
TECHCAST

Welcome to the new *Collection Management* column Techcast. With less time to devote to collection responsibilities, librarians are casting about for ways to streamline collection management activities. At the same time, rapidly changing information technology brings forth new software, Web applications, and innovative practices that will potentially streamline and enhance collection management activities. The Techcast column will highlight these new ideas and help librarians assess them for use at their own institutions. Our first Techcast column examines the WorldCat Collection Analysis Interlibrary Loan Analyses module. We hope that this article and those that follow will guide librarians searching for ways to improve and streamline collection management activities.

Margaret Mellinger
Column Editor

Further Reflections on the WorldCat Collection Analysis Tool

Hilary Davis
Annette Day
Darby Orcutt

ABSTRACT. This article focuses on a recent enhancement to the WorldCat Collection Analysis tool, the Interlibrary Loan Analyses module, exploring the possibilities that this enhancement offers for strategic collection development. The study concentrates on the tool as a way to assess the impact of the recent growth at North Carolina State University in biomedicine and human medicine programs. The research contained in this article originated from a session presented at XXVII Annual Charleston Conference, which was held in Charleston, South Carolina, on November 8, 2007.

KEYWORDS. WorldCat Collection Analysis, strategic collection decisions, interlibrary loan data, interlibrary loan analyses

Hilary Davis is Collection Manager for Physical Sciences, Engineering, and Data Analysis. She holds an MLS from the University of Missouri-Columbia and an MS in biology from the University of Missouri-St. Louis (E-mail: hilary.davis@ncsu.edu). Annette Day is Associate Head of Collection Management. She earned her MLS from Leeds Metropolitan University in the United Kingdom and holds a BS in computer science and mathematics from Leeds University, United Kingdom (E-mail: annette.day@ncsu.edu). Darby Orcutt is Senior Collection Manager for Humanities and Social Sciences. He holds an MS in library science, an MA in communication studies, and a BA in speech communication and religious studies (E-mail: darby_orcutt@ncsu.edu). All are at North Carolina State University Libraries, 2 Broughton Drive, Box 7111, Raleigh, NC 27695-7111.

Collection Management, Vol. 33(3), 2008
Available online at <http://www.haworthpress.com>
© 2008 by The Haworth Press. All rights reserved.
doi: 10.1080/01462670802045566

INTRODUCTION

A 2006 article (Orcutt and Powell) reviewed the WorldCat Collection Analysis (WCA) tool as used at the North Carolina State University (NCSU) libraries. It detailed many shortcomings and necessary capabilities of any collection analysis tool for contemporary and future collection assessment. In their review, Orcutt and Powell noted that the WCA tool enabled fast gathering of data related to the holdings of a single institution (i.e., a single Online Computer Library Center (OCLC) holding symbol) and allowed comparison to a group of two to five others to quickly produce exportable title lists in targeted areas. They also reported that there were some limitations to the WCA tool. For one, the WCA's and WorldCat's use of matching on accession number instead of matching at the title level produces duplicative results. Instead of a concise list of unique titles, many multiple records for same or like items were often returned with up to 80% duplication. In addition, they reported that the tool supports "single-institution analyses rather than the cross-institutional comparisons needed for the cooperative decision-making increasingly desired in practice" (Orcutt and Powell, 2006).

This article focuses on a recent enhancement to the WCA tool, the Interlibrary Loan (ILL) Analyses module, exploring the possibilities that this enhancement offers for strategic collection development. The study concentrates on the tool as a way to assess the impact of the recent growth at NCSU in biomedicine and human medicine programs. The NCSU Libraries had anecdotal evidence of the demands placed on its collections by this growth trend but needed evidence to justify increased spending focus and funding requests. The WCA ILL data study described herein provided an opportunity to flesh out broad-level observations that originated in part as a response to the 2006 study conducted with the Triangle Research Libraries Network (TRLN) consortium. TRLN comprises NCSU, Duke University, the University of North Carolina at Chapel Hill, and North Carolina Central University, with combined resources approaching 14 million volumes and total collections budgets of more than \$29 million. The TRLN study aimed to produce actionable interpretations of our users' collections needs within the consortial context (Triangle Research Libraries Network 2006).

