*.³ John E. Newhagen's essay "Above the Fold: The Value of Paper Newspapers" presents several cogent examples of how "the preservation of the physical

form of yesterday's papers will be critical to preserving the meaning and content of the messages they embody" (87). A recent example is Laurel Leff's *Buried by The Times: The Holocaust and America's Most Important Newspaper*, in which Leff, a former newspaper journalist and current professor of journalism, exhaustively studied the *New York Times's* coverage of the campaign against European Jews from 1939 to 1945, analyzing how decisions about the content and placement of news article on the genocide had the effect of minimizing public awareness and outrage.⁴

Part three, "Enduring Value," features the perspectives of a variety of preservation professionals and scholars seeking to explore the issues of how we should select materials for preservation and how to preserve them (and perhaps re-preserve them through reformatting) in the appropriate format for their audience, while taking into account the resources available for such preservation efforts.

Richard J. Cox, of the School of Information Science at the University of Pittsburgh, who contributed the afterword to *Who Wants Yesterday's Papers?*, suggests that the answer to the question posed by the book's title is not, as the Rolling Stones concluded in their 1967 recording, "Nobody in the world," but rather that everyone wants them, and they want all of them.⁵ According to Cox, the major error of such critics as Baker is the failure to recognize that selection is an imperative, because not everything can be saved. Cox opines, "archivists and librarians have not sufficiently explained themselves or provided adequate reasons for some of their preservation activities. . . . we cannot take for granted that we or our discipline will be understood or appreciated by external observers."⁶ As a catalyst for discussion of these important issues, *Who Wants Yesterday's Papers?* poses many more questions than it answers and provides avenues for further exploration and debate on the issues of the way society uses and values information. —Susan Herrick (sherrick@law.umaryland.edu), University of Maryland School of Law, Baltimore

References

1. "J. S. Bach: Concerto in G Major for Violin, Viola and String Orchestra," www.rbpmusic.com/bach.htm. (accessed Jan. 5, 2006).
2. Nicholson Baker, *Double Fold: Libraries and the Assault on Paper* (New York: Random House, 2001).
3. Michael A. Bellesiles, *Arming America: The Origins of a National Gun Culture* (New York: Knopf, 2000).
4. Laurel Leff, *Buried by The Times: The Holocaust and America's Most Important Newspaper*. (New York: Cambridge Univ. Pr., 2005).
5. "Yesterday's Papers Lyrics—The Rolling Stones Lyrics," www.rolling-stones-lyrics.com/papers.html (accessed Jan. 6, 2006).
6. Richard J. Cox, *Vandals in the Stacks? A Response to Nicholson Baker's Assault on Libraries* (Westport, Conn.: Greenwood Press, 2002), 194, 196.

Brief Reviews

The Next Library Leadership: Attributes of Academic and Public Library Directors. By Peter Hernon, Ronald R. Powell, and Arthur P. Young. Westport, Conn.: Libraries Unlimited, 2003. 192p. \$50 paper (ISBN 1-56308-992-0).

Acknowledging the shortage of librarians entering the management ranks of academic and large public libraries, this book addresses a topic of great interest to the profession and fills a need in the library literature. Beginning with a brief chapter supporting the claim that there is a shortage of librarians, the book continues with a literature review that provides a good introduction to the topic of leadership in libraries and makes some distinctions between the attributes needed by managers versus those needed by leaders.

Three subsequent chapters list and analyze the qualities needed by Association of Research Libraries (ARL), Association of College and Research Libraries (ACRL), and public library directors. Each of these chapters treats the topic as a separate research study that begins with lists of potential attributes, a survey of leaders in the profession, and follow-up interviews of selected individuals. Whereas the lists of desired attributes arrived at through the surveys were agreed upon by those surveyed, there are clear differences of opinion when individuals are queried. The next few chapters compare the lists of attributes identified by ARL, ACRL, and public library directors, address tools that can be used to assess the leadership qualities and abilities of individuals, and review methods for attaining leadership skills, such as work experience, leadership institutes, and mentoring, although the authors do not endorse any particular method. Finally, a brief chapter concludes with a discussion on the use of head-hunters to identify and recruit potential leaders.

