The Dark Side of Collection Management: Deselecting Serials from a Research Library's Storage Facility Using WorldCat Collection Analysis

Suzanne M. Ward
Purdue University, ward@purdue.edu

Mary C. Aagard
Purdue University

Follow this and additional works at: http://docs.lib.purdue.edu/lib_research
Part of the Library and Information Science Commons

http://docs.lib.purdue.edu/lib_research/134

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
The Dark Side of Collection Management: Deselecting Serials from a Research Library’s Storage Facility Using WorldCat Collection Analysis

Suzanne M. Ward
Mary C. Aagard

ABSTRACT. To address the situation of its nearly full storage facility, the Purdue University Libraries developed guidelines for the deselection of very low-use serial titles and conducted a pilot project to identify material for withdrawal. The library school intern hired to conduct the pilot used WorldCat Collection Analysis to create subject lists of titles to be considered for withdrawal. These lists also contained value-added information about how many benchmark institutions owned and duplicated the local holdings of each title. Working with subject specialists, the intern developed criteria to guide a major serials deselection project. After the success of the internship in six subject areas, the libraries funded a longer-term position to complete the deselection project.

KEYWORDS. Deselection, weeding, storage facility, WorldCat Collection Analysis, low-use serials, collection management, Purdue University

Suzanne M. Ward, MA, AMLS is Professor, Head of Resource Sharing and Mary C. Aagard, MLS is Collection Project Librarian at Purdue University Libraries.

Address correspondence to Suzanne M. Ward, Purdue University Libraries, Purdue University, 504 West State Street, West Lafayette, IN 47907. E-mail: ward@purdue.edu

Available online at http://www.haworthpress.com
© 2008 by The Haworth Press. All rights reserved.
doi: 10.1080/01462670802368638
BACKGROUND

Purdue University is a state-supported, land grant research institution with an enrollment of 39,000 for FY 2006/2007. The university is best known for its program strengths in engineering, agriculture, business, and the sciences. The Purdue University Libraries’ collection strengths reflect the focus on teaching and research in these areas. The libraries currently consist of 11 college and departmental libraries and an on-campus compact shelving storage facility called the repository. The total volume count for the entire system was 2,543,864 in FY 2005/2006; approximately 622,000 of those pieces were housed in the repository. In 2005/2006, 11,391 repository items circulated (either for in-room use or for use away from the collection); this is a rate of 1.8%.

The repository contains approximately 104,000 linear feet (almost 20 miles) of shelving and is nearly at capacity. The libraries’ administration is pursuing alternatives for acquiring additional storage space; current efforts have resulted in options available only for the short term. In the meantime, space continues to be at a premium both in the libraries’ active collections and in the repository.

The increasing pressure on the need for space was the most obvious driver for examining the existing material in the repository to assess what was still low-use and what was potentially extremely low-use to potentially no-use. However, this assessment was more generally viewed as part of normal collection maintenance activity; while the active collections are routinely reviewed for material to withdraw or to transfer to the repository, once material arrived in the repository it typically was never reviewed again. It is often difficult to think about weeding and making deselection decisions when adding material to a collection. For research libraries, some weeding is essential, though it is often viewed as a necessary evil—the “dark side” of collection development.

When Purdue’s repository opened in 1982, with its 20 miles of empty shelving, there was doubtless a natural tendency to regard the available space as inexhaustible. Some items were transferred there that might, in hindsight, better have been weeded at the time; the potential use of other transferred titles dwindled over time. Of course, many titles are ones that are entirely appropriate to keep: Purdue University publications, backruns of serials in the areas of Purdue’s program and collection strengths, and other items that receive occasional use.

In the past, two formal attempts had been made to review items in the repository. The first, about eight years ago, was an attempt to review
one small section of books that was in call number order rather than in the first-in, first-shelved order of the major portion of the collection. Subject specialists were invited to a kick-off pizza lunch and then asked to review the call number ranges in their areas of expertise. Many responded enthusiastically, but when the time came to review the shelves, it was clear that compact shelving was not intended for multiple people to use simultaneously. After that first day, very few bibliographers returned to review the books.