SETTING THE SCENE

As context for the study, some details about the NCSU Libraries are necessary. As part of the overall mission, the NCSU Libraries support more

than 31,000 students and 8,000 faculty in areas focusing on engineering, science/technology, mathematics, and veterinary medicine. NCSU is ranked third in industry-sponsored research spending, compared with all public universities without a medical school (or, at least, a medical school of the non-veterinary variety). As noted above, the NCSU Libraries are also members of the Association of Research Libraries and the TRLN consortium. All of these elements factor into the questions we asked regarding the tool and how we reviewed and interpreted the data retrieved.

DATA GATHERING AND RESULTS

To investigate the ILL Analyses module as a possible tool for guiding strategic collection development and to understand the impact of programmatic changes on the NCSU Libraries' collections, we posed a series of questions.

First, we wanted to know whether the ILL Analyses from WCA could help us identify subject areas in which there was a clear need for resources that were not already part of our collection.

The analysis represented in Figure 1 highlights the demand for health professions and public health materials, with over 20,000 requests in a 3.5-year period—nearly 17 requests per day. It should be noted that the subjects listed in the figure are based on the OCLC Conspectus of subjects, which is also based on multiple classification schemes. These are the subject divisions provided by WCA. When applying the tool, users have no control over how the subjects are defined.

Once we had identified an area of demonstrated need, namely, health professions and public health, we wanted to know whether the need for resources in that subject was format-specific, that is, for serials, books, or other formats.

Figure 2 shows that for health professions and public health, 95% of requests were for serials and only 3% were for books. Further data analysis provided us with specific title information. For this subject area, we were also able to determine which journals were in highest demand over the past 3.5 years.

Another area to analyze was the publication date of requested materials. Were the articles requested from newer or older publications? For health professions and public health, in the past 3.5 years, there was a clear indication that articles published in the 1980s and 1990s represented the greatest need for serials (Figure 3). This would imply that backfile

FIGURE 1. Number of Requests Across 3.5 Years of Data Provided by the Interlibrary Loan Analyses in WorldCat Collection Analysis Spanning Mid-2003 Through Mid-2007

