

Collaboration and Automation Support Cornell University Library's Collection

Like many of their peers, Cornell University librarians have, out of necessity, embraced the concept of doing more with less for decades. In the last decade, a close collaboration between technical services and collection development librarians has provided a source of salary savings to support the Library as it faces the specter of flat or reduced staffing budgets. In 2004, Cornell librarians and technologists developed a prototype web-based tool that would facilitate the work of both the subject specialists who select new materials for the collections and the technical services librarians responsible for procuring and organizing the new acquisitions. Cornell's prototype was adopted by OCLC in 2006 and is now available to libraries worldwide as WorldCat Selection. At Cornell, the use of a shared, web-based tool, has offered workflow improvements that have reduced the time required to provide library users with new books. Kizer Walker from Collection Development and Scott Wicks from Technical Services share their experiences surrounding the development and use of this technology that facilitates the acquisitions of new books. Libraries willing to alter traditional acquisitions workflows should achieve similar benefits.

The Technical Services Perspective (Scott Wicks)

Easy Acquisitions Made Easy. This was the title of a presentation I offered in 2005 to describe how the web tool we developed at Cornell (ITSO CUL, pronounced *it's so cool*) made it possible for staff to take advantage of technology to execute routine, 'easy' ordering activities of monographs. While the title may sound somewhat silly, what was truly silly were the labor-intensive processes and workflows highly skilled Cornell staff were following indiscriminately for the simple orders as well as the complicated ones that legitimately demanded such staff attention.

Throughout my career as a librarian, the theme of doing more work with fewer resources has been a constant companion. In order to achieve significant savings of staff resources, I've been compelled to look for patterns. Patterns offer predictability. Patterns lend themselves to rules. But most relevant to our work, patterns can be automated. In terms of library acquisitions, simple orders follow a pattern. Once a title has been selected by a subject specialist, the highly trained acquisitions staff at Cornell answer the following basic questions: does the library already own this title? Is there a MARC record to import into the local system to support the ordering function? To which vendor should the order be sent? This simple pattern of questions is applied to the majority of the 30,000 monograph firm orders placed by Cornell staff each year. We asked ourselves "why not take advantage of the pattern, automate the processes, and handle the orders as part of a batch operation rather than process these orders one-by-one?"

To achieve such an automated outcome, there is a need for specific data elements as well as the cooperation of all of the stakeholders involved from the point of selector review to receipt of the order request transmission into the book vendor's system. In 2004, when we began our system development work, these pieces were all in place:

- Selectors willing to shift from a paper-based selection routine to one supported through the world wide web
- Electronic metadata from the book vendor to support new book selection
- Book vendors willing to supply the data (export from their proprietary database) on a regular schedule
- Electronic records of our library's monographic holdings
- Batch loading application software to support loading files of MARC records that result in the creation of bibliographic, holdings, and order records in our library management system (LMS)
- Ability to detect duplicate order requests as part of the batch load protocol
- Standard MARC fields to store fund code, library shelf location, and pricing data (embedded order data, or EOD)
- Ability to transmit orders from the local LMS to the book vendor's ordering system (EDI)

Cornell staff made use of these disparate elements, but because we lacked a system that integrated all of the elements, we were missing an opportunity to leverage the benefits they could support. A selector might locate a title within a book vendor's database, print out the record, write their fund code on the print-out, and send the order request through the mail to the acquisitions staff. In this scenario, the electronic record that supports the selector purchase decision provides no value beyond selection review. With the print-out in hand, the acquisitions staff would need to search the local catalog to determine whether the selector's request is already held by the library. These staff then would need to search for an electronic record to import into the local LMS (Voyager), providing the necessary metadata to send to the book vendor to indicate what we wished to order. Finally, the staff would need to transcribe the fund code and price information from the selector's order request form into a purchase order line item.

