

Some Findings from WorldCat Local Usability Tests Prepared for ALA Annual, July 2009



WorldCat Local provides locally branded access through a simple search box and advanced search options to local, group and global holdings, print, non-print and electronic resources (including full text from electronic subscriptions) for academic and public library users. Search results include relevance ranking, FRBR-based edition clustering, and faceted browsing. Item details include collaborative and workflow features: lists, tags, reviews, and ratings.

Over the past two years OCLC has conducted close to a dozen tests of WorldCat Local, most with academic library users, some with public library users, and many in cooperation with pilot libraries. This summary assembles some key findings from our ongoing testing process.

Contents	Test goals	2	Item Details pages	5
	Test methods	2	• Getting print items	6
	Test partners	2	• Links to electronic resources	6
	Combining local, group, and global scope	2	• Rearrangement of high-value bibliographic data	7
	Journal article content	2		
	Simple search box	2	Collaboration and personal workflow tools	7
	Advanced search options	2	• Reviews	7
	Search result pages	3	• Recommendations	7
	• Order of search results	3	• Tags	7
	• What “relevance” means	3	• Ratings	7
	• The interplay of location and relevance	3	• Lists	7
	• Details on search result pages	4		
	• Editions and FRBR	4	Navigation	8
	• Faceted browsing	5		

Some Findings from WorldCat Local Usability Tests

Prepared for ALA Annual, July 2009



Test goals

Often our usability tests are summative: the goal is to validate designs or changes we've made. Sometimes our tests are formative: the goal is to provoke or correct our thinking about an emerging approach or design at a preliminary stage.

Test methods

We've done usability tests in a number of ways: often task-based sessions with prototype or production systems followed by questionnaires and interviews; sometimes contextual interviews, where the test participant trains the observer. Some of these were relatively large tests, with approximately fifteen participants, and others were small tests with only five participants. Often we conduct tests at colleges, universities or public libraries. Sometimes the user is there or at home, but we conduct the test from Ohio or California.

An agile development methodology is used for WorldCat Local: we design, test, pilot, test, design. It is an iterative and ongoing process.

Test partners

We do these tests in collaboration with staff at pilot sites who are often very involved in planning the tests and analyzing the findings with us. Test partners have included the University of Washington, the University of California (Berkeley, Davis and Irvine campuses), the Ohio State University, the Peninsula Library System in San Mateo, California, the Free Library of Urbana and the Des Plaines Public Library, both in Illinois. Later in July we'll be testing with Northeastern Illinois University and the University of Illinois, Springfield.

Usability tests like these have some advantages. We see what people actually do (in an artificial setting) rather than hear what they say they would do. We can probe interactively, and get participants to reflect. We can observe behavior across different institutions and over time.

These tests have some limitations too. Our samples are in some ways diverse: they come from different institutions, but they are not reliably representative. They're just loosely aligned with our developing personas and profiles for WorldCat Local by rank, discipline, age and gender. Test participants are predominantly from the social sciences and humanities; a few have been from the sciences, engineering and nursing.

Surveys and focus groups that can give a view of broader and more reliably representative samples than the usability tests summarized here are also conducted by various parts of OCLC.

Combining local, group and global scope

Academic users have a strongly favorable response to searching local, group and worldwide collections together.

For some public library patrons, local means this particular branch, not their whole city system.

Public library patrons are interested in what's "nearby" that they can get, without always recognizing the interlending partnership that supports it.

Journal article content

Both undergraduate and graduate student test participants recognized, understood and welcomed the basic concept of combining articles from different sources and combining them with books. They repeatedly mentioned the inclusion of journal article content as something they valued highly. However, most academic participants in one test (nine of fourteen) wrongly assumed that journal article coverage includes all the licensed content available at their campuses.

The addition of more licensed content to WorldCat and support for federated searching of databases outside WorldCat will enlarge the set of sources to search. We have tested metasearch prototypes; however, those results are not included here. Metasearching and multi-database searching in a single interface introduce usability challenges and need further testing.