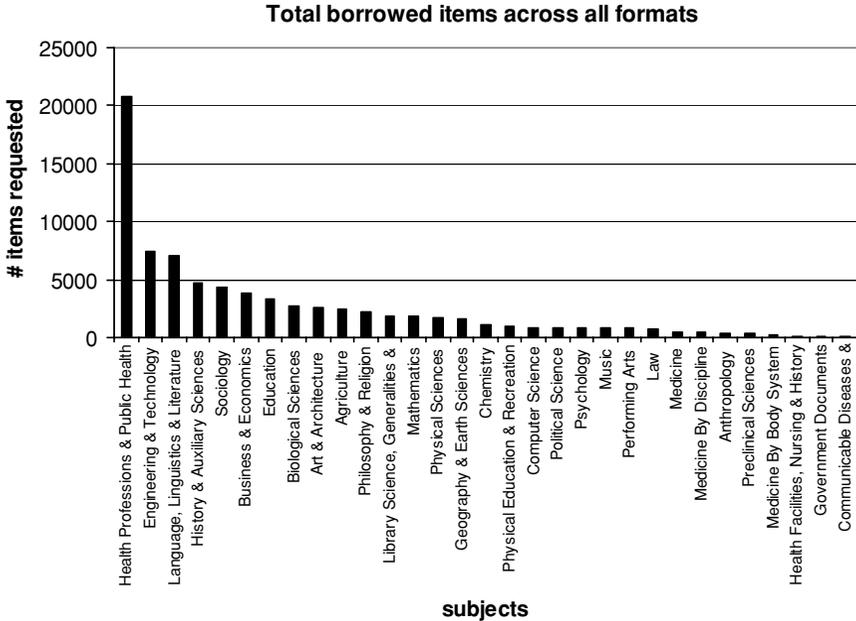


FIGURE 2. Top 10 Most Requested Subjects Based on Format

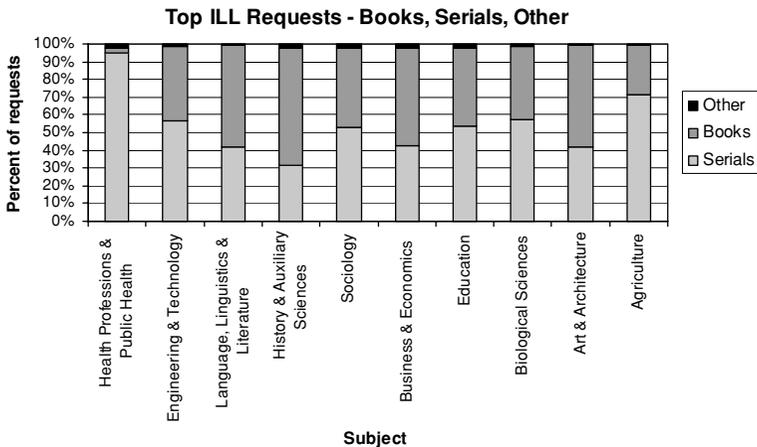
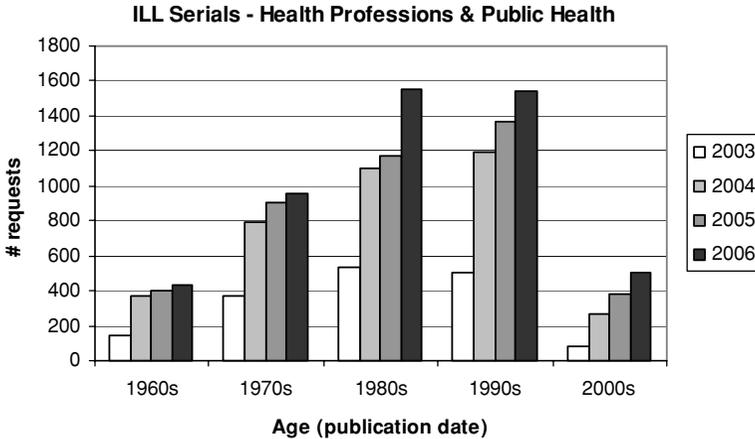


FIGURE 3. Interlibrary Loan (ILL) Requests for Serials in the Health Professions and Public Health Subject, Showing Level of Need and Age of Materials Requested



purchases may be necessary to fill gaps. We used these ILL data to identify which publishers were most highly represented, based on the journals that were requested, and mapped those data to publication dates to help determine which backfile packages were available to fill our collection gaps.

Finally, with regard to our local library consortium, TRLN, we wanted to know how many of our users' ILL requests were being filled by TRLN. Figure 4 shows the top 10 lending libraries, with the top 4 being from the TRLN consortium (representing about 49% of our ILL requests). The ILL Analyses also allows us to see the same data for any specific subject. For example, we were able to see that for requests in health professions and public health, the top four lending libraries are from the TRLN consortium (representing just over 60% of our ILL requests for this subject).

On the flip side, the ILL Analyses module allows us to look at our activity as a lender. Figure 5 shows the top 10 libraries requesting materials in health professions and public health for both books and serials. We found that we lend the greatest amount of content in the same area in which our users generate the most requests: health professions and public health. Because of the very broad nature of the *Conspectus* headings, interpretation of our lending data is a little difficult, but it is likely that we have a specialized

