

Report of the WorldCat Local Special Collections and Archives Task Force

16 December 2008

Executive Summary

The WorldCat Local (WCL) Special Collections Task Force was convened by OCLC to make recommendations to improve discovery of special collections and archival materials in local implementations of WCL. On behalf of the national special collections community, the task force expresses its thanks to OCLC for having quickly convened the group in response to serious concerns raised by numerous special collections experts within the context of early implementations of WCL. The task force has validated those concerns and recommended solutions to each.

The rare books, manuscripts, archives, and other special collections materials in research libraries have become widely recognized as the hallmarks of distinction of individual academic libraries, as their general collections become increasing homogenized through the acquisition of licensed content and digitized books. Current initiatives of key organizations such as ARL, CLIR, and the Mellon Foundation all highlight this recognition.

Academic libraries compose the largest number of subscribers to WCL at this time, and the vast majority of special collections in the United States reside in academic libraries. They have chosen WCL in order to maximize discovery of their holdings alongside other such collections, both locally and globally. Satisfactory implementation on the part of these early adopters of WCL will be critical to the product's continued success.

WCL can enable discovery of rare and special materials only if the rich descriptions and access points created by specialized catalogers in their decades of expert work is fully revealed for both display and indexing. Key characteristics of these materials are hidden from view in WCL in a variety of ways, to the detriment of successful discovery. All information in the local catalog has previously been determined necessary for discovery in the context of each institution's collections and its overall mission of excellent service. If this does not change, research libraries will necessarily have to maintain local OPACs, to the detriment of discovery of rare and unique materials and at significant expense.

The task force therefore recommends that OCLC:

- Deploy WorldCat Local with the underlying philosophy that each institution's local records will be included in full.
- Index and display all descriptive data and access points in the member institution's local MARC records.
- Include data not represented in OCLC master records, including that in the "institutional records" of ex-RLIN libraries.
- Ensure that materials linked to Internet resources can express the format of the original materials, rather than being represented as Internet resources, for purposes of display and limiting of searches.

Further details of these recommendations appear in the full report.

The task force looks forward to ongoing dialogue and consultation with OCLC.

Jodi Allison-Bunnell, Northwest Digital Archives,
Orbis Cascade Alliance
Nicole Bouché, University of Washington
John Chapman, OCLC
David de Lorenzo, University of California Berkeley
Jackie Dooley, OCLC
Elizabeth Johnson, Indiana University
Mela Kircher, OCLC
Jennifer MacDonald, University of Delaware
Margaret Nichols, Cornell University
Jennifer Schaffner, OCLC
Nina Schneider, University of California Los Angeles
Beth Whittaker, Ohio State University

Specific Issues

The OCLC WorldCat Local Special Collections and Archives Task Force has identified an array of issues for special collections materials in WorldCat Local. They are listed in four categories, as summarized here:

Highest significance

1. Record displays are incomplete
2. Local/institutional data is not indexed
3. Principal author is not distinguished from added entries
4. Keyword searching is inadequate for voluminous results sets
5. Definition of “type” for archival materials is inaccurate

Medium significance

6. Archival holdings display as copies
7. “Buy It” link is inappropriate
8. Summary/Abstract field warrants prominence
9. Multivolume works display zero volumes

Lower significance

10. Some archival records display zero holdings
11. Some collection locations are insufficiently specific

Problems with data

12. Many NUCMC records have duplicates
13. Some finding aids have inaccurate multiple links
14. Author names are incorrectly matched in WorldCat Identities

Highest Significance

1. Record displays are incomplete

Description:

WCL does not display significant fields in archival and other special collections bibliographic records, particular 5xx fields. This information is central for understanding of the materials described, such as physical nature and extent, arrangement of materials, rights and reproductions, limits on access and use, preferred citations, etc. Some such fields are critical to users, such as restrictions on access and use (506 and 540).

In some cases, subfield \$a displays, but necessary subfields do not. For example uniform titles (130) display only subfield \$a; this is particularly problematic for rare books, given the high incidence of numerous editions (the Bible, for example). The 655 field is another example of this.

Because this content is generally indexed in WCL, but not displayed, users will not be able to determine why they have retrieved a particular result.

Recommendation:

Display in WCL all variable fields found in full WorldCat records.

2. Local/institutional data is not indexed

Description:

Indexing only the master record renders inaccessible any descriptive fields and access points the owning library has made locally. This includes key rare book indexes such as provenance, form/genre, binding, binder, publication date, and printer/publisher. Users of WorldCat Local therefore cannot compare copies, either within or across institutions, which is a central need of rare book users.

We recognize that this problem exists in part because local changes to downloaded master records are not pushed back up to WorldCat. Institutional records from ex-RLIN libraries also are excluded. We are aware that OCLC is investigating various approaches to import of local data. Whatever solution is adopted should involve minimal, if any, extra effort at the local level.

Recommendation: Index all local/institutional access points in WCL.