Simple search box

Faculty praised it as "Googly." Very few participants demonstrated any concern about whether surnames should precede forenames. Repeatedly, academic users expressed appreciation for a search box where they can "just type anything." Most faculty, student and public library patrons used it often, searching from whatever page they happened to be on either within WorldCat Local or a library's home page.

Advanced search options

In a test with 14 participants (seven undergraduates and seven graduate students) only half of participants ever used the advanced search screen, and in fewer than half of the total searches—nine times in 66 searches.

When participants chose to use advanced search, they often mentioned a desire to avoid a too-large result. In this and other tests they stated a preference for searching by author

Some Findings from WorldCat Local Usability Tests Prepared for ALA Annual, July 2009



to avoid works about a person, or title to avoid finding common title words as subjects.

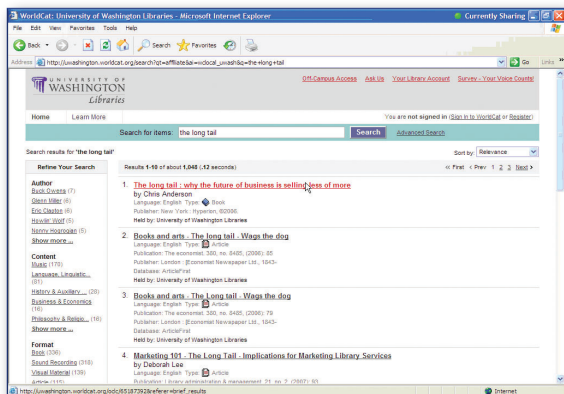
When journal article content was involved, academic test participants often mentioned a desire for a journal title index. We have now built that journal title index.

We have also made some language adjustments for academic users, saying “journal” instead of “serial, magazine, newspaper.”

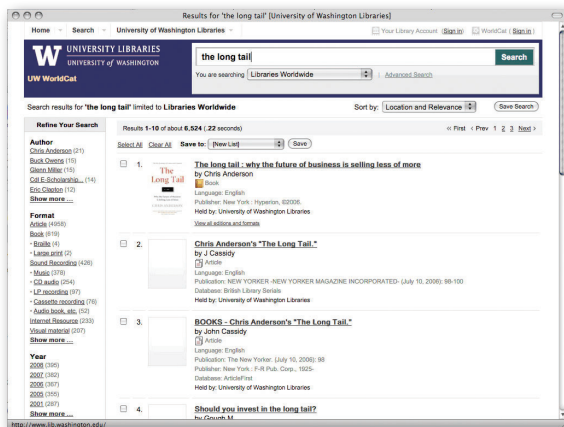
We’ve caused some problems with prompted values, and solved some problems. Early on we made search terms persist on advanced search screen (because web search engines have the user’s search terms persist into the advanced search form), but received many complaints. Now those terms are not persistent. We are also looking at alternatives for revising a search or starting a new search.

Search result pages

We had questions and got findings about the order of search results, editions and FRBR, what details we show, and facets.



Search results, May 2007



Search results, July 2009

Order of search results

We investigated what the default order should be in an academic setting, whether it differs between advanced researchers and undergraduates, what the default order should be in a public library setting, how well the default ranking supports finding known items, and how well it supports topical searching.

We found that both for students and for scholars, in both the known item case and the topical search case, the expected and preferred order was relevance. Nearly all participants expected it; a large majority preferred it. A small minority preferred date descending, and those participants were able to change the sort order. Most academic test participants recognized the order of search results in WorldCat Local as “relevance.”

For public library users, it is not yet clear from our testing what the expected order is. Some public library patrons had a hard time determining what order search results were in.

When searching for known items, test participants expected the item to be on page one of the search results. When it wasn’t, the most common behavior was to search again, using more words or quotation marks.

What “relevance” means

For topical searches, test participants—advanced researchers, faculty and graduate students—stated that the “best” items would be at the top of the list. This is “relevance” in a special sense, one that includes something like renown.