A larger effort took place starting in 2002. Information technology staff created an online list of all serial titles in the repository. Call numbers were divided between the subject specialists; the selectors were asked to review all the serial titles in their areas and to tag titles for withdrawal. This process was slightly more successful than the previous attempt, but interest in the project, and thus the degree of participation, differed widely among selectors. Some conscientiously reviewed their areas; some started but never finished; some never started. In hindsight, a process that asked busy librarians to review all the repository serials in their subject areas was perhaps doomed to failure or at least to uneven application. Many titles were ones that the libraries would of course retain, but they had to be reviewed along with the ones that might be candidates for withdrawal. This second effort slowly fizzled out as librarians moved on to other projects competing for their attention.

Despite these two failures, it was clear that something had to be done to identify serial titles that could be withdrawn. A new methodology would have to be devised to make retention decisions faster and easier. The new process would have a better chance of success if the subject specialists’ involvement, an essential element, was reduced to the less onerous time commitment of reviewing only those titles that were likely candidates for withdrawal rather than all the titles. However, the repository titles suggested for deselection were rarely used. A very small percentage of the serial titles had electronic equivalents, so electronic access was not used as a decision point.

Several factors coalesced to suggest a better approach:

- Concern that the repository might run out of space before new storage areas became available.
- Knowledge of hundreds of serial titles in the repository with scattered holdings, short runs, etc.
- Introduction to OCLC’s WorldCat Collection Analysis (WCA) tool, whereby a library can compare its holdings with those of its peers.
• A visit by Bernard Reilly, President of the Center for Research Libraries (CRL), who spoke about CRL’s role as a provider of low-use, specialized, and/or foreign-language research materials. He mentioned that CRL might be interested in reviewing locally withdrawn titles for possible acquisition by CRL.

• Discussion by the libraries management team that offering an internship for a library school student would be a desirable and beneficial undertaking.

• An inquiry by a former staff member, then a library school student, asking whether the libraries had a summer internship opportunity.

A potential solution presented itself: develop general guidelines varying as needed by broad subject areas for at least partially automating the identification of candidates for deselection. Examples might be repository journal titles for which the libraries held 20 or fewer volumes and for which there were at least two more or less complete runs available at several consortial partners. These consortial partners have major collections, and it would be unlikely that significant amounts of material would be weeded in the foreseeable future. Also, the partners all offer rapid-response interlibrary loan service at no direct per-transaction cost to Purdue. So in those rare instances that Purdue patrons requested items that were no longer held locally, fast resource-sharing arrangements could be made. The guidelines would be applied to the generation of OCLC’s WCA lists of repository serial titles from which obvious “keeper” titles would be removed; the reviewing selectors would only mark those remaining titles that should not be withdrawn. Employing a library school intern offered the perfect opportunity to (1) develop the guidelines and deselection criteria using WCA; (2) test the guidelines and criteria with subject specialists in selected subject areas; and (3) complete the practical process of reviewing and withdrawing titles in those selected subject areas that met both the criteria and the bibliographers’ approval. An intern could focus on this work for 180 hours over a summer, time that Purdue librarians would find difficult, if not impossible, to squeeze from their busy schedules.

The outcome would be a methodology for carrying out a better deselection project in the repository that could then be either routinized as part of normal work or handled as a special project with dedicated staff over several years. Such a methodology might also be adaptable for transfer/withdraw decisions in the active collections.
FIRST STEPS

When faced with the situation of a nearly full repository, the libraries could approach the problem in many different ways. First, the libraries could build or acquire a new storage facility. Second, the libraries could digitize the collection where allowable by copyright considerations and provide electronic access to the material. These two options were not the most attractive because the titles in the repository had already been designated as low-use titles. In fact, many titles in the repository are virtually no-use, so why go to the expense of digitizing a collection or even housing a collection that is almost never used? Next, the libraries could instate a moratorium on transfers to the repository, not a popular option when most of the active collections are also at or near capacity. Finally, the libraries could undertake a deselection project. Obviously, the problem of a full storage facility can be approached from a variety of angles and a combination of approaches can be used to remedy the situation. For the purpose of the pilot project, the libraries focused on deselection.

