

Catalogers Group

Minutes

November 1, 2001

Present: Angela Riggio, Beth Feinberg, Caroline Miller, Elie Chammou, Gia Aivazian, Janice Matthiesen, Jean Rashedi, Joan LoPear, John Riemer, Luiz Mendes, Nancy Norris, Rebecca Aiken, Renée McBride, Rita Stumps, Valerie Bross

Dublin Core Presentation (part 3)

John Riemer

Practice Exercise

Working in small groups for about 15 minutes, there was time to applying DCQ to slide 1087 from the Arts of the US collection, using the information found in the printed book catalog:

1087. WRIGHT, Frank Lloyd (1869-1959).

John Storer House,

1923. Los Angeles (8161 Hollywood Blvd.), Calif.

Int., fireplace. Property of Dr. & Mrs. Otto

Neurath. AC 428B

(Cf. DC view of OCLC# 44281127 in CORC). Comments coming from the discussion afterwards included:

John Storer is likely the person who commissioned the house Wright built; he is a possible candidate for Contributor.namePersonal. The date the house was built might go into Date.created.MARC21-Date, if it weren't for the fact that one is cataloging the photograph and/or its digitization. The only thing published was the digitization, so Date.issued.MARC21-Date would correspond to 1999. The street address of the house could go in Coverage.spatial, which maps to the

free-text 522 field, and data in that element can be prefaced with "Location:". The fact that the image represents an interior view, and includes a fireplace, can be brought out in a Description or Description.summary element. A Contributor access point for the building owners of about 40 years ago may not be warrant, as ownership might have changed, but a Description.note can capture the information. "AC 428B" is the photograph negative number, according to the printed catalog's introduction, and this can go in Description.note. That printed catalog can be cited in Relation.isReferencedBy.

[There was not time for practicing on slides 1745 and 1861, OCLC#s 44901251 and 44856370, respectively.]

Summary: you're likely to find more than one right selection for data elements to use and you will face ambiguities in how exactly to enter the data in the DCQ you choose. Considering that one of the raps on Dublin Core is that it does not prescribe for the content of the data, you could just relax and figure that whatever you decided to do was right. You would probably be consciously or subconsciously guided by what you do in MARC, and you would apply decisions consistently.

Mapping

A few comments about mapping: There is always some loss of data when you map from one metadata scheme to another. If this weren't the case, you probably would not have two different schemes in hand.

CORC's DCQ-MARC crosswalk is not the only one in existence. LC has one at <http://lcweb.loc.gov/marc/dccross.html>. When I first turned to it for help, it struck me as a theoretical document, the advice one would come up with based on flipping through MARC documentation looking at the various tags available. An example of how they used to differ was Coverage.spatial. LC said 522, while CORC used 691 in the early days. I was reluctant to use 691 since I didn't think that local-sounding tag would stay in the national record, especially if it had 65X fields. For the Arts of the US project, I decided I liked 522 better, since it was possible to record the exact street address of the building that was photographed.

To this day, Coverage.temporal does not go anywhere in CORC MARC. LC equates it to the 513 subfield \$b, <http://www.oclc.org/oclc/bib/513.htm>, though the time period context of that field seems to be what's in an annual/quarterly report (another stretch!).

If you are curious to know how something in DC will map to MARC, or vice versa, you can always find out by using CORC as a test bed. Tell the system you want to "create" a single record, starting from either DC or MARC, depending which way you want to test the mapping. Then flip to the other side and have a look.

My favorite memory of the Arts of the US project has been my bright grad assistant's asking me one day if she could quit flipping back and forth between the CORC DC and MARC views, since she realized she should do all the inputting on the MARC side. I was stunned. I had gone to so much trouble to maximize how much data entry could occur on the DC side. I replied, "Well, okay, but just don't tell anybody I *made* you do so." Then I laughed at my conclusion that, for some people at least, Dublin Core is "MARC Training Wheels!"

Subject Analysis

The FAST (Faceted Access to Subject Terminology) that you have heard mentioned at past Catalogers Group meetings is meant to be a compromise between totally uncontrolled keyword access and complex LCSH subject heading strings. Correct order of LCSH subdivisions can take a while for catalogers to master.