Academic test participants demonstrated more trust than curiosity about how this might be accomplished. Most faculty and graduate students could suggest several factors—relevance, popularity, and location—that might be involved. But attention to that was something they said they regarded as librarians’ expertise, not theirs. They trust their libraries to have made the right choices about how to do this for them.

Some participants remarked with surprise on the desired item not being first in search results or being preceded by an item where they couldn’t see the matching terms in the brief search display (because they were in contents notes or elsewhere). This ranking, however, was not usually an obstacle to participants getting the desired item.

The interplay of location and relevance

The default result ranking in our first tests was a subtle one. Local holdings were listed first, going from most to least relevant, followed by group holdings going from most to least relevant, followed by worldwide holdings going from most to least relevant.

Some Findings from WorldCat Local Usability Tests

Prepared for ALA Annual, July 2009



Undergraduate test participants generally report feeling well served by this ranking that puts items from the local institution at the top of the list. Participants were generally not aware that items from that collection were promoted to the top of the list.

Faculty and graduate student test participants generally did notice that local materials were at the top of the list. But they generally did not anticipate that ranking by both location and relevance would unpredictably put highly relevant items out of sight, after less relevant ones. This was a concern for scholarly test participants who were often not interested only in local materials.

In response, we created a sort option for relevance without location, still weighted by holdings but without local, group and global grouping. Now that option can be set by an institution as the default. At the same time, we added a more overt control for location: a pull-down menu near the search box. In subsequent tests, some undergraduate participants used the pull-down menu to scope results. For those who did, it did what they expected. All graduate students and faculty used this control often and proficiently.

Details on search result pages

Many participants, from both academic and public libraries, stated that they expected to see their search terms in the items at the top of the search result page. Matches in titles are visible, but matches in subjects, which happen often and can be equally important, are not visible now. Test participants expected the evidence of matching to be visible, and tended to disregard items in which their search terms were not visible.

Test participants at public libraries often needed to see holdings for the local branch, not the whole city system, on the search results page. Loading local holdings records is a remedy when this is a problem.

An indication of what's available is a basis for selection. Several participants in various tests expressed a preference to see an indication of circulation status on the search results page. However, nearly all test participants were able to determine whether locally held materials were available by examining the item details page.

Many participants expressed a desire to see an abstract or summary on the search results page. This is a change we are evaluating now.

Authors were an important basis for selection for advanced scholars. Faculty and graduate students indicated that recognizing a reputable author in their field was something they used to decide whether an item in a search result should be looked into further. None of the undergraduate

students in our contextual interviews indicated this was a factor in their selection process.

Editions and FRBR

We investigated whether test participants were misled by seeing only one manifestation on search results representing all other manifestations of that work; whether we show the right manifestation on search result pages; whether participants were confused by seeing more than one manifestation on search result pages when FRBRization is imperfect; whether participants get the latest edition of a work when they know they need the latest edition, but don't know the year of that edition.

We found that in general work-level granularity for search results based on FRBR rather than edition-level granularity fit the expectations of both academic and public library test participants. Even scholars had delegated attention to which edition to start with to us. Putting the details of one edition in search results didn't stop students or scholars from finding other editions when they wanted them. Incomplete FRBRization has a negative impact on performance, not understanding.

Undergraduate test participants, primarily in the humanities and social sciences, reported that they seldom (seven participants of ten participants) or never (three participants) looked for a specific edition of a book.

We repeatedly heard editions sometimes matter to scholars—the last, the first, the last during the author's lifetime. Historians expressed a preference for the first edition of primary sources and the most recent edition of secondary sources.

We found that the default edition matters. Even advanced researchers accepted the default when not directed to find a specific edition, without recognizing that it was not the edition they later described as preferring (i.e., the earliest or latest edition). We are now investigating whether this default should be the most recent edition that's available locally or the most widely held edition that's available locally.

When participants did have a specific edition in mind and had information (publisher, date, editor) to identify it, they generally searched with that information and saw the corresponding edition in the search result, since we show the matching edition. Nearly all academic test participants—both undergraduates and graduate students or faculty—have been successful at this sort of task.

