Against the Grain

Volume 16 | Issue 6 Article 7

December 2004

Collection Analysis Using Circulation, ILL, and Collection Data

Jennifer Knievel
University of Colorado at Boulder

Heather Wicht
University of Colorado at Boulder

Lynn Silipigni Connaway
OCLC Online Computer Library Center

Whither the Book?
from page 22

unfortunately an economic failure, and without a new business model, it is headed the way of the Dodo. The monograph we have known and loved is now long in the tooth, but there's a new one, who is a direct descendant, moving into the electronic mainstream. All it needs is recognition, academic recognition, and it may well turn the economic corner to become a self-sustaining member of the academy, too.

As **Eileen Lawrence, Alexander Press**, puts it

> This new monograph comprises not only the traditional narrative of the old monograph, but also a range of primary and secondary documents—manuscripts with introductions, contextual essays, images, journal articles, speeches, instructional materials, glossaries, links to related Websites, and so forth. It delivers all that the old monograph did, but it takes the reader out of the limits of linear exploration and allows a scholarly exploration previously impossible. It provides a model that makes it profitable for small and targeted scholarly investigations to survive in the marketplace.

It may well be that, just as the dinosaurs were succeeded by those uppity mammals scurrying beneath them, the Monograph may again prevail over that Baby Huey called Serials, but only when economics needs macro-conceptions more than micro-conceptions, synthesis more than analysis, vision more than accretion. Evolution may well see pigs fly!

Further Reading

BOOKS

Briscoe, Peter. *Reading the Map of Knowledge: The Art of Being a Librarian*. Grand Terrace, CA: Palo Verde Press, 2001.

ARTICLES

Albanese, Andrew Richard. "Moving from Books to Bytes." *Library Journal* 126, issue 14 (September 1, 2001).

"The Impact of Serial Costs on Library Collections." *ARL Bimonthly Report* 218 (October 2001): 9. *http://www. arl.org/newsltr/218/costimpact.html*

Chodorow, Stanley. "The Once and Future Monograph." Presented at The Specialized Scholarly Monograph in Crisis Or How Can I Get Tenure If You Won't Publish My Book? September 11, 1997. *http://www.arl.org/scomm/epub/papers/chodorow.html*

Mann, Thomas. "The Importance of Books, Free Access, and Libraries as Places-and the Dangerous Inadequacy of the Information Science Paradigm." *Journal of Academic Librarianship* 27, no. 4 (July 2001): 268-281.

WEBSITES

Ayers, Edward L. *The Valley of the Shadow*. University of Virginia. *http://valley.vcdh.virginia.edu/*

Sklar, Kathryn B. and Thomas Dublin. *Women and Social Movements in the United States, 1660-2000*. Alexander Street Press. *http://www. alexanderstreet2.com/wasmlive*

Collection Analysis Using Circulation, ILL, and Collection Data

by **Jennifer Knievel** (University of Colorado at Boulder); **Heather Wicht** (University of Colorado at Boulder); and **Lynn Silipigni Connaway** (OCLC Online Computer Library Center, Research)

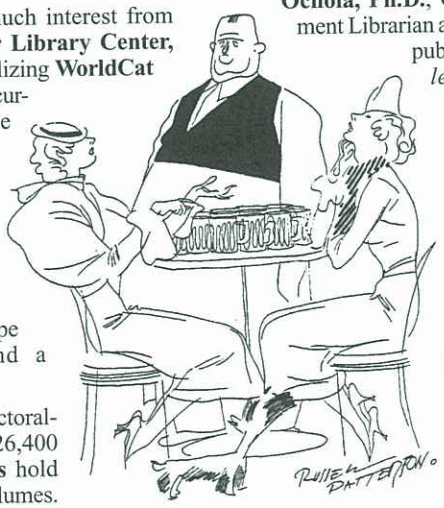
Introduction

This study was initiated to compile statistics for collection development decision making at the **University of Colorado at Boulder (CU)**. Some of the factors in decision making that are of current importance at **CU** are remote storage and budget cuts. The combination of these factors has made efficient collection management increasingly important. Like many university libraries, the libraries' shelves have been filled to capacity for some time. In 1998, **CU** began a remote storage project, and today, approximately 425,000 volumes are stored in a shared remote storage facility in Denver.