Overall this book will be useful for librarians who would like to pursue leadership opportunities, and also for libraries that are recruiting for leadership positions.—Rebecca L. Mugridge (rlm31@psu.edu), Pennsylvania State University, University Park

An Ounce of Prevention: Integrated Disaster Planning for Archives, Libraries, and Record Centres, 2nd ed. By Johanna Wellheiser and Jude Scott, with the assistance of John Barton. Lanham, Md.: Scarecrow; Toronto: Canadian Archives Foundation, 2002. 283p. \$33 paper (ISBN 0-8108-4176-2).

Wellheiser and Scott have thoroughly revamped and expanded Wellheiser and Barton's 1985 first edition, an indispensable manual of its day. Retaining their Canadian-centric focus but encompassing a worldwide scope, the authors incorporate up-to-date approaches to disaster prevention, protection, preparedness, response, recovery, rehabilitation (for collections, records, facilities, and systems), and post-disaster planning for water-damaged collections and records, including CDs and computer media as well as paper-based materials.

In the wake of damage from Hurricanes Katrina and Rita in 2005 on the United States Gulf Coast, disaster recovery rises to the forefront of many librarians' and archivists' concerns. The problem is that we are still talking, usually after the fact, about how to recover from a disaster, not how to protect collections and buildings before a disaster strikes, nor are we preparing to mitigate disaster. It takes money, time, and unrelenting commitment and attention to create, adopt, absorb, and practice disaster prevention and recovery. Once the adrenaline rush of the disaster is behind us, we complacently go back to life as usual until the next calamity. Most manuals deal with the best-case scenario of treating the damage within forty-eight hours to avoid mold, and *An Ounce of Prevention* espouses that approach. When the roof is gone, the staff has evacuated to another state, and whatever security forces in charge do not allow anyone to enter the area for weeks or months, not just hours, what hope is there for recovery of mold-damaged collections? The reality of catastrophic occurrences such as Katrina and Rita raises the question of the value of following instructions in a disaster recovery manual designed for a burst water pipe or leaky roof. Whole mindsets have to change so they do not fall back into the it-cannot-happen-to-me syndrome.

An Ounce of Prevention lays out the groundwork in logical progression for librarians and archivists to start today to plan for all the stages of a disaster. Not only is it essential reading, but an absolute necessity to consult as a practical handbook while writing an institution's disaster plan. The detailed chapters go far beyond most published compilations of conference papers, which often focus on "how I did it in my shop." Wellheiser and Scott include two eye-opening tables dealing with the costs—per book and per box—of recovering and rehabilitating fire- and water-damaged books and archival documents. This book joins its predecessors—Sally Buchanan's *Disaster Planning, Preparedness and Recovery for Libraries and Archives*, Judith Fortson's *Disaster Planning and Recovery*, and Camila Alire's *Library Disaster Planning and Recovery Handbook*—in providing useful guidelines for pre- and post-disaster procedures.¹ —Susan Hamburger (sxh36@psu.edu), Pennsylvania State University, University Park

Reference

1. Sally Buchanan, *Disaster Planning, Preparedness and Recovery for Libraries and Archives* (Paris: General Information Programme and UNISIST, United Nations Educational, Scientific, and Cultural Organization, 1988); Judith Fortson, *Disaster Planning and Recovery: A How-to-Do-It Manual for Librarians and Archivists* (New York: Neal-Schuman, 1992); Camila Alire, ed., *Library Disaster Planning and Recovery Handbook* (New York: Neal-Schuman, 2000).

Out-of-Print and Special Collections Materials: Acquisition and Purchasing Options. *The Acquisitions*

Librarian, no. 27. Ed. Judith Overmier. New York: Haworth, 2002. \$49.95 cloth (ISBN 0-7890-1674-5); \$29.95 paper (ISBN 0-7890-1683-4).

Guide to Out-of-Print Materials. *ALCTS Acquisitions Guides*, no. 12. By Narda Tafuri, Anna Seaberg, and Gary Handman. Lanham, Md.: Association for Library Collections & Technical Services, in cooperation with Scarecrow, 2004. 50p. \$18.50 paper (ISBN 0-8108-4974-7).

Despite the similarity of their titles, *Out-of-Print and Special Collections Materials: Acquisitions and Purchasing Options* and *Guide to Out-of-Print Materials* are two very different sorts of book, both in terms of purpose and integral quality. The former collects—or rather, amasses—a dozen articles of varying quality, from the intriguing to the banal. Adding to the confusion, the writers clearly had no shared vision of audience in mind. The uneven result is a haphazard assortment of articles, some of which—based on one's experience and interest—may well be worth reading, and several of which should be ignored.