We also identified four libraries to serve as benchmark institutions: three libraries within the consortium of which Purdue is a member, the Committee on Institutional Cooperation (CIC), plus the CRL. As consortial partners for this project, we chose Indiana University; the University of Illinois, Urbana-Champaign; and University of Michigan; they all provide quick and reliable interlibrary loan service and they all have reputations as “keeper” libraries. In addition, Indiana University is a sister in-state institution with which Purdue has a special relationship beyond the CIC consortial one. Because of the missions and reputations of these libraries, it is unlikely that they would undertake any major deselection projects in the near to medium future. We planned to check these libraries’ holdings for the titles suggested for local deselection. Further, we predicted that those few future requests for this material would almost all be for articles; they can be delivered electronically within a day or two.

Before approaching the subject specialists, the intern’s supervisor had a few basic deselection parameters in mind for the pilot project. Many of these were based on options within WCA. For example, the date range options were used as a guide to set publication cutoff dates for reviewing titles. The initial parameters used to determine which titles would be considered for deselection were as follows:

- Titles held in their entirety only in the repository (not the repository portion of a title held partly there and partly in an active collection).
• Titles with an official start of publication in 1989 or earlier.
• Titles held by Purdue and at least two of the four benchmark institutions.
• Titles without current subscriptions (including succeeding titles).
• No U.S. government documents.
• Titles not published by an Indiana institution.
• Titles with a relatively small number of volumes held.

WCA uses the OCLC conspectus to "provide a framework to systematically inventory and describe library collections" (OCLC 2005, 1:3). The conspectus divides titles into 32 divisions, which further break down into about 500 categories and then approximately 7,000 subjects. Depending on the subject area and the number of titles held in each area, different levels of subject areas would be selected when it came time to compile the title lists of repository serials for potential deselection. Figure 1 shows the breakdown of WCA divisions.

The supervisor felt that another set of distinctions needed to be made based on material language because subject specialists may make deselection choices differently depending on the language of the titles. To address this issue, each subject was divided into three language lists: (1) English; (2) French, German, Spanish; and (3) other languages. The middle-tier languages could be changed according to the selector's needs; for example, the mathematics bibliographer wanted to have the middle language group contain French, German, and Russian. The last language group, "other languages," included the other non-English languages that WCA includes as language selections. Within each of those language groups, titles were ranked according to their compliance with the deselection criteria. First-rank titles fit the criteria exactly, second-tier titles fit most of the criteria, and third-rank titles fit some of the criteria. For example, the second and third tiers might include titles that were duplicated by only one benchmark institution, or there might be several title changes and dozens of volumes of a particular title. These titles would require more careful review by the subject specialist because they only fit some, not all, of the deselection criteria.

COLLABORATION

While creating the lists and organizing the information to be given to the subject specialists, the intern employed the expertise of several other
library staff members. The data exported from WCA were used to pull data from the Purdue catalog, such as local call numbers, BIB ID numbers, and holdings information. A library staff member in Access Services used Microsoft Access to create queries to pull the needed information from the catalog system. These data were then merged with the exported data from WCA to create a more detailed picture of the titles selected for review.

Five subject specialists in a broad range of subject areas participated in the pilot project. Certain subject areas were selected because, for the most part, a single selector could review the subject lists. In future subject areas, several selectors may need to review the lists because of the interdisciplinary nature of some subjects. The supervisor and intern selected the following subject areas for the pilot phase: agriculture, comparative literature, education, linguistics, mathematics, music, and physics. After
creating the different language lists and ranks of titles for each subject, the supervisor and the intern met with each selector and talked about the deselection criteria and explained how the spreadsheets were organized. A date was set for each selector to review the lists and to return deselection decisions to the intern.