However, in our first tests no undergraduates or graduate students were successful at finding the latest edition of a work if they didn't already know the date. This task is one instance of identifying a particular edition without knowing in

Some Findings from WorldCat Local Usability Tests Prepared for ALA Annual, July 2009



advance any words that could be used to find it by searching. Performance was almost uniformly poor for faculty, too.

We added a “view all editions and formats” link, first at the top of the item details page, and then also on the search results page. After that, undergraduates as well as advanced researchers generally found the latest edition without knowing in advance the year of that edition (even when, as often happens, that edition is not the one shown in brief search results).

Most advanced researchers expressed a desire to have all available editions listed (on the editions tab, as it then was) in reverse chronological order without regard to location. Public library patrons didn’t expect to see non-English editions, especially not at the top of that list. We created a new editions page, where users can sort by date and location, refine by format, see languages and scope by location. We changed sort of editions to remove grouping by local, then group, and then global holdings. We are considering adding more edition information for highly used/search formats and editions on the item details page.

Faceted browsing

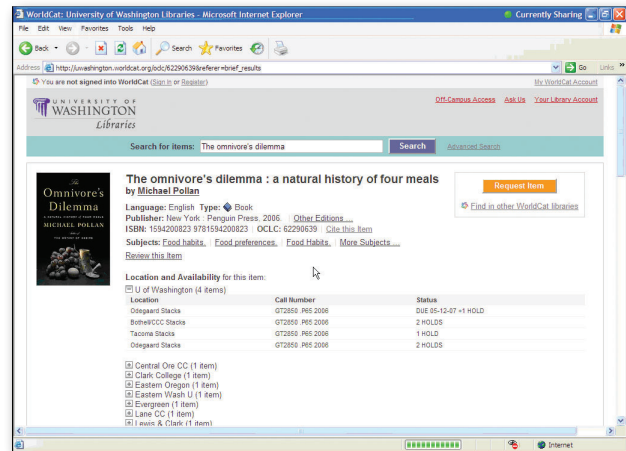
We investigated whether facets are noticed by test participants, whether they work as expected, and whether there are any obstacles to use.

Participants said and demonstrated that they noticed facets. They sometimes used facets, and facets generally worked as expected.

However, facets were more often praised than used. In subject searching in particular, facets were not often used by undergraduate or public library test participants. Three out of ten users in a test of subject searching used facets. The facets they used were language, format and year. Participants in this test did not use the subject facet (then called Content). When asked why, several said they couldn’t be sure what they would be eliminating if they did.

Subject terms at first appeared in a facet labeled “content.” Now that facet is labeled “topic.” We added a way to reverse facet choices under each facet. We changed the order of facets to put the more-often-used ones closer to the top of the page.

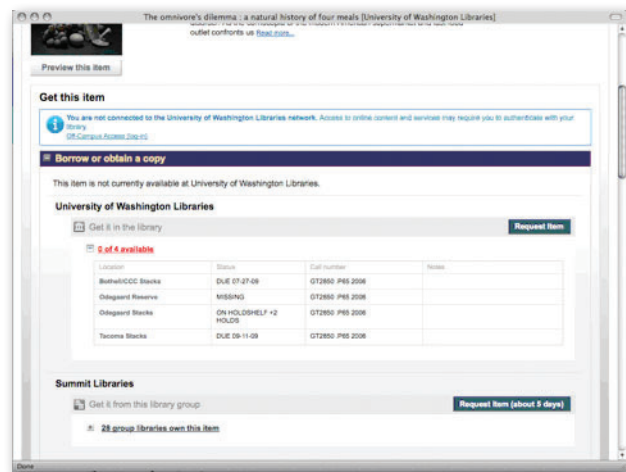
Item details pages



Item details, May 2007



Item details, July 2007



Item details, July 2009

Some Findings from WorldCat Local Usability Tests

Prepared for ALA Annual, July 2009



Some things worked well even with the earliest versions we tested, including getting print items from the local collection, at least for academic library users. Some things were problematic, including links to electronic versions and supporting academic workflows for coping and pasting citations. We undertook a major revision of this page, testing high-fidelity prototypes part way through the design process.