FIGURE 4. Top 10 Lending Libraries Across All Subjects and All Formats

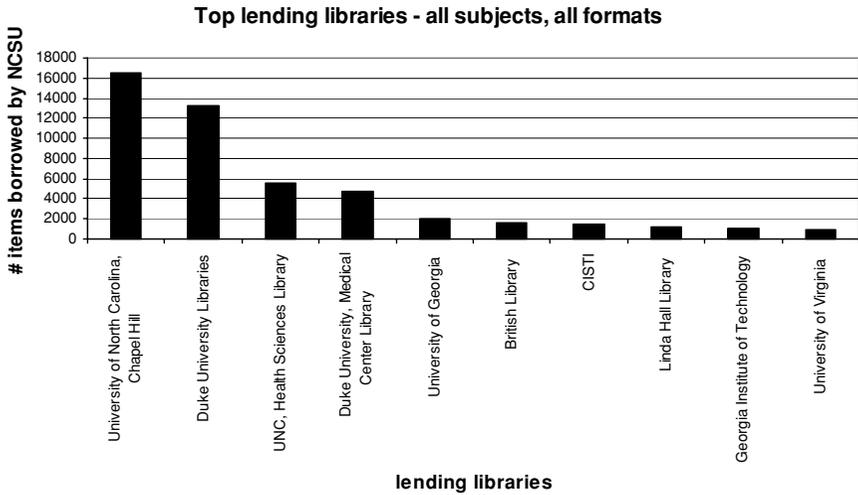
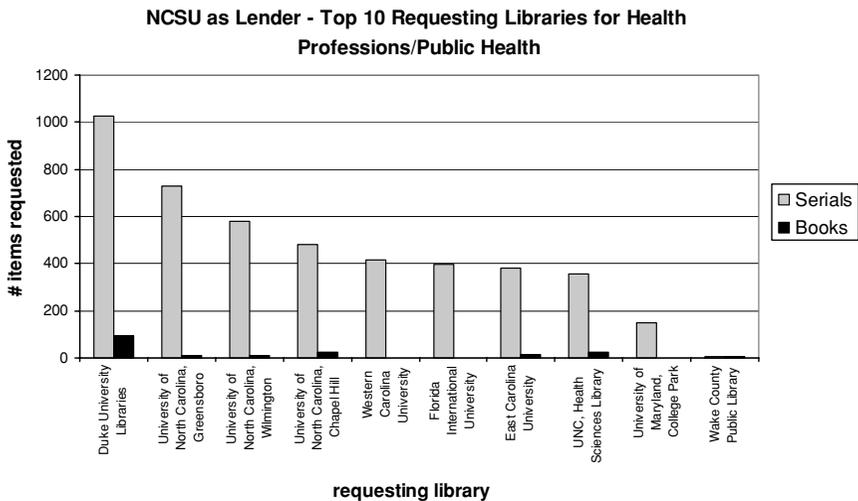


FIGURE 5. The Top 10 Borrowing Libraries in the Subject of Health Professions and Public Health for Both Books and Serials, From the Perspective of the North Carolina State University Libraries as a Lender



collection in the subject area among our history of science and toxicology and psychology collections to explain these data.

Analyzing our ILL activity with regard to TRLN provides some interesting insights into the cooperative collection development practiced by the consortium. At the NCSU Libraries, we have a relatively new research collection. It has only been in the last 15 or so years that we have had adequate funding and a dedicated collection management department to strategically build our expanding research collection. The cooperative collection building through TRLN has enabled us to provide materials to our users that supplement our collections' strengths and foci. Over time, the focus of research has changed, and materials that were supplemental to NCSU research have now become central. The demand for health professions and public health materials clearly demonstrates this change in focus. Working with these data, collection managers, in consultation with faculty and researchers, are embarking on a program to ensure that the NCSU libraries' local collections meet the developing needs of its constituents in these areas. Using this analysis to streamline and sharpen the collaborative collection building of TRLN, the ILL Analyses module of the WCA tool may provide valuable guidance in developing more meaningful cooperative collection building strategies.

WCA ILL ANALYSES AND OUR LOCAL ILL DATA

While the WCA ILL Analyses module was able to give us some meaningful data, as previously discussed, we found that to gain a full understanding of our ILL activity and its implications for collections we needed to use the WCA data in conjunction with our local ILL data. Our local data, gathered from the NCSU Libraries' ILL system, Illiad, gives us a useful level of granularity, indicating the college or department the requestor is from and whether he/she is a faculty member, graduate student, staff member, and so on. This level of granularity is not available from the WCA. Another limitation/difficulty we encountered with the WCA was the ability to retrieve ILL data based on a specific call number range. Our ILL librarian can provide us with locally generated data to tell us the ILL activity in a specific call number range. With the WCA, we had to use the broad *Conspectus* headings. However, the WCA does have the advantage of easily providing a big-picture overview of our ILL activity over a broad time period. This is not easy to obtain with our local data. While we have those precise granular data, pulling back from that detailed level to create

a broad overview would require much data manipulation and would be time-consuming.