**USING WCA**

As an example, one of the subject areas selected during the pilot phase was physics. Using WCA, the intern selected the division “physical sciences” and the category “physics, general” to search for serial titles owned by both Purdue and the other benchmark institutions. The intern searched each title to see whether it might be a deselection candidate according to the parameters. As earlier noted, the WCA preset values guided the deselection criteria. For example, the year 1989 was used as the cutoff for start of publication because WCA groups publication dates together and the next date range is too recent (1990–1994). WCA also limits language selection to 53 languages based on the *MARC Code List for Languages* and includes an “other” selection for languages that are not included in this list. Format selections are based on 17 different material types based on the MARC record leader that WCA offers for selection.

For the purpose of this project, the intern used the “Uniqueness” tab within WCA to create a list of titles owned by Purdue and also held by two, three, or four of the benchmark institutions. Figure 2 shows a screenshot of WCA showing the total number of titles held by Purdue and the number of titles shared by two, three, four, or five other institutions selected for comparison.

After a list of titles was generated, the bibliographic information was exported from WCA to a Microsoft Excel file. The exported information included title, author, Library of Congress control number (LCCN), Dewey Decimal number, International Standard Serial Number (ISSN), edition, publisher, physical dimensions, and language. Figure 3 shows a few lines of sample data from one file with some of the exported columns omitted.

Some of these information columns, such as physical dimensions and LCCN, were deleted and not used in the analysis. The data generated from the libraries’ online public access catalog (OPAC), including local call number and holdings, were then added to the WCA-generated data.

Next, the intern searched each title’s holdings in the Purdue Libraries’ catalog and in the benchmark institutions’ catalogs to create accurate notes
FIGURE 2. Screen Shot of WorldCat Collection Analysis
FIGURE 3. Sample Exported Data from WorldCat Collection Analysis

<table>
<thead>
<tr>
<th>OCLC Control No.</th>
<th>Title</th>
<th>ISSN</th>
<th>Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>6616104</td>
<td>Physis</td>
<td>0031-9414</td>
<td>Firenze, L.S. Olschki.</td>
</tr>
</tbody>
</table>

on holdings information. When using WCA, the intern found that Web browsers that allow tabbed pages are the most effective ones for managing different Web pages within WCA. (We used Mozilla Firefox for the project.) Because not all institutions have their serial holdings in OCLC, each title must be searched in the owning institution’s online catalog. WCA provided a link to the title record in the OPAC, and it was easier to open each page in a new tab within the browser rather than having each catalog displayed in a new window. The intern developed a shorthand, including using the OCLC symbol for each library, to note duplications and title changes (Figure 4).

After all the data were collected, each title was ranked according to how well it fit the criteria decided on by the supervisor at the beginning of the project. First-rank titles fit the criteria exactly. Second-rank titles fit all the criteria except that only one institution, instead of two, duplicated Purdue’s holdings completely. Third-rank titles fit the criteria except that they did not have any exact duplicates within the benchmark institutions. Figure 5 shows an example of this list.

If WCA indicated that Purdue owned a title as did two other benchmark institutions, but those other institutions did not adequately duplicate Purdue holdings, then these titles were not included in the deselection lists. Also, the intern removed from the deselection lists any title that OCLC indicated
FIGURE 4. Sample Exported WCA Data with Added Catalog Data

<table>
<thead>
<tr>
<th>OCLC Control No.</th>
<th>Title</th>
<th>BIB ID</th>
<th>Call No.</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1639132</td>
<td>Thin Films</td>
<td>56623</td>
<td>532.605 T346</td>
<td>v. (1-2); 1968-1972; UIU, IUL dup all</td>
</tr>
<tr>
<td>6616104</td>
<td>Physis</td>
<td>64402</td>
<td>505 P569</td>
<td>v. 1-26 (27-28); 1959-1991, EYM dupes all, IUL most</td>
</tr>
</tbody>
</table>

being held by fewer than 25 WorldCat institutions, regardless of duplication among benchmark institutions.

FEEDBACK FROM SUBJECT SPECIALISTS

Each of the five subject specialists who worked on the pilot project was enthusiastic about the deselection approach. Some responded that they appreciated that the intern had done all of the preliminary work; selectors only looked at lists of titles that were probable candidates for deselection. Most “definite keepers” had been removed from the lists in the preparation stage.

