CORES Action Plan: MARC 21 and MODS metadata elements

[May 29, 2003, submitted by Rebecca Guenther as input to a D-Lib Magazine article on the CORES Resolution.]

The Library of Congress plans to explore the possibility of providing persistent URIs for its metadata elements. This will include, initially, "http:" URIs for MARC and MODS content designators (elements, subelements, etc). As a first step, this will take the form of a statement that details the convention for naming elements. LC will explore the use of URNs as element names as a more appropriate means to provide persistent identifiers for elements, particularly since naming, rather than retrieval, is the primary objective of a metadata element identifier, and resolution is not an issue. URIs may also be provided for values in controlled MARC/MODS value lists, also initially by publishing a naming convention, and perhaps later explicitly in mark-up (with the exception of the MARC relators as detailed below).

As a result of a request by the DCMI Usage Board to make available the terms in the *MARC Code List for Relators* for use with the Dublin Core Creator and Contributor elements, the Library of Congress will establish URIs for the values on this list as an initial test case. This will initially take a form compatible with DCMI (using RDF) however in the long term it may be replaced by a more comprehensive solution. This work has already been largely completed, although a few details need to be worked out.

Establishing URIs will enable an unambiguous reference to any MARC/MODS element.

1. Assignment of MARC/MODS URIs (short-term).

1.1. MARC
Because there are different MARC formats (i.e., bibliographic, authority, holdings, classification, and community information), it is necessary to identify both that the namespace is MARC 21 and what the format is.(Note that a namespace has been established for MARCXML.) URIs for MARC elements might be constructed as follows:

http://www.loc.gov/marc.[format].[fieldname].[subfield]

Bibliographic format (subfield):
http://www.loc.gov/marc.bibliographic.245.a
would be a URI for subfield $a of field 245 of the MARC bibliographic format.

Bibliographic format (field):
http://www.loc.gov/marc.bibliographic.245

Bibliographic format (indicator):
http://www.loc.gov/marc.bibliographic.245.i1

Bibliographic format (008 element):

http://www.loc.gov/marc.bibliographic.008-s-03
for sound recording 008 character position 03.

Authority format (008 element):
http://www.loc.gov/marc.authority.008

1.2. MARC Code lists

MARC Code List for Relators. Each term on the Relators list would be assigned a URI taking the following form:
http://www.loc.gov/marc.relators.adp
where the last element is the code that represents the term on the list. The above would be the URI for "adaptor".

At a later date, this pattern would be followed for the values on other code lists, where the URI would be constructed with the code attached to a namespace. An unresolved question will be establishing these for MARC language codes, since they are equivalent to ISO 639-2/B codes. This will be considered at a later date.

1.3. Source codes

There are many places in MARC 21 that use a code to identify a source, or in Dublin Core terms, an encoding scheme. The DCMI Library Application profile recommends registering these. Currently there are several encoding schemes that have been registered to be used with Dublin Core elements that also exist in the MARC namespace (examples: lcsh, lcc). If LC assigns URIs for each of these values the redundant namespace identification will need to be worked out (i.e., are they redundantly registered in the DCMI namespace or is the MARC 21 namespace referenced? What will happen with the current overlap?).

1.4. MODS elements

Some MODS elements are already used in the DCMI Library application profile, so URI assignment for these would be useful. Those that are referenced in DCMI-Lib AP are:

Edition
Location

The namespace for MODS is:
http://www.loc.gov/mods

The above could be specified as follows:
http://www.loc.gov/mods.originInfo.edition
(Note that edition is a subelement of originInfo)
http://www.loc.gov/mods.location
It needs to be considered whether a version number should be included in the URI.

MODS elements that are subelements could be assigned URIs as follows:

http://www.loc.gov/mods.titleInfo.partNumber

2. Assignment of MARC/MODS URIs (long-term)

LC plans to assign persistent URIs utilizing URN namespaces. It is currently investigating options for naming elements within this mechanism, rather than assigning http: URIs. Examples might be:

URN:[urn namespace id]:marc.bibliographic.245.a
URN:[urn namespace id]:mods.originInfo.edition
URN:[urn namespace id]:marc.relators.adp

3. Assignment of identifiers to metadata schemas

LC plans to register handles to be used as persistent names for XML schemas which it maintains. These handles will be expressed as URLs which resolve through LC's proxy server (http://hdl.loc.gov). For example:

http://hdl.loc.gov/loc.standards/mods
http://hdl.loc.gov/loc.standards/marc21.slim

4. Timeframe

LC plans to make the relator list available (with URIs for each entity) after the DCMI Usage Board meeting in June 2003. It should be possible to establish persistence policies and URI assignment policies by the second half of 2003, pending investigation of URN namespaces. Minimally, these will be made available as a tool for referencing MARC and MODS metadata elements. Assigning URIs to MODS elements will probably take less time than to all MARC elements. It is expected that assigning URIs may take a year or so.

Rebecca Guenther
Library of Congress
Network Development and MARC Standards Office