GUIDELINES FOR THE CREATION OF DIGITAL COLLECTIONS

Best Practices for Descriptive Metadata

This document sets forth guidelines for creating descriptive metadata for items in CARLI Digital Collections (CDC) including minimum standards and required metadata fields. It was created by the CARLI Digital Collections Users Group (DCUG), standards subcommittee.

For questions about this document, please contact CARLI at support@carli.illinois.edu .

Definition of Descriptive Metadata

Descriptive, or bibliographic, metadata describe the content of an object. Other types of metadata provide information about the structure and arrangement of the object (structural metadata) or its lifecycle as a digital object (administrative or technical metadata). Structural and administrative metadata are not the subject of this document.

Dublin Core

Libraries managing digital collections in the CDC environment may use any descriptive metadata scheme to describe the objects in their collections, but Dublin Core is strongly recommended as a minimum standard. This document provides best practices for the use of the Dublin Core metadata scheme only; other metadata schemes, such as VRA Core, MARC, METS, etc., are not detailed here.

The chart on the following pages provides a basic overview of version 1.1 of the Dublin Core element set and indicates which elements are recommended or required for digital collections to be included in CDC.

All metadata fields in CONTENTdm, regardless of which scheme is used, have the option of being mapped to Dublin Core via the Web Administration module. Mapping fields to Dublin Core facilitates cross-collection searching within CDC and the sharing of item-level metadata outside of the CDC environment. Not only do many of the advanced cross-collection searching options in CONTENTdm require the use of Dublin Core, but the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) relies on unqualified[1] Dublin Core for metadata exchange between organizations. OAI service providers are currently harvesting CARLI Digital Collections, so to make your collections searchable through these metadata aggregators,

---

[1] The Dublin Core standard includes optional element refinements, or qualifiers, that provide additional specificity. For example, the Dublin Core element "date" has the qualifiers "created" and "dateAccepted." Since protocols like OAI-PMH only support unqualified Dublin Core, the record must still be useful if the refinements are removed. Dublin Core refinements are discussed at length on the Dublin Core Metadata Initiative site and in the Collaborative Digitization Program Dublin Core Best Practices document (see additional information section).

metadata fields must be unqualified Dublin Core, or, if you are using a different metadata scheme, your fields must be mapped to unqualified Dublin Core.

Required Fields

CARLI Digital Collections uses CONTENTdm as its digital asset management system.   Several of the metadata fields outlined in the chart are listed as "required." Please note that all required fields apply only to institutions using the CDC implementation of CONTENTdm on the CARLI servers. Institutions not using CARLI's CONTENTdm may find these documents to be useful, but are not required to follow the recommendations in their own collections.

For additional information about metadata and recommended best practices:

Dublin Core Metadata Element Set, Version 1.1
<http://dublincore.org/documents/dces/>

Collaborative Digitization Program: *Dublin Core Metadata Best Practices*:
<http://www.cdpheritage.org/cdp/documents/cdpdcmbp.pdf>

Introduction to Metadata, from the Getty Institute
<http://www.getty.edu/research/conducting_research/standards/intrometadata/>

Metadata Guidelines for Collections using CONTENTdm at the University of Washington
< http://www.lib.washington.edu/msd/mig/advice/default.html>

Digital Libraries: Metadata Resources, from the International Federation of Library Associations and Institutions (ILFA)
<http://www.ifla.org/II/metadata.htm>