Some selectors expressed concern that colleagues in some subject areas might be more willing to deselect titles and to participate in the project
FIGURE 5. Sample Ranked Titles

<table>
<thead>
<tr>
<th>OCLC Control No.</th>
<th>Title</th>
<th>Notes</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1639132</td>
<td>Thin films</td>
<td>v. (1-2); 1968-1972, UIU, IUL dup all</td>
<td>1</td>
</tr>
<tr>
<td>7139127</td>
<td>Kinam: revista de fisica</td>
<td>v.1(1979)-v.6(1984), IUL dups all, EYM some</td>
<td>2</td>
</tr>
<tr>
<td>2267258</td>
<td>The Rockefeller Institute review</td>
<td>v. (1), 3; 1963-1965; EYM dups some; title change, Rockefeller University Review, v. 4 (5-6); 1965-1968; UIU dups all</td>
<td>3</td>
</tr>
</tbody>
</table>

than those in others. Perhaps other librarians would be less likely to be as enthusiastic about the deselection process and might decline to participate. One selector went so far as to say that her participation depended on the assurance that all subject areas would eventually be included in the deselection process. The supervisor explained that administrative support would be sought to validate the project and to create some solid guidelines for subject specialist participation if the project advanced past the pilot phase.

WCA gave the intern the ability to compare Purdue’s holdings easily with other institutions’ holdings. The tool let her quickly see where Purdue’s holdings might be duplicated by the benchmark libraries. Undertaking a deselection process like this would have been prohibitively time consuming without WCA. That is not to say that this project did not include a considerable amount of time involved in human analysis of the data. Because not all libraries have their serial holdings uploaded into WorldCat, and because WCA works with a “snapshot” of OCLC’s
FIGURE 6. Percentage of Deselected English-Language Titles from Physics, General Category

<table>
<thead>
<tr>
<th></th>
<th>Rank 1</th>
<th>Rank 2</th>
<th>Rank 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of titles</td>
<td>11</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td>Total No. of titles</td>
<td>10</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>deselected by</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bibliographer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of</td>
<td>91</td>
<td>38</td>
<td>24</td>
</tr>
<tr>
<td>deselection</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

holdings that are updated quarterly, the intern had to check the catalogs of each holding library to look for duplicates. Other institutions, such as North Carolina State University, have used WCA to conduct analyses on monograph collections in different subject areas. They reported many limitations in the product when using the tool to analyze monograph holdings that this project did not encounter (Orcutt and Powell, 2006).

After all of the subject specialists returned their lists, the intern found that 54% of all titles had been chosen for deselection. More surprisingly, 82% of first-tier titles were deselected. See Figure 6 for a summary of the final withdrawal decisions in one subject area.

At the conclusion of the pilot project, the intern presented the results to the libraries’ faculty at a collections forum. Intrigued by the pilot’s success, several other librarians immediately offered to start working in their own subject areas. Others better understood the potential for using WCA to analyze other aspects of Purdue’s collection.

WITHDRAWAL PROCESS

As serial titles were deselected, the intern realized that some of the weeded volumes would fill gaps in the benchmark libraries’ holdings. We also realized that other libraries around the world might be interested in acquiring the deselected titles. The acting assistant dean for collections and information resources contacted her counterparts at the project’s benchmark libraries (Indiana University; University of Illinois, Urbana-Champaign; and University of Michigan) to ask whether they would be interested in acquiring lacking volumes that the Purdue Libraries were weeding. The intern sent an explanatory e-mail with an attached spreadsheet of deselected
titles to each library and included a deadline for response. Librarians at all three libraries were enthusiastic about this opportunity. The Purdue University Libraries have also begun to work with CRL to offer them deselected titles. Serial titles that were already completely duplicated by the benchmark libraries and CRL were offered to other libraries through our Gifts and Exchange office.

Selected volumes were shipped to the benchmark institutions using the daily consortial interlibrary loan courier service. A distinct paper flag for each institution was placed in each volume, so that upon arrival at the receiving library, staff would know where to route the volumes.

