

CARLI Board Meeting 2/29/08

Explanation for CARLI Board of new metadata requirements proposed by the Digital Collections User Group

The Standards Subcommittee of the Digital Collections User Group (DCUG) has spent several months reviewing metadata requirements for CARLI Digital Collections. Currently, the only requirements are that the metadata record for each item contain a rights statement, including contact information for the library, and a field identifying to which collection the item belongs.

Due to the popularity of the Dublin Core metadata scheme and that its use is required for any collection wanting to be harvested by an OAI Service Provider (e.g. Illinois Harvest), the Standards Subcommittee tasked itself with providing CARLI libraries a best practices document for Dublin Core metadata. Included in this document are three additional requirements that DCUG recommends be instituted for all current and future collections in CARLI Digital Collections.

The goal of these new requirements is to make CARLI Digital Collection more useful to the end user by providing richer, more descriptive metadata records.

1. All items must contain a Title field, and the information in the field must be descriptive of the object (i.e. not a file name or other non-descriptive text).
2. All items must contain a Type field, and, within the field, state what type of resource each item is (e.g. Text, Image, MovingImage). Institutions will be instructed to use the Dublin Core Metadata Initiative (DCMI) Type vocabulary as the controlled vocabulary for this field.

When CARLI Staff creates a CONTENTdm collection for a library, the Type field can be pre-populated with a list of appropriate terms from the DCMI Type vocabulary. The libraries can simply select the appropriate term in each record they create. If records are imported in bulk, they can refer to the DCMI Type vocabulary list when creating the metadata.

3. All text items or items with spoken elements must contain a Language field populated with the language code appropriate for the item according to ISO639-2.

Again, CARLI Staff can supply this field with a pre-populated controlled vocabulary list of language codes to make the process easier for libraries.

Updating Current Collections to the New Requirements

CARLI Staff will work with all institutions that currently have collections in CARLI Digital Collections to update their collections with the new requirements. In the case of the Type field requirement, many collections are comprised of only one type of item (i.e. all images, all

manuscripts) and the Type field can be batch-populated accordingly. For mixed-type collections, the different types will need to be identified and appropriately marked.

As most of the current collections are image-based, the Language field requirement will not affect as many collections. For those where the Language field is appropriate, CARLI Staff will work with libraries to create and populate the field.

It is estimated that the Type and Language field additions will take no longer than three months to complete.

A few collections in CARLI Digital Collections do not have descriptive titles for some or all of their items. CARLI Staff has identified these collections and will work with the primary CONTENTdm contact at each of the institutions to determine what course of action is most appropriate for the collection. In several cases, there is another field provided in the metadata record that contains information that could be used in the Title field, and CARLI Staff will discuss this option with the institutions. If an alternative field cannot be used for the Title, then the library will be charged with creating descriptive titles for each of their items.

Since the process of creating new titles could be very time-consuming for a library's staff, DCUG proposes that the deadline to complete this change is a MINIMUM of one year from the time the library is first contacted about it.