As you recall, in FAST the 4 basic parts of the subject headings are broken out into separate fields. The topic, place, and genre parts are still governed by controlled vocabulary. The time period is free to match the exact date range applicable to the resource.

In the Arts of the United States project, I felt free to have the grad assistant put the place name in Subject.geographic.LCSH. AAT terms did not feature topical or geographic subdivisions.

Authority Work

Controlled vocabulary is just as much a good idea here as with any other kind of cataloging. The limit of the authority work I have usually found in DC-based projects is taking the time to search for and match any established headings that already exist. Establishing any new names or subjects through NACO/SACO has always been beyond the time and staff resources available for the projects. For the Arts of the United States project, I sorely wanted to turn in all the new building names to SACO. (I had to ask my LC liaison if building names are established based on the newest name. He said yes, and he offered to let me base the 670 fields solely on the image in hand, plus an Internet search to see if

a newer name for the building exists, after the guidebook was published in 1960.)

One consideration for identifying and later rounding up sets of records, for sharing purposes, is including a handle like a devised corporate added entry (I should say Contributor.nameCorporate) in each of the descriptions. The current CORC system retrieval limit is 1000 records, so if more than that will be involved in a project, then some other provision will be needed. In the Arts of the US project, there were over 4000 records, so the 710 field

710 2_ Arts of the United States Collection (GALILEO (Georgia statewide project))

‡5 GAGAL

has to be searched in conjunction with the LC subject heading assigned to each chapter of 200-300 records.

Include the DC Records in the OPAC?

One question likely to arise is whether the records belong in the OPAC.

In favor of Yes:

One-stop-shopping convenience for OPAC users—do not have to discover and search a separate group of records

The authority control energy expended on them makes for compatibility; having them in same place would increase chances of heading maintenance occurring.

If 'Held by CLU' in CORC, then one would expect the records to be present with all the others.

In favor of No:

The deconstructed subject strings would not collocate well with precoordinated ones, particularly the place names.

Single searches by patrons in the future might be able to return integrated results from records residing in different files, as if they were all from the same source.

The scope of what each bib record covers would vary tremendously—from an entire journal run of 100 years to a single photo or drawing.

The quantity of item-level descriptions could overwhelm the other kind in a few years.

(Does the MARC record leader and fixed field contain the data that would enable us to sort search results in a more meaningful way, e.g. journals first, articles next, followed by monographs, and then this type of item and then that type of item?)

(If you catalog the resources in CORC, using an authorization number tied to 'CLU' symbol, then you would either have to go out of your way to prevent them from reaching the OPAC, based on ELvl 3 or 042 dc, or do a post-load deletion. The latter might be based on some parameter such as a special 049 code.)

To CORC, Or Not?

If you were approached to start up and give guidance to some DC metadata project would you advise that it be conducted in CORC? Or in a separate, local system instead (with the possibility of importing records to CORC later)? What are the pros and cons?

For CORC:

Lets others know of the resource's existence.

Gain record contribution credits (at least \$1.99/ea.)

Likelihood of URL and authority control maintenance.

Against CORC:

Cost to get the records out of the system when the work is done (export is 38 cents/ea.)

More pressure on us for timely answers to questions the project participants ask

Greater need to be monitoring what's going on with the data entry, perhaps

Have to be comfortable with non-catalogers feeding records to the OPAC.

(Is it really "non-catalogers," if people like us are guiding the enterprise?)

Some digital library projects involving Encoded Archival Descriptions (EAD) might not entail item-level descriptions and, thus, not much CORC activity.

What We Can Do With DCQ

I wanted to step back for a minute to imagine out loud what the arrival of DC can mean to the catalogers in a cataloging department. A couple of years ago where I worked a day-long conference for librarians was being planned on the topic of "outreach to the teaching faculty." The conference caused great consternation to the cataloging department because the funding was coming from a library-wide fund named for a former library director and the assumption was that "outreach" only pertained to reference and collection development librarians.

