CONTENTdm and WorldCat Digital Collection Gateway: Lessons Learned by University of Illinois at Chicago

Compiled by Kristin Martin, Metadata Librarian, Jan. 2011

Link to records representing images from UIC collection, "Image of Research:"
http://www.worldcat.org/search?q=on:IAY+uic_ior+CNTNT

Link to records representing images from UIC collection, "Century of Progress World's Fair:"
http://www.worldcat.org/search?q=on:IAY+uic_cop+CNTNT

1. The collection-level record allows all of the items in a collection to be brought together in WorldCat. All of the item-level records link back to the collection-level record and thus to each other. Some caveats:

    a. If libraries have already created a collection-level catalog record by hand in WorldCat, they should not have an additional collection-level record created through the Digital Collection Gateway. Instead, they will need to figure out whether/if they want to add that existing record's link to their new item-level records. Theoretically, it should be possible using the constant data feature in the Digital Collection Gateway, but since UIC didn't have collection-level records already, this wasn't tested.

    b. Display lists in WorldCat will group like items together, so if you have a number of images with similar names and metadata, WorldCat will group them together and treat them as "editions" of one another. The result is that the total number of items seems smaller than it really is even though every item does have an individual record.

        i. E.g., the Images of Progress collection has 1421 individual items. The results list in WorldCat makes it appear as if it has only 1149 items. All 1421 items have records; however, some are not available on the first level of browsing.

    c. Libraries that use Worldcat Local will observe that the links cause the interface to transition back and forth between the general WorldCat.org interface and the library-specific WorldCat Local Interface. This has been reported to OCLC.

2. The CONTENTdm metadata may need to be adjusted to allow for better mapping.

    a. Only fields mapped to a Dublin Core element will be able to be mapped to MARC. All new CARLI CONTENTdm collections have their fields mapped to Dublin Core elements. Any fields added or edited by your institution may have to be mapped.

    b. To the original CARLI CONTENTdm metadata, UIC added a format field to all records in the collection so that WorldCat would recognize the format as an image, making it easy to identify the images by the facet, "Narrow by Format."

        i. The Digital Collection Gateway can be used to add constant data to all records, including both the creation of new fields and the addition of prefixes and suffixes to existing fields. However, OCLC is still working out some bugs in regards to both of these features, so use them with caution.

    c. All fields are first mapped to their respective Dublin Core element, then they are available to be mapped to MARC. If a field in an item is blank in CONTENTdm, and it is one of several mapped to the same Dublin Core element, it can be mapped to a specific

MARC tag, but when there is no content, the mapping process will select the next field mapped to the same Dublin Core element and fill in that information with that field's data.

    i. For example, in the collection, "Image of Research" there is a field for awards given to a particular image. If no award was assigned, the field was left blank. This is not a problem in CONTENTdm as that field won't display to the public if there is no content in it. However, this created a problem for the mapping to Dublin Core. This field, named "Award", was mapped to DC Description, and was called Description1 within the Digital Collection Gateway, as there were multiple fields mapped to DC Description. Once in the Digital Collection Gateway, UIC added the constant data prefix "Award" to provide some context to the note and mapped it to a MARC notes field.

Example:

CONTENTdm Field	CONTENTdm Data	CONTENTdm DC mapping	DCG	Apply Constant Data	MARC
Award	2008 President's Choice	Description	Description1	Award:	5XX Award: 2008 President's Choice
Discipline	Moving media	Description	Description2	Discipline:	5XX Discipline: Moving media

    ii. However, when no award was given for the image, the Award field in CONTENTdm was blank, but the mapping process moved the field that was mapped to Description2 up to Description1 and so forth. This meant that rather than not mapping any data to the record in the DCG, instead the next Description field (which happened to be the student's discipline of study) was mapped to the MARC notes field and now had a prefix "Award," which was misleading. To solve this problem, UIC had to go back to CONTENTdm and add in content in the blank fields in order to get mapping to MARC to work correctly. If you are ultimately going to map multiple CONTENTdm fields to a single MARC field, be aware of this limitation and check your metadata in CONTENTdm.

Example:

CONTENTdm Field	CONTENTdm Data	CONTENTdm DC mapping	DCG	Apply Constant Data	MARC
Award	<none>				
Discipline	Moving media	Description	Description1	Award:	5XX Award: Moving media

3. The mapping process is slow, tedious, and buggy. It may seem like it can do everything you need it to, but most likely it can't. UIC's best advice is to map what you can, document any errors or bugs, and move on. The complete record will still be available in CONTENTdm and re-syncing is an option if the bugs get fixed in the future. For example:

    a. When deleting a mapping from one field, content from another field disappeared too.

    b. When mapping one collection, multiple 500 fields were allowed. Upon mapping another collection, only one 500 field was allowed.

    c. Personal names were mapped into the 100 field. The names were in inverted order in the original CONTENTdm metadata. Upon mapping they were flipped to direct order.

4. Currently, when the records are created, there are two links in each record that go to the individual object in CONTENTdm. One uses the CONTENTdm reference URL and one uses the WorldCat Persistent Identifier. OCLC said not to worry about this double entry and they will correct it in the future so that there is only one URL to the individual object in the record.

5. UIC did not load these records into UIC's Voyager or I-Share. At this point, for UIC, the records are available through WorldCat Local, which we think is sufficient.

    a. Advantages: More complete contents in I-Share and better accessibility to digital collections.

    b. Disadvantages: These records are not AACR-compliant; names, subjects and other access points are not in authorized form. They can be very messy and ugly.