*Collected Researches* (OCLC accession number 13571935) is a perfect example of one of the titles that met the deselection assumptions and was subsequently marked for withdrawal by the selector. The Purdue Libraries owned 10 volumes of the title, v. 15–25 (1920–1935). However, another title, *Monthly American Journal of Geology and Natural Science* (OCLC accession number 10497117) was retained by the selector because of its age (publication started in 1831) and its historical value, even though the holdings were duplicated by two of the benchmark institutions. A Chinese-language title, *Yen chiu pao kao. Memoirs of the College of Agriculture, National Taiwan University* (OCLC accession number 2564498) was withdrawn and several issues were sent to CRL to fill in gaps in their holdings of the title.

**NEXT STEPS**

Since the pilot project proved so successful in identifying serial titles for potential deselection and in making the best use of the subject specialists' review time, the supervisor proposed a two-year project to continue the work on a full-time basis. The project would need a coordinator as well as dedicated clerical assistance for such tasks as call number lookups; checking and recording the benchmark libraries' holdings; preparing withdrawal forms; removing volumes from the shelves; and packing volumes for shipment to other libraries. The proposal was approved, and the intern, having graduated from library school, was hired as the collection project librarian in January 2007. She hired a clerk in March 2007.

The major focus of the two-year project is to complete the review of all serial titles in the repository. If there is time, we would like to test adaptations of the process for reviewing serial titles in the active collections and/or for those split between an active collection and the repository. In
addition, we would like to evaluate whether the process is effective for identifying monographs potentially eligible for withdrawal. Although we are using WCA for a purpose that was perhaps not envisioned when the tool was designed, we feel that WCA is a useful tool to support a process of making deselection decisions.

Another major concern is the possibility of digitizing the titles that remain in repository. In June 2007, the CIC announced their partnership with Google to scan and digitize select collections from across the CIC libraries for the Google Book Search Project. Some of Purdue’s collections, and even the titles deselected based on their availability at the benchmark institutions, are expected to be digitized in this endeavor.

**PHILOSOPHICAL ASPECTS**

Traditionally, research libraries have retained almost all material they have ever acquired: the “just in case” approach. This practice arose decades ago, when identifying other libraries’ holdings was labor-intensive and when requesting and delivering material though interlibrary loan was a slow and cumbersome process. The authors believe that it is time to reevaluate this “keep everything” mindset, even at large academic libraries.

We work in an environment that is vastly different from 50 or even 20 years ago. Electronic access, especially to journal articles through either local subscriptions or via interlibrary loan, satisfies most library patrons and is in fact preferred by many of them. (Brady et al., 2006). Many consortia or regionally close libraries have built or are considering shared storage facilities so that each institution need no longer maintain low-use material, especially low-use material that is duplicated among members. “Just in time” access is acceptable to patrons who often receive items, especially electronically delivered articles, within about the same amount of time whether they are held in a local remote storage facility or at a consortial partner library.

This article described how Purdue University librarians considered serial titles that, by virtue of their presence in the repository, had already been designated as low-use material. Titles that were relatively rare (not duplicated by the benchmark libraries and/or held by few OCLC libraries) were retained not only for occasional future use by our own patrons but also so that Purdue could provide access to others through interlibrary loan. Further, subject specialists decided to retain some titles that otherwise met the deselection criteria.
While long serial runs were automatically retained during the work of this pilot project, some of these may be considered for withdrawal in the future based on decisions involving stable electronic access to the same content and/or access to regional, consortial, or national print runs archived elsewhere. Of course, some of these longer runs with electronic equivalents may continue to be housed at Purdue as part of our consortial obligations.

So while immediate space constraints provided the impetus for undertaking this low-use serials deselection project, larger changes in access to older print serials make it an appropriate activity for a research library. Our systematic, careful, and thoughtful approach assures that subject specialists make the ultimate decisions about which materials should be retained and which can be withdrawn with minimal inconvenience to our users.

REFERENCES


FURTHER READING