There also has been much interest from **OCLC Online Computer Library Center, Inc.** members in studies utilizing **WorldCat** holdings data. Such studies currently in progress at the **OCLC Office of Research** include library collection comparisons with gap/overlap analyses, identification of unique or last copy, and the determination of intellectual or audience level derived from type of library holdings and a weighted formula.

CU is a Research I, doctoral-granting institution with 26,400 FTEs. The **CU Libraries** hold approximately 3 million volumes. We use an Innovative Interfaces integrated library system for circulation and a **CLIO** database to track interlibrary loan (ILL) data. We are in the process of implementing the **OCLC ILLiad** software. It is already being used in ILL lending, and we plan to begin using it in ILL borrowing in early 2005.

The goal of the study was to gather and analyze holdings, circulation, and ILL borrowing data for monographs, and to compare the three data sets by common subject categories. Analyzing and comparing all three data sets makes it possible to get a more accurate picture of the usage of the monographic collection. **John Ochola, Ph.D.**, Collection Development Librarian at **Baylor University**, published a study in *Collection Management* that analyzed and compared monographic holdings, circulation and ILL data.[1] **Ochola's** intent was to use the resulting data to support decision making for the selection of monographs to be placed in remote storage. This study has incorporated some aspects of the methodology of the **Baylor** pilot project.



Scope

This study evaluated books owned by **CU Libraries**, as indicated by **WorldCat** holdings. The holdings, circulation, and ILL borrowing data from the **CU Law Library** were excluded, as the **Law Library** maintains separate integrated library and ILL systems. **CU's WorldCat** holdings were compared to book circulation data and ILL borrowing requests for books from January 1, 1998 through December 31, 2002. Only titles that circulated one or more times were included, as there was no code to identify non-circulating items. The ILL borrowing requests were harvested from the **CLIO** database. Canceled ILL borrowing requests were eliminated when the requested item was owned by the **CU Libraries**. Foreign language monographs, government documents, dissertations and theses, manuscripts, and music were not included in the data.

Approximately 20% of the ILL borrowing requests were for foreign language books. Most of these requests lacked sufficient subject classification data, which would have skewed the results of the study. Since many of these ILL borrowing requests were not filled through **OCLC**, they did not have associated **OCLC** numbers. Obtaining subject data for these ILL borrowing requests from other sources would have been extremely time-consuming; therefore, it was decided that a separate investigation was needed to specifically address these foreign language monograph requests.

continued on page 25

Collection Analysis ...
from page 24

Methodology

The **CU Libraries** worked with the **OCLC Office of Research** to determine a definition of the book in order to programmatically identify the **CU Libraries'** holdings in **WorldCat**. Based on the MARC record, a book was defined by the following criteria:

1. Bibliographic level is monograph (fixed field code "m")
2. Record type is language material (fixed field code "a")
3. No physical description fixed field present, except for text (007 position 001 is coded "a," "b," or "d")
4. If defined, the form of item can only be defined as large print or regular print reproduction (if not left blank, 008 position 23 can only be coded "d" or "r")
5. No 245 $h, which is a general materials designation used for non-book formats
6. Publisher name and/or ISBN must be present (260 $b and/or 020 must be present)
7. Book must have more than 49 pages; books with less than 49 pages are considered manuscripts (300 $a must include >49 pages)

The **WorldCat** holdings, circulation data, and ILL borrowing requests were mapped to more than 600 subject categories by **Library of Congress Classification (LCC)** number. The study incorporated the call number ranges and associated subject categories established by the **North American Title Count (NATC)**. **NATC** is a statistical collection analysis tool developed by the **Association for Library Collections & Technical Services (ALCTS) NATC** subcommittee.