Librarians new to collection development or acquisitions work may appreciate the primer-style contributions. Linda Fidler's "The Acquisition of Out-of-Print Music" clearly and succinctly reviews the vocabulary, formats, publishing landscape, and methods of print music acquisition. Stephen C. Wagner's "Acquiring Materials in the History of Science, Technology, and Medicine" performs similar service for its subject matter, but it also includes discussion of the development of collections policies and the selection of materials. Also of interest to new or new-to-rare-books librarians may be Marilyn Bailey Ogilvie's "Books from Abroad, One Collection Development Strategy." Focusing on the Western European trade, Ogilvie anecdotally shares common wisdom well-known to those experienced in the rare book world but that is seldom recorded in print for neophytes.

Several articles offer exemplary approaches for collection building, thoroughly describing a specific and perhaps unique collection, but in ways that may serve as models for others. In "Underground Poetry, Collecting Poetry, and the Librarian," Michael Basinski offers an extensive overview of underground poetry and its history, spelling out along the way implications for a special collection. In "Acquisitions in the James Ford Bell Library," Carol Urness details a true-to-life example of how collection scope and policies have been defined and renegotiated over time. Elaine M. Doak's "chat" titled "Decisions, Decisions, Decisions: A Tale of Special Collections in the Small Academic Library" offers another case study, one less exemplary perhaps only for its greater frankness.

While the Purdue University Libraries' Books on Demand pilot project and subsequent studies of its results are both strong and strongly provocative, the content of Suzanne M. Ward's "Books on Demand: Just-in-Time

Acquisitions” may be found elsewhere in her co-authored articles. Collection development librarians should read this material in one of its iterations.

Less helpful are the literature review articles. Editor Judith Overmier’s own “Twenty Years of the Literature on Acquiring Out-of-Print Materials” unfortunately provides mere summary rather than a critical overview. Svetlana Korolev’s “Gifts to a Science Academic Librarian” is little if anything more than a summary of others’ work.

Out of place in this book is John Alhouse’s too-brief (approximately two pages) article, “Using Older Materials in Support of Teaching.” Barbara Patterson’s “Four Factors Influencing the Fair Market Value of Out-of-Print Books: Part 1” offers no new insights, and Overmier and Wallace Koehler’s “Part 2” follow-up study provides only an example of what not to publish. Their own conclusion that their sample size is insufficient for reliable results neither fails to dissuade them from publishing their study, nor stops them from relying on this data for other conclusions.

Unlike the former book, *Guide to Out-of-Print Materials* fulfills an explicit purpose, “to provide an overview of both the traditional and online resources available to the acquisitions librarian in locating and acquiring materials that are considered to be out of print” (1), covering resources related

to books, videos, sound recordings, and serials. Presented in outline format, the guide provides a wealth of crucial information in concise, well-ordered snippets. Comprising mainly vendor contact and bibliographical resource information, its coverage with regard to English-language materials and the English-speaking world is excellent, and thus it deserves a place in any library acquisitions department.

The strong value of the guide’s information intensifies its greatest weakness—its print format limits its usefulness as a quick reference. The reader is bound by the categories into which the entries are sorted, with not even an index to aid in finding information regarding a specific vendor, Web site, or other entry. For example, while the Web sites of AddALL.com and Bookfinder are (correctly) classified as metasearch engines, this separates them from the listings for such close cousins (in practice) as Abebooks and Alibris. In addition, some Web addresses and other contact information, as might be expected, are already out of date. Librarians using this book will quickly long for an electronic format, which could be more easily searched and kept current (not to mention hotlinked). A must-read for those new to the area, the guide will yet hold value for even the most senior acquisitions librarian.—*Darby Orcutt (darby_orcutt@ncsu.edu), North Carolina State University, Raleigh*

Archival Products
B/W
P/U Jan ‘06, p 30

Library Technologies
cvr3
P/u Jan '06 c3



ALCTS

Association for Library Collections & Technical Services

a division of the American Library Association

50 East Huron Street, Chicago, IL 60611 • 312-280-5038

Fax: 312-280-5033 • Toll free: 1-800-545-2433 • www.ala.org/alcts