Getting print materials from the local collection

Can users identify what's available and where?

For academic library users, this worked well. While a few test participants expressed a preference for seeing information about availability on the search results page, going on to the item details page for this information proved not to be an obstacle. Location and availability information was visible and recognized.

Public library users wanted to see holdings for their local branch, not the whole city system. This granularity can be an issue, as it was in our early public library tests, depending on how holdings are represented. Loading local holdings records is a remedy.

Getting print materials from the group collection

In our early tests participants couldn't tell whether copies would be available from the consortial collection.

We added capacity to get availability information and set the fulfillment policy and workflow for consortial holdings, so WorldCat Local can offer a button to get a copy from the consortial collection when it is available and propose another approach when no consortial copy is available.

Academic test participants in general had a high awareness of—and reliance on—the group collection supported by a consortium or statewide system their institution might be a part of.

Some public library users were concerned almost exclusively with what was held in their local branch. Others regarded all materials they could have speedy and sure access to as equivalent. Not all these users recognized the consortial arrangements that made that possible.

Getting print materials from the worldwide collection

Undergraduate test participants saw the way to request items from outside the local consortium but, in early tests didn't seem interested in doing it. That was probably an issue with the test design. There are some barriers and some forms to fill out. These participants were not, in fact, deeply interested in the items we had them search for, nor did they really expect to get them. Activity logs show that interlibrary loan is quite accessible through WorldCat Local. Interlibrary

loan was of limited interest to most of our public library test participants.

We've moved the controls for getting local, group and global materials to clarify where the material is coming from. We're testing them in the summer of 2009 in public and academic libraries in Illinois, and we'll test them again with the University of California.

Links to electronic resources

In early tests, accessing electronic resources proved to be very problematic. At first, half of our test participants overlooked the big button for viewing online or checking for electronic resources. We moved that button.

In the next round of tests, participants demonstrated patterns of misunderstanding about the many links and icons associated with electronic resources. Those links were still dispersed on various parts of the page. Participants were not very successful in knowing when electronic resources were available, or retrieving them when they were.

We redesigned the item details page to gather together and prioritize previously scattered links for electronic fulfillment and information about availability. We removed some confusing icons. Also, WorldCat Local now allows for customization of the words shown on the open URL resolver button so they resonate with what a user sees on other pages from their library.

We tested jpeg prototypes of the revised design with six graduate students from a variety of disciplines. We saw improvements. That design is in production now, but there are still some problems. There can still be a mix of links on this page: ones that lead to fulfillment because a library has a subscription or because they're open access, and ones that are dead ends, because neither of those is the case. We recognize that this is a usability issue that can mislead users into thinking they have access to something that we don't, and we've initiated action to resolve the issue. We continue to listen to feedback to make the display of electronic links accurate for users. This poses many challenges as we try to accommodate different library cataloging practices as well as meet the expectations of the users of the system. In addition, there have been comments about how we show and hide links on the item details page. By default, we show three links and if there are more they are collapsed and the user must click on "show more" to expand and see them. We're now investigating ways to indicate which links carry the local library's guarantee of reliability and, when there are more, how to show more.

Some Findings from WorldCat Local Usability Tests

Prepared for ALA Annual, July 2009



Rearrangement of high-value bibliographic data

In our redesign of item details pages we also aimed to promote high-value bibliographic details.

When participants examined item details pages, the summaries, abstracts and tables of contents on those pages were highly valued. Participants often selected them with the mouse pointer and sometimes read them aloud. Subject headings sometimes played a similar role, supporting evaluation rather than discovery of related items.

We moved summary information to the top of the page in the redesign of the item details page. We are evaluating moving other information – perhaps contents and physical description – to the top of the page. Survey data will help us determine what information is most valuable to users in this area.