For the purposes of analysis, presentation and discussion, the **NATC** categories were then consolidated into the 24 conspectus divisions. The conspectus was originally developed by the **Research Libraries Group (RLG)** and later adopted by the **Western Library Network (WLN)** and **OCLC**.

The ILL borrowing requests lacked subject data. Since they were largely processed through **OCLC**, the researchers were able to map them to **WorldCat** bibliographic records by **OCLC** number. Selected fields from the matched **WorldCat** bibliographic records were then integrated into the existing ILL data.

When the holdings data were harvested from **WorldCat**, the researchers noted a significant discrepancy between the number of **CU** holdings set in **WorldCat** and the number of bibliographic records in the **CU Libraries'** online public access catalog (**OPAC**). They learned that the **CU Libraries** had recently discovered a problem in the process of setting holdings, and were in the process of refreshing **CU's** holdings. The researchers then harvested a current holdings file for comparison and analysis. As is common in many libraries, **CU Libraries** do not set holdings on some items; therefore, a discrepancy between **CU's** local and **WorldCat** holdings still exists.

Since the ILL data were harvested from the **CLIO** database and not the **Libraries'** integrated library system, it was impossible to programmatically limit the data to books, as defined previously. As a result, many non-book materials, such as audio/visual items, sheet music, manuscripts, theses, dissertations, and foreign language materials were removed manually from the data set.

Results

The results were categorized and analyzed in four ways:

1. Overall holdings
2. Average transactions per item
3. Percentage of items circulated
4. Ratio of holdings to ILL requests

Overall Holdings

We collected our overall holdings according to **WorldCat**. We did not count multiple copies, in order to keep multiple copies of any particular title from falsely inflating a record of our holdings in any subject. Though we mined our holdings from **WorldCat**, they could also be pulled locally from a library system. It is not surprising that many of the subjects with very large collections are also subjects with a very high publishing output, such as history and literature. Some other subjects were rather small, not because they are not active subjects, but because publication in those areas is concentrated in the journal literature (e.g., chemistry) or non-book formats (e.g., musical scores).

Average Transactions Per Item

An average of transactions per item in each subject area was calculated. Each circulation was counted as a transaction. All of the transactions were assigned to a subject area, and the number of circulated items in that subject area were calculated. The number of transactions by subject was divided by the number of items circulated by subject to calculate the average number of transactions per item by subject. For example, if there were 10,000 transactions within a subject and 5,000 items circulated within that subject, it would be calculated as an average of 2 transactions per item. The highest average transactions per item in the study was 7.4, which represents an extremely active subject area. It is important to use the **NATC** subject data to evaluate where the activity is occurring. In this case, all of the circulation was occurring in two of the subfields of the discipline. The lowest average in the study was 3.1 transactions per item, which represents a low rate of transactions and may warrant an evaluation of the collection to determine if it is relevant. It is important also to evaluate the kind of usage one might expect in the subject areas. This particular subject area includes a high number of reference materials that are likely to be used in house and not be circulated.

Percentage of Items Circulated

The percentage of items circulated within the collections was calculated by dividing the number of items circulated within a specific subject by the number of items held by the **CU Libraries** within that specific subject area. For example, if the **CU Libraries** held 8,000 items within a specific subject area and 2,000 of those items circulated, then 25% of the items in that subject area circulated. The highest percentage of items circulated in the study was 43.3%. This indicates that almost half of the books circulated during the study period. Circulation was widely distributed across this subject area. The lowest percentage of items circulated during the study period was 14.9%, representing a circulation of less than one quarter of the books in that subject area. This is the same subject area discussed above and could indicate that use was mostly in-house because there are a high number of reference materials included within this subject area.