Here are excerpts of a memo I wrote to the two AULs, March 7, 1999:

"In thinking about what 'faculty outreach' could mean specifically for a librarian in Cataloging, it occurred to me that there are probably a number of professors who would be interested in receiving suggestions on how (better) to organize a collection of information objects in their possession or under their purview.

"One example is something sitting on my 'to do' list--a request from [a librarian] who runs a small library in the School of Environmental Design: she has a colleague who would like to find a way to organize for retrieval about 200 old slides. Other examples might include personal or departmental 'stashes' of image or text files, traditional books and journals, biological specimens, etc.

"An outreach encounter might consist of sitting down with an interested faculty member and conducting something analogous to a 'reference interview' to learn the nature and extent of the material, the range of methods by which s/he would like for it to be retrievable, and how widely accessible s/he desired the items to be once they were organized.

"The cataloger could then take some time to search the literature to learn how similar material has been organized elsewhere and to investigate what software might be available to assist with the task. While it would not be possible for the Department to take on any of the work, the cataloger could provide helpful advice on how the professor's staff might accomplish the necessary work, and the cataloger could stay in touch as subsequent questions arose. When the work gets completed and if use of the materials by other members of the campus community is welcomed, the cataloger could add to the OPAC a single collection-level record (analogous to that for a manuscript collection) describing the resource as well as pointing to it, should it be online.

"If members of the Cataloging Department were seriously interested in offering this service, then word could be left with colleagues in Public Services and Collection Development, such that if they encounter a desire or a need for [this type of service], they should give out contact information for the Cataloging Department. The Head of Cataloging could have on hand a list of turn-taking 'volunteers.' If initial success stories are forthcoming, judicious 'advertising' might be employed through Libraries/campus communication channels.

"Implementation of this type of outreach would enable the teaching faculty to mine an even greater range of expertise the Libraries has to offer. Growth for librarians in the Cataloging Department could occur in the areas of listening skills; learning to think in terms of getting work accomplished through others versus doing it oneself; and the mind-stretching exercise of contemplating other methods and levels of bibliographic control, beyond the particular category of MARC record used in one's primary job responsibilities."

In our setting, I am not inclined to start out with such a program for using Dublin Core. We have a Digital Library Coordinator, Howard Batchelor, who has asked me lots of challenging questions since I've arrived. What I would like is, the next time he asks for help or advice on a project, I could poll Catalogers Group and ask for volunteers who would work with me as a small subgroup, to help figure out the "answers."

There pretty much are not any DC experts at this stage, just a need to figure out something reasonable. If you were interested in participating, you would not have to count yourself in or out based on the subject matter of the project and how well that matched the print resources you're responsible for. (If you took that restrictive of a view, your turn might never come up.)

I happen to have a standing invitation from Howard to give advice on the Archive of Popular American Music (APAM) project. He would greatly appreciate any guidance, along the lines of the DC input standards I showed you for the Arts of the US project. The Music Library is about at the point where they can no longer contemplate going back to all the existing records to make changes to them. If you were to look at some the sample records and data available at <http://digital.library.ucla.edu/sheetmusic/index.html>.

(If you search for Irving Berlin's "Legend of the Pearls," for example, what DCO would you select for the data presented and what changes would you suggest?)

[John planned to send out an email asking who would be interested in forming a small subgroup to look at APAM.]

Announcements

Caroline asked people to let her know if they are interested in forming a small group to discuss and send comments in response to the call for comments from the CC:DA Task Force on ISO Harmonization. Its charge is located at: <http://www.ala.org/alcts/organization/ccs/ccda/tf-iso1.html>, and it is interested in advice on areas where AACR2 should and should not be brought more in line with ISO standards.

Likely agenda topics for upcoming meetings (Nov. 15 and 29) are

--a training session for all department staff in use of the new Cat Client (call# searching and keyword searching are undergoing major changes) (Cat Client Training Group)

--a report on the Nov. 14th AdCon Retreat on Guiding and Enabling Access to Digital Collections (John)

--short reports from those attending the Internet Librarian 2001 conference next week in Pasadena