WorldCat Local Known-Item Searching Focus Group, January 10, 2010
Notes

On January 10, 2011 OCLC conducted a focus group with member library staff to discuss the nature and extent of problems with known-item searching in WorldCat Local. Participants talked about the range of users and items for which known-item searching is important, expectations about simplicity, and about a desire for control. Interrelated problems including what users know, the role of relevance ranking, the role of local and non-local holdings, the effects of edition clustering, and the choice of default edition to display were considered; solutions that might make searching more precise or results more revealing were proposed. These notes summarize the discussion and end with some use cases drawn from discussion.

Participants

Anya Arnold, Orbis Cascade Alliance
Diana Brooking, University of Washington
Sara Brownmiller, University of Oregon
John de La Fontaine, Occidental College
Allie Flanary, Portland Community College
Linda Fredericksen, Washington State University, Vancouver
Cynthia Johnson, University of California, Irvine
Jesse Koennecke, Cornell University
Theresa Mudrock, University of Washington
Christine Turner, University of Massachusetts, Amherst
Mark Vargas, St. Xavier University

Arnold Arcolio, OCLC
Jasmine de Gaia, OCLC
Mindy Pozenel, OCLC
Bob Schulz, OCLC

Discussion

There was consensus that searching for known items is basic for all academic users: undergraduates, graduate students, faculty, medical students, and circulation staff were all mentioned. It's an issue not just advanced users or expert searching.

"A lot of times they're working from a printed list...they're working from a syllabus."
"...pretty much all the students search by known item searches. They have a list of items they need to find and they type in the complete titles."

Our focus group participants all came from academic libraries, including community college libraries. Public library users might have various other goals, different sources of information, and different ideas about what constitutes a known item.

Known-item searching should be simple.

"If they have a known item, that to them should be the simplest type of search--it should be able to find it automatically."

Users of all kinds put something in the simple search box and expect to see the corresponding known item at the top of the result list. What's most important to patrons is finding known items "without having to put in a bunch of Boolean stuff." "Typing in a title or the title or beginning of the title and then finding...getting." A good discussion of the importance of advanced features like
Boolean operators is summarized below. There was agreement that meeting user expectations for a simple search and result is the baseline.

**What searchers know is not always the exact title**, and they sometimes include something additional, often an author name, in the search box.

"The user often won't know what the known item title is"
"Our syllabi are often wrong."
"The list is often based on popular title."
"70% of the time I'd categorize what they [student users] have as known building block search—they know parts of things. And they put them all in. This dude created it, and these words are in it. And I think that's also a reasonable assumption. If I knew who wrote it and these three words were in the title, I should somehow be able to marry those in a reasonably succinct fashion."
"It's a really common undergraduate research behavior that we see—and also in faculty...a dialog we have a lot is 'Okay, what you've put in the box isn't actually the title'...but it is a reasonable assumption."
"It's also a problem with cataloging procedures and what is the real title as opposed to the popular title: conference proceedings and symposia are wonderful examples of we are going to call it this but you are never going to think of it that way, but that's the official title."

A number of different problems, not always easily distinguished, were identified.

"Most everything's problematic."

Books, journals, articles (sometimes desired, sometimes eclipsing a desired book), media, and conference proceedings were all identified as potentially problematic. Although one example focused on a musical score, the difficulty of finding known items was described as "pretty general," and not focused on particular formats or subjects.

"At some point it affects everybody."
"Any book title that is a single word is extremely problematic."
"Titles that have a lot of common words in them."
"The known item that they are searching for is actually a course reserve...those don't even exist in this system"
"Searching for titles of journals. There's so much now that you're hitting is every article published in that journal, not find the location of our journal, if they're looking for something like that."

1) **When the matching known item is not held locally, it is often out of sight somewhere after the first page.**

"At what point does library become more prominent, at what point does relevance become more prominent?"
"The tipping point should be in favor of the title."

2) **Less relevant local items sometimes precede the expected item.**

"I agree with what the speakers have already said: just in terms of just trying to find a simple book even, if they know the exact title, since it's usually based on relevancy searching, other odds and ends are coming up."
"I think people are either just typing in the title--if they are a little bit more savvy they're typing the in the title and within quotations. A lot of times I don't see them actually if it's a book just looking in the book section. I think they're just going and searching broadly. So then you have the added distraction of book reviews coming up, generally before the book."
3) **Sometimes an available local copy is not discovered.**

The default edition WorldCat Local shows for the FRBR work cluster (the locally-held edition that's most widely held globally, not the locally-held edition with the most copies or the most available copies) is sometimes not available. Users often request this default edition rather than a more available local edition.

"If we own an edition that is not one of the most commonly held, it's often buried in that list of 'view all formats and editions'."

"Our particular custom edition for whatever reason is might be different is buried in the editions list because we're one of two holding libraries."

When there are multiple FRBR clusters for the same work, users often use and request the first one, sometimes one with just a few European holdings, rather than the one with hundreds of American holdings.

"This is not an isolated situation, it happens to them very frequently"

**Some proposed remedies would increase precision.**

- Limit what fields are searched as keyword to author and title.
- Match the search string as a phrase first, then roll over to joining words with and, then join words with or.
- "Reference librarians so wanted that 'title begins with' search."
- Implement stop words.