The lack of integration of these disparate systems created significant time delays and multiplied the labor required to execute each order. Additionally, multi-step ordering process created multiple opportunities for error: an acquisitions staff member might misread the selector's handwriting and assign an incorrect fund code, or she might import a different record into the LMS than the original one appearing on the print-out provided by the selector. Figure 1 illustrates the traditional workflow involved in placing monographic orders at Cornell without the benefit of WCS's integrated system.

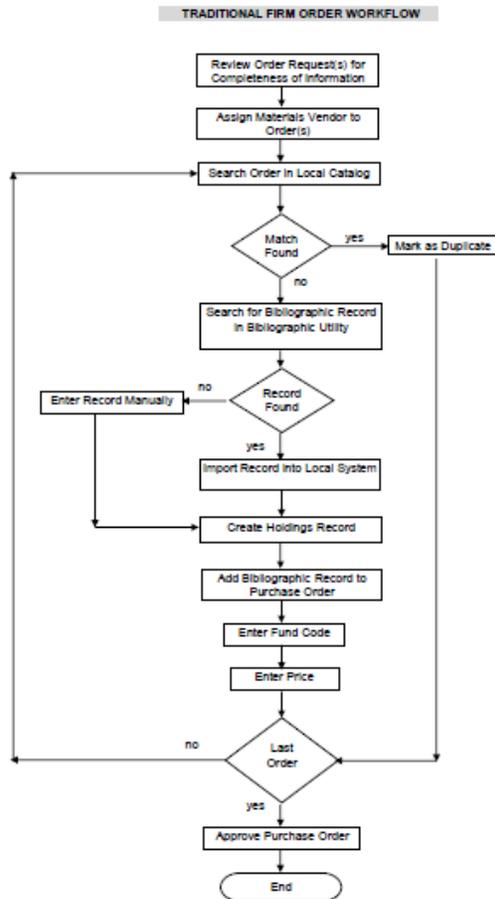


Figure 1

As a manual, one-by-one process, the time and labor involved in repeating these steps in excess of 30,000 times each year was substantial at Cornell. What is more, some components of a more automated ordering system were available to us, but useless in isolation. For example, even though an existing LMS application could support the detection of duplicates, the loading of the bibliographic and holdings data, and the creation of a purchase order as part of a batched activity, the software was not employed in the traditional workflow. It seemed a crime to have the capability, but not to take advantage of the functionality.

Cornell's response to the 'crime' was to develop ITSO CUL in 2004 (www.library.cornell.edu/backstory/v1n1/itsofeature.htm). Basically, ITSO CUL made it possible to aggregate and repurpose the electronic records each book vendor utilized to describe books they wished to sell us. Most of the major vendors in use by Cornell had their own database, each with a unique user interface. With ITSO CUL, selectors were able to create subject profiles that spanned multiple vendors' content but that could be served through a single, uniform web interface. When selectors logged into the system, they viewed all vendor records corresponding to their subject responsibilities. The selectors assigned fund codes, were able to assign prices, and key in notes to the acquisitions staff. Overnight, selector orders were harvested and loaded into the local LMS. When the

acquisitions staff began their workday, they performed some minimal review work before approving the pending purchase orders created by machine.

Building on this success, in 2006 Cornell collaborated with OCLC to make such a web tool available to the broader library community. The result of this collaboration was WorldCat Selection (www.oclc.org/selection/).

Contrast the traditional monographic order workflow illustrated in Figure 1 with what Cornell experiences by using an integrated system, displayed in the significantly simpler Figure 2. Once a selector has chosen one or more titles to order in WorldCat Selection, the acquisitions staff have very little additional work to do to complete the ordering process. The staff transaction required to place monographic orders is the same whether placing a single order or a batch of hundreds of orders.