Ratio of Holdings To ILL Requests

This is perhaps the most difficult data to analyze and interpret. This ratio may be comparing apples to oranges since it is comparing what the **CU Libraries** own to what the **Libraries** do not own, but it can be a useful measurement of borrowing in relation to the size of the collection. The ratio is calculated by comparing the total items held by the **CU Libraries** within a subject area to the total number of ILL borrowing requests within that same subject area. For example, if **CU** owned 12,000 items within a particular subject area and processed 1,000 ILL borrowing requests within that same subject area, the holding to ILL ratio would be 12:1. This indicates that the **Library** borrowed one book in that subject area for every twelve books the **Library** owned within that subject area. The highest ratio in the study was 9:1, indicating that for every nine books owned within that subject area, the **CU Libraries** borrowed one book within that subject. The ratio indicates very high borrowing in this specific example. It warrants investigating if the local collection is meeting users' needs, and if acquisitions in that subject should be increased. The lowest ratio in the study was 144:1, indicating that for every 144 books owned by the **CU Libraries** within that subject, one book within the subject area was borrowed through ILL. This ratio indicates there are very few requests within this subject area, and can be interpreted in several ways. It could indicate a lack of interest by users in this subject area, or that the **CU Libraries'** collection adequately meets the users' needs.

Sample Subject Analysis: Sociology

How can these analyses be interpreted for use of the collection within a specific subject? The **CU Libraries'** sociology collection can provide an example for data interpretation. The **CU Libraries** have 43,437 holdings within the subject area of sociology, which represents a medium-sized collection. There were 6.0 transactions per item, which indicate a very high number of transactions. Forty-one percent of this subject collection circulated and the subject had a 26.0:1 ratio of holdings to ILL requests, which is also high. Reviewing the more detailed subject data for sociology can help identify the narrow subject areas with high transactions, which would be helpful in making collection and off-site storage decisions.

continued on page 26

Collection Analysis ...
from page 25

Analyzing The Collection

There are several steps to follow to replicate this analysis for another institution. One must determine the operational definition of a book and a time period for the analyses; identify the required data; analyze and compare these data; and interpret the findings.

When defining a book, it must be determined whether to include or exclude government documents, dissertations, microprint, non-circulating items, foreign language books, and branch libraries. The definition of a book needs to be consistent throughout the data sets. Determining the time period for the study may be influenced by the availability of the data for all three data sets. This study was limited in scope to the time period that ILL data were available.

Identifying the types of circulation data available is critical to the development of the study. When identifying the required data for the study, it is necessary to determine if bibliographic records or item records will be used. Using item records will introduce multiple copies to the data set. The **CU Libraries'** circulation system only tracks total transactions, not the dates of each circulation instance; therefore, it was not possible to isolate the transactions that occurred before the specified study period. The ILL borrowing requests also need to be reviewed to confirm that the data set does not include requests for books that are owned by the library.

In order to analyze, compare, and interpret the data, there must be a common identifier for every book circulated, borrowed, and owned by the library. The **OCLC** number was the only field common to all three data sets in this study. The **OCLC** number was used to acquire an **LCC** number for each book owned by the library, circulated, or requested through ILL. The data sets were combined in a database so that the **LCC** numbers could be used to classify the books into **Conspectus** and **NATC** subject areas and to analyze and compare the three data sets.

Conclusion

This comparison of holdings, circulations, and ILL borrowing requests for books can be a useful method for evaluating a collection, both for the purposes of collection development and for off-site storage decision making. As library budgets continue to decline or remain flat, it becomes increasingly important that collection decisions are based on data, not simply on intuition.

References

1. John N. Ochola, "Use of Circulation Statistics and Interlibrary Loan Data in Collection Management," *Collection Management* 27, no.1 (2002): 1-13.