**Some proposed remedies would offer additional choices in the search process.**

- Include options to choose an index (e.g., title, author) in place of the choice of local, group, or global scope; the default choice could depend on the material-type tab chosen in the tabbed search box.

**Some proposed remedies would increase the visibility of expected matches.**

- Show the five most relevant items in each relevance level, collapsing the others.
- Bring close (left anchored) matches from level 2 to the top yet indicate that you don't own it.

**Some proposed remedies would increase the visibility of more available local editions.**

- Work level records.
- Default setting for edition page sort (library first, show the one that my library holds the most copies of).

Experience leads users to expect **known-item searching should be easy.**

"I took the list of books I had here and looked them up in WorldCat then on Google, not Amazon, on Google, and probably 90% of them were the first hit in Google, as a book."

"I did a search for a specific known item title and then it was--it didn't even show up until the third page. But then he did the same search in Amazon and it was the first thing that came up."

The predominant view was that **our primary aim should be to make known-item searching simple.**

"I worry that we're going to overcomplicate the searching process, instead of making it easier
for users to find what they want."

"They are assuming that when they put 'examining pedagogical content knowledge' in the box that there are invisible quotation marks around it."

"What theoretically looks like such a simple process is actually--to use it efficiently is incredibly complex at times, when you have to tell people, do this, and then do that, and then use this facet, and then you'll be lucky, it's number two."

"I had an example search I just did last week: examining pedagogical content knowledge. It's not like a one word title, "Seed." Type that in, now way would it come up, right? First to get rid of the articles, the first thing I do is to click the books facet, because I know it's a book. And then everything that comes is totally not that title [not held locally, item number 5]. So then I type in ti:, and that didn't work; ti:, quotes around it, and finally I force it to the top. Having to finesse that, to know to ti:, to know to put the quotes--what Mark is suggesting: having something there so the user doesn't have to know to supply those things."

"If it's for the novice user, then it should find something if I type it in, what I want. If it's for the advanced user then it should be able to do all the Boolean." "I mean even faculty members, I mean people that have come up through the old system where we've have trained them to use Boolean and everything. They think of themselves--that they are advanced users, they know, they understand, and they try, and it doesn't work. So we're always telling them in this particular database, yes, use Boolean. In that particular database, don't, make sure you use quotations."

"We [librarians] are accustomed to being able to machine things in certain ways, and we carry along our own sets of assumptions and assertions about what you can stick in a box and make things do."

"What if the system could work at both levels, without having to train people with a special language--which I think is ultimately what we hope this would be?"

"We compare it to the catalog, and it's not a catalog--which is the whole problem with known item searching. It's a discovery tool, and we have to learn to think about it differently. And maybe we have to start thinking we need to throw out some of our 'this is how we search' so that we can be open to new ways of discovering. At any rate this is sort of philosophical. It is a real problem for us. I think we get stuck in 'but this is how you do it...it should do this.'"

"It hasn't been sold to librarians yet. We are so frustrated with it, I think, that we just..." "The libraries aren't really taking to us about what their frustrations are. They're just moving to the local catalog and teaching the local catalog."

**Features like Boolean operators are less critical if the search box works as most users expect.**

"I think there's a flow. My library faculty want to use advanced search operators because the known item search has failed, and they're attempting to mediate. I could get them--they still would want Boolean to work--however, I could get them to let go of the urgency if I could truthfully tell them they'd be mediating fewer problems about finding 'I put the title and it didn't come up.'"

Participants expressed a strong interest in learning about Omniture to inform them about what is happening for their users.

"I would love a session to help us work with the Omniture data."

"Some help interpreting, because there's a lot there."
"We've now lost touch with what's going on in our catalog because we can't figure Omniture out is kind of...sad."

"I get the murmurings from librarians, but nothing specific. So we need to do, I guess, our own user group usability focus groups. But I'd love to be able to look at that data and be able to make sense of it for the non-mediated searchers."

"Through our QuestionPoint software we have a trace, a trail, of how it began, what the patron was doing, and we can see it morph through the process, so we do have some evidence, but it's only the patrons, like Theresa said, have problems, come to the desk, come to the reference chat..."

Appendix: Some proposed user stories

These use examples that came up in the focus group, or were supplied previously. The most common cases should come first in this list.

1) As a University of Washington student, searching at uwashington.worldcat.org I want to get a copy of Lady Chatterley's Lover. (It's locally held, but hard to find now.)
2) As a University of Washington student searching at uwashington.worldcat.org I want to get the book Seed. (It's not held locally, but I don't know that.)
3) As a University of Washington student searching at uwashington.worldcat.org I want to get a copy of Huckleberry Finn. (I have not doubt that's the right title.)
4) As a University of Washington student searching at uwashington.worldcat.org I want to get the book Examining Pedagogical Content Knowledge.
5) [A case where the most available edition is not default, which may change over time.]
6) As a University of Washington student searching at uwashington.worldcat.org I want to know where bound volumes of the journal Science are.