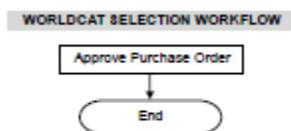


Figure 2

With an integrated system that includes WorldCat Selection (which outputs MARC records embedded with order information), LMS bulk loading applications, a locally developed program to assign orders to appropriate book vendors, and a simple cron job to automate the execution of scripts after the conclusion of the business day, Cornell has been able to achieve three major benefits. First, the Library has realized a recurring annual staff savings in excess of two full-time employees or approximately \$100,000 in equivalent value. Second, the impact of automation through the integrated system has reduced the wait time between the selector's decision to buy a title and the receipt of the order within the book vendor's system. What previously might have taken between five to seven days in a paper-based, one-by-one ordering workflow can now be executed within twelve hours when the order is processed within the integrated system. A third benefit is related to an improved level of accuracy in matching the selector's expectation to the resulting order. The metadata used to indicate to the book vendor what specific manifestation of a title is desired by the library is derived from same record viewed and chosen by the selector. The fund code information is assigned by the selector from a pull-down menu thus avoiding any miscommunication owing to illegible handwriting or missing information. And because the selected record moves directly from the selector into the LMS, there is no opportunity for record loss as could occur when a printed order request is mailed to the acquisitions staff where it could be misplaced if not lost in transit.

The integrated system is used for the 'easy' acquisitions which account for a slightly more than half of the monographs ordered each year at Cornell. In addition to these easy orders, Cornell acquisitions staff still contend with several thousand more complex monographic orders that are not served out

from WorldCat Selection, orders that follow the traditional workflow. Selectors continue to use publisher catalogs to alert them to content not held at Cornell. They send to the acquisitions staff computer print-outs, e-mail messages, and specialized book dealer catalogs with their selection decisions. They receive purchase requests from faculty and students that form the basis for rush orders.

After designing a web tool for the easy acquisitions, Cornell University Library staff, in collaboration with colleagues at Columbia University Libraries (2CUL), are beginning to explore mechanisms that will capture and automate several steps for a large percentage of those remaining monographic orders that continue to follow the traditional acquisitions workflow. We will be certain to share our findings with the broader library community.

The Collection Development Perspective (Kizer Walker)

Cornell selectors receive information about newly published books from several sources: the Library receives hundreds of publishers' catalogues, which an administrator sorts by subject and forwards to the appropriate selectors; selectors receive email notices directly from publishers and from booksellers; selectors subscribe to subject-specific reviews or email discussion lists where new titles are announced; the vendors that supply books to the Library send new title announcements; and the Library of Congress (LC) sends notification of newly catalogued titles. The WorldCat Selection tool is now our exclusive channel for these last two sources of bibliographic information, and for me and many of my colleagues it is the source we use the most for identifying new book titles and ordering individual books for the library.

I am Cornell's subject specialist for German literature and German history, as well as Classical Studies, Ancient Near Eastern Studies, and Archaeology. There is a profile set up for me in WCS that directs bibliographic records for new publications in these areas to my view of the database, based on the Library of Congress subject classification system that most U.S. research libraries use, as well as various keywords and vendor-specific subject codes. I can customize my profile at any time to exclude records, or take in additional records, by keyword, language, publisher, publication date, etc.

Until we began to receive the vendor and LC records in electronic form in 2004 (via the prototype for the WCS tool developed here at Cornell), selectors worked with stacks of paper notification slips, a different size and format from each source. Confirming whether the library already held a title, checking the availability of a title at other libraries, or seeking more than the basic bibliographic information about a book that these slips contained, entailed keying information from the paper slip into a database search. This cumbersome back-and-forth between paper and computer screen was then duplicated by the acquisitions staff who processed the orders. WCS provides a single, uniform interface for all of these new title announcements. I no longer have to worry that I might be placing an order from the announcements for something that the Library already owns, because WCS not only filters out announcements that match records in Cornell's catalog, it also runs a continual live status check against the catalog, so that if the Library acquires a book after it appears in my WCS queue, the system updates the record and marks the title as "held."