3. Principal author is not distinguished from added entries

Description:

The user cannot tell which name at head of the display is for the principal author or creator. This is a particularly severe problem for rare and archival materials, since bibliographic records often include numerous added entries, many of them for entities that are not authorial in nature. These include printers, publishers, editors, translators, illustrators, former owners, correspondents, and others. For archival collections, it is not unusual to have twenty or more added entries.

Recommendation:

Display only 1xx headings at head of the record. Display all 7xx headings in the "About the Authors" section.

4. Keyword searching is inadequate for voluminous results sets

Description:

WCL includes keyword search capability only, not left-anchored string searching or browse options. A search on a voluminous author or a common title retrieves so many results as to be unusable, including headings for many names other than the one of interest.

Recommendation:

Add string searching or browse options such as those included in WorldCat.

5. Definition of "Type" for archival materials is inaccurate

Description:

Limiting by "archives" as a format retrieves only a subset of archival materials. We are aware of two reasons: 1) these materials can legitimately be assigned a variety of material type codes (p=mixed, t=manuscript, k=visual, etc.). 2) The meaning of the second indicator in the 856 field has changed over time, such that the database is inconsistent. Some archival materials erroneously appear to be Internet resources.

Recommendation:

Adjust system specifications so that all records for archival materials are retrieved when limiting a search by format. Do not base assignment of "type" on the second indicator of the 856 field.

Medium Significance

6. Archival holdings display as copies

Description:

In records for archival collections, each box of material is treated as a “copy,” and the order in which boxes display cannot be manipulated to retain a logical order. The current display creates user confusion about the nature and extent of the material in the collection.

Recommendation:

May possibly be remedied by basing the “copy” designation on holdings records rather than item records.

7. “Buy It” Link is inappropriate

Description:

Inclusion of the “Buy It” link is nonsensical for archival materials, since they are unique by definition; each is held by a single institution, and they are never available for sale. Users will be confused and are less likely than otherwise to realize that the collection is available only at the owning institution.

“Buy It” often is inappropriate for rare books as well, since many of them never or rarely come to market, and few are available from Amazon.com. Unfortunately, MARC records provide no way to systematically identify rare books. Perhaps the issue can be revisited as “Buy it” functionality becomes more sophisticated.

WCL recently provided institutions the option to turn “Buy It” off in their local application. This is an all-or-none decision, however; any institution that wants to retain the functionality for current publications would not be able to turn it off for archival materials.

Recommendation:

Remove “Buy It” linking functionality for archival materials. Identify these records by appropriate material type values (see recommendation related to “type” designations).

8. “Summary/Abstract” field warrants prominence

Description:

The summary/abstract (520) is critical to understanding the nature of archival collections. A user must open the Details tab to see it, which makes it easy to overlook.

Recommendation:

Move “Summary/Abstract” field up to the basic record display following “Related subjects.”

9. Multivolume works display zero volumes

Description:

Volume holdings for multivolume works are preceded by a zero (the 2nd indicator of the MARC tag), which suggests that the library has zero volumes.

Recommendation:

Do not display the second indicator of the MARC tag for the 852 field.

Lower significance

10. Some archival records display zero holdings

Description:

Some records for archival collections display no holding library. This may be based on the OCLC subscription level for the owning institution.

Recommendation:

The task force is aware that staff in several OCLC departments is investigating this problem and hope to resolve it soon.

11. Some collection locations are insufficiently specific

Description:

In the University of Washington instance of WCL, the specific library or branch holding a collection is obscured under the consortial name “Summit Libraries.” To find the specific location, the user must click on the “+” under “Summit libraries, which is an extra step that some users will not understand. The problem is not present in the Summit instance of WCL, only the UW instance.

The problem apparently exists because Summit is a “virtual group.” Holdings are generalized among all locations/institutions that are part of the group. WCL doesn’t have the info to share.

Recommendation:

This is a low priority based on an informal survey of UO staff. OCLC perhaps should look at how consortia catalogs such as Melvyl (Univ. of California) and Summit differ from solitary library catalogs and implement a strategy for how locations data should be handled across all WorldCat local implementations.

Problems with data

12. Many NUCMC records have duplicates

Description:

NUCMC records often duplicate records contributed by member libraries. In some cases the situation is similar to what happened when OCLC absorbed WLN and the merge was oriented more strongly to non-original records at the expense of original records. Most ex-RLIN records have been loaded into WorldCat; more duplicates may occur as this process finishes.

A byproduct of this problem is that for works with both a member-contributed record and a NUCMC record, the list of library holdings includes the Library of Congress because NUCMC records were input by the Library of Congress.

Recommendation:

OCLC staff should continue to research methods of mitigating this problem.

13. Some finding aids have inaccurate multiple links

Description:

Duplicate links, i.e. a link by “web resources” and link by “Internet resources found.” This confuses anyone who wants to go from the MARC record to an online finding aid.