Collaboration and personal workflow tools

We did some formative testing of collaboration and personal workflow tools with jpeg prototypes and graduate students. Our challenges in introducing these collaborative features are not unique. Identity and responsibility are critical pieces for all Web sites and not unique to our tests or to libraries in general. However, academics are an audience especially aware of reputation, and academic users have spheres in which these identities and responsibility are created, outside of the library catalog.

User-contributed reviews got mixed reactions from graduate students. Response to reading reviews was mostly positive. Six of seven participants would read one. Only one of seven would write one. It matters who does write them. We saw a strong preference what participants called “editorial” reviews over user-contributed reviews in several tests. Among public library patrons (eight of them in two Illinois libraries), half said they would contribute reviews.

We changed the design of the Item Details page so that reviews are collapsed by default.

“Readers recommend” (based, at the time, on recommendations from WeRead) got less mixed—largely negative—reactions. One graduate student had a positive reaction because, he said, these recommendations might be based on connections he could not otherwise have made. Five graduate student participants had negative reactions that ranged from doubt to a sense of violation. Negative reactions had more to do with who the other readers might be than with the idea of recommendations (which were praised in MedLine) or with the specific items recommended (which were ignored). To be valued, recommendations need to come from a meaningful source. What we often heard

mattered was not just level—faculty recommendations often meant more than student recommendations to test participants—but also institution and department. And not only—or even mainly—the searcher’s own institution, but rather an institution the participant recognizes as highly regarded in his or her particular discipline.

We removed WeRead recommendations, and instead we added a recommender that looks at author, subject and series title to recommend items. These findings prompted us to conduct focus groups with librarians to explore possible sources of more authoritative reviews and recommendations.

Graduate students’ reactions to tags as a way of discovering related items were mixed, but mostly positive. Five out of seven participants would use others’ tags as a way to search. (Three participants began with a negative view and moved to a positive one.) Test participants said they valued the way tags “go beyond official subjects.” They had different ideas about what this way might be: some participants expecting tags to be broader, others narrower, or more current, than library terms. Two of seven participants would not use tags provided by others; both questioned the expertise or viewpoint of those unknown and too various others. One participant said she would add tags of her own, one said she would not. The most critical users—scholars—saw more promise in tags. The one test participant with a strong preference for the “official” subjects provided by librarians was an undergraduate.

We tested star ratings. Graduate student reactions to ratings were almost all strongly negative. Six participants would not rely on a rating. Five participants would not rate an item. The main objection here was not about who is speaking, but one about purpose. Participants pointed out that they were not reading for leisure (a context in which they might rely on ratings) but rather constructing an argument or developing a position. The fact that something is disagreeable or mistaken would not make it useless. Public library users, whose goals include leisure, might have very different reactions to ratings.

We removed Star ratings from the bibliographic area for WorldCat Local.

Graduate students valued lists. They would be valued more highly if the author were known or author’s institutional affiliation were recognized and respected. Nevertheless, in one test only one of four users created a bibliography in WorldCat Local without being told to do so. Instead, users cut and pasted references into a Microsoft Word document. Advanced scholars expressed the same preference.

The new design of the item details page now in production provides for a highlight/copy/paste workflow.

Navigation

Navigation past the second or third page of search results did not happen often, even for scholars. Although we had participants do topical searches in their areas of expertise, we could not get them to look past the first two pages of search results. Completeness of these search results and what was going on at the tail of the ranking were not a concern for our test participants.

Selection is often made from brief search result pages, without going further. When participants do go further, summary, abstract and contents notes are most valued.

All participants relied heavily on the browser's back button for navigation. Some participants expressed a desire for a more explicit way back to search results. We added a "return to search results" link to the item details.

We recognize that academic library users and public library users have different needs, goals and expectations. Many changes we have made so far have been based on the needs of academic users. As we develop and differentiate personas and profiles for academic users on the one hand and public library users on the other, we expect to create different design themes to reflect these differences. So, for example, reviews and ratings can be more prominent for public libraries using WorldCat Local than for academic libraries.