The WCS records are rich in hyperlinks, so many of the factors I want to consider when making a purchase decision can be easily accessed directly from the record: clicking the name of the author runs a search for other titles by that author in Cornell's catalog; another link searches the title in the WorldCat union catalog, so that I can see which other libraries own a copy of the book. These links appear in every WCS record, but some of the vendors who supply the records provide additional links that can aid my decision about whether to purchase a book. German-language publications are central to all of the academic fields I cover, so many of the records I deal with are from Harrassowitz, Cornell's vendor for publications from the German-speaking countries. Many of the records from Harrassowitz contain links to tables of contents and publishers' marketing text for the books, supplied by the German National Library. Whether or not such links are provided in the record, however, it is easy enough to open a new tab in my browser and copy and paste a title or author name or ISBN into a Google or Amazon search. When I work with WCS, I usually have a tab open for this kind of free searching for cases in which I want to look for a review of the book in question, or to see if an Amazon preview of the book is available, or if I want information about an unfamiliar publisher.

Certain WCS features allow me to collaborate with colleagues in selecting books for the library. When other selectors receive the same record that comes to me, I can see the names of the other recipients and an indication of whether they have ordered or rejected the book. If I receive a record in WCS that I think is more suitable for a colleague, I can forward it with a note.

In many respects, I appreciate being able to access publication announcements from multiple sources in one place. WCS is as portable as my laptop and it lets me order a few books for the library while I'm waiting at the airport or between sessions at a conference. But the unified presentation also has drawbacks: the WCS view tends to homogenize the books that it offers for my evaluation. The visual cues that a publishers' catalog provides about the book, its publisher, and the intended audience are stripped away, along with the context of the rest of the publisher's list. In some cases, even with the links to external data, I find I have to work harder to assess the book and its likely value to the collection. I still receive other notification sources for new books, but because the ordering process in WCS is simpler and quicker, and I seem to have less and less time at my disposal, I let most of the publishers' catalogs I receive pile up unopened on my desk.

It is important to note that WCS emerged at a time in U.S. research libraries when the work of collection development was shifting from small teams of selectors who focused full-time on building the collections in their subject specialties to a more fragmented model in which collection building is distributed among large groups of librarians for whom selection is only one piece of a multifaceted job. Multitasking selectors only have time to give sporadic attention to lists of new publications – and WCS is good for selecting in this distracted mode. WCS lets me see a greater variety of new titles than I would have time for if I had to rely only on catalogs, publishers' mailings, and other scattered sources. But I do not feel confident that I am seeing a *comprehensive* picture of what is published in the fields I cover. Constraints on my time, a shrinking collections budget, an expanding universe of publications, and changing expectations about the nature of library holdings all mean that the meticulous, all-inclusive bibliographer's approach to collection building is unrealistic for me and out of synch with my library.

WCS supports the kind of collection building that is possible for me today – it is an important factor in maintaining strong research collections in my fields.

Considering the central role WCS plays in my work as a selector, technical limitations of the tool are particularly frustrating. The chief problem is the speed of the system: the response time when issuing a command or moving between screens in WCS is much slower than what one normally encounters and expects on the Web. Searching within the system is particularly sluggish. OCLC has improved the speed somewhat since releasing WCS in 2006, but this remains a significant issue.

For Cornell librarians in both technical services and collection development, ideas about what constitutes quality and value in library service have changed over the past decade in response to changing economic realities and user needs. Manual, one-by-one handling of content no longer counts as a major measure of quality library work; we now focus on mechanisms to achieve time and labor savings through automation. Handwork-intensive selection and ordering is now mostly reserved for building and serving the special, unique, and very often hidden collections that support Cornell's research mission. The cost and time savings we achieve through automation of standard acquisitions can be redeployed to support these special collections and to build new services that our users demand – or that we anticipate they will demand in the near future. Such new services might include enhanced support for discovery of grey literature or other unpublished content, including data sets generated by Cornell faculty and students. Through automation of standard acquisitions, the technical services staff achieves its goal to do more with less. With collaboration, we achieve the goal for the whole library.