Recommendation:

Remove link for “Web Resources” in main record, as it is duplicated in “Internet Resources Found.” Clarify what the difference is between “Web Resources” and “Internet Resources.” OCLC might develop guidelines for implementers of WCL, based on “mistakes made by others” so as to avoid later implementations simply repeating design flaws that others have sorted out.

14. Some author names are inaccurately matched in WorldCat Identities

Description:

Imperfect disambiguation of names in WorldCat records lead to creation of “fictitious authors” in the current beta version of WorldCat Identities. The feature that links multiple disambiguated names to a single LC authority record at LC may contribute to the problem.

Recommendation:

Continue OCLC Research efforts to perfect matching algorithms in Identities. Encourage the cataloging community to report mismatches to assist in this effort.

Charge

Task Force: Special Collections and Archives in WorldCat Local
Meetings: Sept.-Oct. 2008
Report: Nov. 2008
OCLC Response: Jan. 2009

Problem:

Special Collections and Archives metadata records present unique issues within WorldCat Local. The types of materials represented by the records have significant differences from other materials. Some of the user needs for discovering, assessing, and accessing these materials are also different than for other materials. As a result, the way WorldCat.org and subsequently WorldCat Local manage and display Special Collections and Archives metadata is often not appropriate.

Charge:

This taskforce is charged with

a) identifying the highest priority problems in the user experience with Special Collections and Archives metadata in WorldCat Local

b) proposing solutions to these problems. If more than one possible solution is set forth the working group will give its recommendation as to the preferred solution.

c) classifying solutions in two categories:

1) can be fixed with metadata currently in place in WorldCat.org or services already existing in WorldCat Local,

Example: WorldCat Local can already pull local notes from the record in the local system for an integrated display with the OCLC master record.

2) changes the way WorldCat.org or WorldCat Local currently access metadata.

Example: WorldCat.org and WorldCat Local do not use Institution Records.

Outcomes:

The task force will meet for no more than two months on a schedule set by the task force. The task force will issue its recommendations to OCLC in a three weeks after the last meeting. OCLC will examine the recommendations and respond to the task for within 6 weeks. OCLC's response will include what actions it will take based on the current infrastructure and development schedule and priorities.

Task Force Members

Jodi Allison-Bunnell

Program Manager

Northwest Digital Archives, Orbis Cascade Alliance

418 Woodford

Missoula, Montana 59801

jodiab@uoregon.edu

406-829-6528

fax: 860-540-8281

Nicole Bouché

Pacific NW Curator

Special Collections Division

University of Washington Libraries

BOX 352900

Seattle, WA 98195-2900

phone: (206) 685-2856 / (206) 543-1929

fax: (206) 543-1931

nb4@u.washington.edu

John Chapman

Product Manager, Cataloging & Metadata Services

OCLC

6565 Kilgour Place

Dublin, Ohio 43017

1-800-848-5878

chapmanj@oclc.org

David de Lorenzo

Associate Director

The Bancroft Library

University of California Berkeley

Berkeley, CA 94720-6000

510.643.3203

510.643.2548 FAX

ddeloren@library.berkeley.edu

Jackie Dooley

Consulting Archivist

OCLC Programs and Research

647 Camino de los Mares Suite 108-240

San Clemente, CA 92673

949-295-1529
dooleyj@oclc.org

Elizabeth Johnson
Head, Technical Services
Lilly Library
Indiana University
1200 East Seventh Street IUB
Bloomington, Indiana, 47405-5500
(812) 855-2452
johnson1@indiana.edu

Mela Kircher
Associate Product Manager
WorldCat Discovery Services
OCLC
6565 Kilgour Place
Dublin Ohio 43017
v: 614-764-6459
f: 614-718-7364
1-800-848-5878
kircherp@oclc.org

Jennifer MacDonald.
Senior Assistant Librarian and the Coordinator, Special Collections Cataloging
and Processing Unit
University of Delaware Library
181 South College Avenue
Newark, DE 19717-5267
302-831-1512
[jmacdon@udel.edu](mailto:jsmacdon@udel.edu)

Margaret Nichols
Head, Rare Materials Cataloging Unit
Division of Rare and Manuscript Collections
2B Carl A. Kroch Library
Cornell University
Ithaca, NY 14853
(607)255-3530
mnr1@cornell.edu

Jennifer Schaffner
Program Officer
OCLC Programs and Research
777 Mariners' Island Blvd. #550
San Mateo, CA 94404
650-287-2140
schaffnj@oclc.org

Nina Schneider
Head Cataloger
Clark Library, UCLA
2520 Cimarron St.
Los Angeles, CA 90018
323-731-8529
nschneider@humnet.ucla.edu

Beth Whittaker
Head, Special Collections Cataloging Department
Ohio State University
5783 Ackerman Library
600 Ackerman Road
Columbus, Ohio 43202
247-7463
whittaker.50@osu.edu