Something Ventured, Something Gained: Acquiring Resources for a New Doctoral Program

by **Charles S.L. Marlor** (Head, Acquisitions & Serials Dept., Central Connecticut State University) <marlorc@ccsu.edu>
and **Lynn Johnson-Corcoran** (Collection Development Librarian, Central Connecticut State Univ.) <corcoran@ccsu.edu>

One of the **Lively Lunch** programs at the **2004 Charleston Conference** dealt with the trials and tribulations experienced by the Library at **Central Connecticut State University** in purchasing materials and resources for the school's first doctoral program. Herewith is a brief history of the school and the ensuing saga.

Central Connecticut State University (CCSU) is the oldest public institution of higher education in Connecticut and one of four regional universities that comprise the **Connecticut State University** system. **CCSU** was established in 1849 as the **New Britain Normal School** with its mission being the preparation of schoolteachers. In 1933 the **Normal School** became **Teachers College of Connecticut**. In 1955 a graduate school was established and in 1959 the name changed again; this time to **Central Connecticut State College**. **Central Connecticut State College** became **Central Connecticut State University** in 1983. Its current enrollment is 9,000 FTE.

Since **CCSU** was originally founded as a school of Education, it has always been strong in providing programs for aspiring and practicing teachers. The department of Educational Leadership offers a variety of masters programs, including an MS in Educational Technology and Educational Leadership. It also awards a sixth year certificate for administrators and supervisors. The doctorate program in Educational Leadership was licensed in December 2001, after much ado with the **University of Connecticut** that, up to then, was the only public institution of higher education in the state authorized to award a doctorate. The EdD program at **CCSU** and **SCSU** is designed for educators who are working full time and can only attend classes evenings, weekends, and summers. The program is designed to prepare professionals in the field of Education to become principals, superintendents, and curriculum and assessment specialists. Twenty-five students per year are currently accepted into the program and the first cohort will graduate in May of 2005.

We on the Collection Development and Acquisitions end of things at **Central Connecticut State University's Elihu Burritt Library** think of year one as the "Reactive/Chaotic" year. Special funds had been allocated for library materials during the first three years of the new program. Beginning in July of 2002, $70,000 was allotted to purchase both print and electronic materials that would bring the library "up to snuff" and that would ultimately ensure the EdD program's accreditation. We learned of the existence of these special funds through our Library Director who had just returned from a meeting with the University Provost. The Provost had just remembered to tell her that the library would get $70,000.00 per year for the next three years. She also told us that the hiring of an "Education Librarian," who would be dedicated exclusively to the provision of library services to the EdD cohort, was also under consideration. This position never materialized, but the anticipation of it helped add to an already strong sense of limbo. Having been told not to worry and being the obedient little librarians that we are, we awaited further instruction and direction from the faculty in the Educational Leadership program. Unfortunately, we gradually learned that silence and confusion were our only rewards for patiently waiting for others to begin to express their "vision" as to how this chunk of change was to be spent. After eight months of merciless (yet futile) hounding of many for information on what to do with this money, the University budget office did what budget offices do near the end of a fiscal year. The EdD program's collection development money was redirected to the account from which the **University** pays its utility bills. After much ringing of hands, pulling of hair, and providing an extensive wish list to the University's Finance Director a portion of the money was returned. As a result of this intervention, we were able to purchase about $5,000 worth of monographs, upgrade **Ebsco's Academic Search Elite** to **Academic Search Premier**, and subscribe to **PsychArticles**. These orders were placed at the virtual speed of light so that the small portion of money returned would be expended by the close of the fiscal year. Through our first year experience we learned that being complacent librarians waiting for direction from others was an enormous mistake. At the beginning of the second year we took the opposite approach to the issue. We let people see our proactive, creative, we're-gonna-roll side, and the results were quite satisfying.

We call year two our "Proactive/Productive" year. The **University** finally named a Director for the EdD, **Karen Beyard**. **Dr. Beyard** is a senior member of the faculty who happens to hold an MLS. She is also incredibly smart and has more energy and drive than most people on the planet. We realized by now that the pro-

continued on page 28