CARLI Dublin Core implementation recommendations				
Term or element	Notes/description/details	CARLI status (required, recommended, optional)	Common vocabularies or encoding schemes	Additional explanation for CARLI's CONTENTdm environment, examples, and similar to which MARC field information.
Title	A name given to the resource.	Required	N/A	Title should be descriptive of the object, should not be a file name or other non-descriptive text. Similar to 245 field in MARC.
Creator	A person or entity primarily responsible for making the content of the resource.	Recommended	Library of Congress Name Authority File, ULAN	Similar to MARC field 1xx (100 or 110). Use lastname, firstname format. Example: Roddy, Alison
Subject	The topic of the content of the resource.	Recommended	Library of Congress Subject Headings, Art and Architecture Thesaurus, Thesaurus for Graphic Materials, MeSH, Rare Books and Manuscripts	MARC fields 6xx. Example: Opera --17th century --Italy --Venice. The CONTENTdm default vocabulary is TGM-I, but LCSH and MeSH can be obtained through CARLI. Users may also define other vocabularies, including locally-generated ones. Contact CARLI for details.
Description	An account of the content of the resource. May include table of contents or abstract.	Recommended	N/A	Similar to MARC fields 300 and 5xx.
Publisher	An entity responsible for making the resource available.	Optional	N/A	Similar to MARC field 260
Contributor	An entity responsible for making additional contributions to the content of the resource.	Optional	Library of Congress Name Authority File, ULAN	The Contributor will generally not be the primary author or creator. Similar to MARC field 7xx
Date	A date associated with an event in the life cycle of the resource. Best practice is to input the date in the format YYYY-MM-DD according to the W3C-DTF scheme.	Recommended	W3C-DTF	It is possible to create a local field for date in CONTENTdm and map it to Date. This is helpful if there is more than one date, for example an original publication date and a digital conversion date. Similar to the MARC field 260, fixed field 008/07-15.
Type	The nature or genre of the content of the resource.	Required	DCMI Type vocabulary	Image, MovingImage, Text, etc. See: http://dublincore.org/documents/dcmi-type-vocabulary/
Format	The physical or digital manifestation of the resource.	Recommended	Internet Media Type (IMT)	jpeg, jp2, tiff
Identifier	An unambiguous reference to the resource within a given context.	Recommended	URI	Examples of identifiers that could be used for this element: call number, bar code, file name of image, other unique numbering scheme.

CARLI Dublin Core implementation recommendations				
Term or element	Notes/description/details	CARLI status (required, recommended, optional)	Common vocabularies or encoding schemes	Additional explanation for CARLI's CONTENTdm environment, examples, and similar to which MARC field information.
Source	A reference to a resource from which the present resource is derived. Source is not used to describe the nature of the relationship (see the Relation element), but to provide a pointer to the resource itself. Best practice is to use a string or a number from a formal identification system. If the digital object is a book, for example, the Source field should contain the ISBN.	Optional	URI	Examples of source information that could be used for this element: ISBN, original URL, accession number
Language	A language of the intellectual content of the resource.	Required if a text item or an item with a spoken element (audio, video). Recommended for other types of materials.	ISO639-2	Like MARC fields 008/35-37 and 041. See: http://www.loc.gov/standards/iso639-2/
Relation	A reference to a related resource. Like the Source field, the Relation field is often a number or other type of formal reference (URI, ISBN, etc.), though it can also be a title. Relation is used to describe things such as: the collection in which the object belongs, a different version of the object, something the object is based on, etc.	Required for collection name. Optional for other information.	N/A	CARLI requires a Collection field and requires that the Collection field be mapped to Relation. Include the entire name of the collection in this field for every image. Example: Digital Image Collection (University of CARLI). The relation field may be repeated for additional information other than collection information.
Coverage	The extent or scope of the content of the resource. Usually refers to geographic coverage or a time period (temporal).	Optional	Thesaurus of Geographic Names (TGN), W3C Date and Time Formamt (W3C-DTF)	Can be indexed as a date if proper form is used. For TGN, see: http://www.getty.edu/research/conducting_research/vocabularies/tgn/. For the W3C-DTF, see: http://www.w3.org/TR/NOTE-datetime
Rights	Information about rights held in and over the resource. This should be a short and simple statement defining terms of use for the digitized resource.	Required	N/A	CARLI requires that rights information be provided for all digital content; contact information for digital collections is also required.