Ideally, both sets of data are necessary, as they complement each other well. This combined approach enabled us to see the complete picture of our ILL activity and pointed us to some useful interpretations for our collections. The WCA was able to provide us with a snapshot of where there may be gaps in our collections, which we were then able to flesh out with the more granular local data. Using those local data, we could identify specific departments needing materials, and we could speak to them to learn more of their needs and demands on our collections.

DATA INTERPRETATION

Regarding data interpretation, it is important not to look at the data alone but to ensure that it makes sense with what you know historically and currently about your collection and also the programmatic and research activities on your campus.

One of the first things that we had to bear in mind was that these data may show us false gaps. For example, a book that was requested via ILL does not automatically mean that the NCSU Libraries does not hold that title in the collection. It may mean that we do not have enough copies to meet the demand. We found this to be the case when we examined the title lists generated by the analysis for some subjects. This is useful information to have, and we can approach meeting the demand in several ways in consultation with our patrons. Those ways include purchasing additional copies, buying an online version when available, or perhaps putting the book on reserve if the demand is being generated by a specific class.

We also needed to think about how our patrons use ILL. There may well be a segment of need that we do not see in the data. It is quite possible that we have not captured all the gaps in our collections from this analysis because not all of our constituents use ILL. They may use their own network of colleagues at different institutions to fax or e-mail them an article they need. They also may in some cases choose to purchase the article themselves from a publisher's Web site if either they are unaware that the library can get them the article or they perceive the process as too time-consuming when they need the article immediately.

As was previously mentioned, it is important to understand your collections to make sense of the data and to be aware of any anomalous results that may arise. The NCSU Libraries has a fairly young research collection,

and it is only in the past 15 or so years that we have been funded to develop a strong and rich research collection with a dedicated collection management department. In light of this historical context, the gaps we saw in our collections that pertained to serials from the 1970s to the 1990s made sense.

Finally, we also interpreted the results in the light of the information landscape in which we are working. Serials and e-resources are consuming a large proportion of budgets. As a result, monograph collections are beginning to feel that squeeze. In addition, as publishers' print runs get smaller and smaller, collection managers often find that if a book is not purchased within a few months of its publication, it becomes out of print; thus, it is harder to find and more expensive to purchase. The financial squeeze, coupled with short print runs, could explain why the demand for monographs focused on more recently published titles.

CONCLUSIONS

Regarding the ILL Analyses enhancement, within the context of the NCSU Libraries we found the tool to be extremely useful for providing a big-picture overview of ILL activity over a broad time period. We were able to use those overview data to generate effective graphs, which helped us describe collection needs to administrators. The overview data as communicated via graphs helped administrators understand the justification for concentrating funding on a specific area or for increasing budget requests for library resources.

Much local manipulation of the data from the WCA ILL Analyses module was needed to create the graphs/charts and title lists that we used in our analysis. As a result, we feel that there is a learning curve associated with using this tool to mine, manipulate, and interpret the data. At the NCSU Libraries, we are encouraged to use data sources such as the WCA and are given the time to work with such tools and to develop the skills to use them effectively. In other libraries, we know that many pressures and demands on librarians' time make this sort of focus difficult. Therefore, some possible enhancements to the tool would be to increase the output capabilities. Enhancing these output capabilities would make the data more usable and applicable and enable users to make graphical presentations more complex and flexible, such as those that have been presented in this article. Currently, there are some very basic graphs and charts that can be output from the tool, but they do not describe the data in as much detail

or as clearly as those we generated ourselves using the source data from the WCA. Alternatively, perhaps OCLC could offer this tool via a service model whereby OCLC would run the analyses and work with the library to produce output tailored to their needs.

REFERENCES

- Orcutt, Darby and Tracy Powell. 2006. Reflections on the OCLC WorldCat Collection Analysis tool: we still need the next step. *Against the Grain* 18, no. 4 (November): 44–48.
- Triangle Research Libraries Network. 2006. TRLN OCLC Collection Analysis Task Group: Report to the Committee on Information Resources (CIR) June, 2006. Document online. Retrieved from http://www.trln.org/TaskGroups/CollectionAnalysis/TRLN_CollAnalysis_June2Report.pdf