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 originally created by the CARLI Digital Collections Users' Group (DCUG) and is maintained and updated by the CARLI Created Content Committee (CCC).

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, MODS, etc., are not detailed here.

The list of metadata elements at the end of this document provides a basic overview of version 1.1 of the Dublin Core Metadata Element Set and indicates which elements are required or recommended for digital collections to be included in CDC.

Creating Shareable Metadata

It is important to keep in mind that the metadata you create for objects in your collection is used by the CDC for cross-collection searching and may also be shared with Open Archives Initiative[1] data aggregators to promote further discoverability outside of the CDC environment. Shareable metadata is metadata that is interoperable, meaning that the metadata used in one single stand-alone collection is also meaningful when combined in an aggregated database with metadata from items of other collections. The Digital Library Federation defines several characteristics[2] of good quality shareable metadata:

- **Proper context**--Metadata can stand on its own to give the user context for the item and does not rely on outside resources or description for understanding the item.

---

[1] http://www.openarchives.org/
[2] Introduction to Best Practices for Shareable Metadata
http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/IntroductionMetadataContent

- **Content coherence**--The record needs to be able to stand on its own in a shared environment, and should not contain significant information that only makes sense in the local environment.
- **Standard vocabularies**--Standard vocabularies facilitate cross-collection searching.
- **Consistency**--When metadata elements (their use and syntax) and the use of controlled vocabularies are inconsistently applied, aggregator databases may have difficulty processing the inconsistencies.

Use of Shareable Metadata in the CDC Environment

All metadata fields in CONTENTdm, regardless of which scheme is used, have the option of being mapped to Dublin Core fields in Web Administration. 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 mapping, but the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) also relies heavily on the Dublin Core Metadata Element Set (also known as unqualified or simple metadata) for metadata exchange between organizations. OAI service providers are currently harvesting the Dublin Core-mapped metadata from CARLI Digital Collections, so to facilitate searchable collections through these metadata aggregators, it is recommended that metadata fields for CDC collections are mapped to Dublin Core in Web Administration as appropriate.

For Additional Information about Metadata and Recommended Best Practices:

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

DCMI Metadata Terms (further information on all DCMI terms)
http://dublincore.org/documents/dcmi-terms/

Introduction to Metadata, from the Getty Institute
http://www.getty.edu/research/publications/electronic_publications/intrometadata/index.html

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

Best Practices for CONTENTdm and Other OAI-PMH Compliant Repositories: Creating Shareable Metadata
http://www.oclc.org/content/dam/training/CONTENTdm/pdf/Tutorials/Metadata/Best%20Practices%20for%20CONTENTdm%20and%20other%20OAI-PMH%20compliant%20repositories%20creating%20sharable%20metadata.pdf

Best Practices for Shareable Metadata (DLF)
http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/ShareableMetadataPublic

CARLI Dublin Core Metadata Element Set Implementation Recommendations

CARLI Digital Collections uses CONTENTdm as its digital asset management system. Five metadata fields outlined below are "required" in CARLI Digital Collections. For specifications of the CONTENTdm field properties of these required fields, see "Required Metadata Fields for CARLI Digital Collections" (https://www.carli.illinois.edu/products-services/contentdm/mdatafields). Please note that all required fields apply only to institutions using the CARLI implementation of CONTENTdm. 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. Examples in this document follow best practices; individual collections in CARLI Digital Collections may follow defined local practices instead.

Please note that additional requirements may apply for ingest into the DPLA Illinois Digital Heritage Hub (see https://www.carli.illinois.edu/products-services/contentdm/dpla).

Contributor

CARLI status	Optional
Notes/description/details	An entity responsible for making additional contributions to the content of the resource.
Common vocabularies or encoding schemes	Library of Congress Name Authority File (LCNAF) Union List of Artists Names (ULAN)
Additional information for CARLI's CONTENTdm environment	The Contributor will generally not be the primary author or creator. Use lastname, firstname format for personal names.
Relationship to MARC	Similar to 700 or 710 fields.
Usage examples	Reuss, Martin Big Idea Productions United States. Agricultural Marketing Service.

Coverage

CARLI status	Optional
Notes/description/details	The extent or scope of the content of the resource. Usually refers to geographic coverage (spatial) or a time period (temporal).
Common vocabularies or encoding schemes	Thesaurus of Geographic Names (TGN) (http://www.getty.edu/research/tools/vocabularies/tgn/) Library of Congress Name Authority File (LCNAF) W3C Date and Time Format (http://www.w3.org/TR/NOTE-datetime)
Additional information for CARLI's CONTENTdm environment	Can be indexed as a date if proper form is used (see: http://www.w3.org/TR/NOTE-datetime).
Relationship to MARC	Similar to 043, 045, some specialized 5XX fields, 651, or 650 subfields y and z.
Usage examples	Illinois (state) Chicago (Ill.) 1997-07-16

Creator

CARLI status	Recommended
Notes/description/details	A person or entity primarily responsible for making the content of the resource.
Common vocabularies or encoding schemes	Library of Congress Name Authority File (LCNAF) Union List of Artists Names (ULAN)
Additional information for CARLI's CONTENTdm environment	Use lastname, firstname format for personal names.
Relationship to MARC	Similar to 100 or 110 fields.
Usage examples	Wilson, Gilbert Edwin Chicago Park District Traffic Section

Date

CARLI status	Recommended
Notes/description/details	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.
Common vocabularies or encoding schemes	W3C Date and Time Format (W3C-DTF) http://www.w3.org/TR/NOTE-datetime
Additional information for CARLI's CONTENTdm environment	It is possible to create more than one local field for date in CONTENTdm. This is helpful if there are multiple dates, for example an original publication date and a digital conversion date.
Relationship to MARC	Like the 260 subfield c, or fixed field 008/07-15.
Usage examples	1874-09-28 1967-08 1911

Description

CARLI status	Recommended
Notes/description/details	An account of the content of the resource. May include table of contents or abstract.
Common vocabularies or encoding schemes	N/A
Additional information for CARLI's CONTENTdm environment	N/A
Relationship to MARC	Similar to 5xx fields.
Usage examples	Group of three young women - two seated at a typewriter. Streetscape view of 541 North Fairbanks Court (formerly Time-Life Building) looking to the east along Grand Avenue from Michigan Avenue. The building is clad in Cor-Ten steel; gold-tinted glass was used for the windows. The building was renovated in 1989 by Perkins & Will. View from a height, looking north along Michigan Ave. at night. The Allerton Hotel, the Palmolive Building and the Hotel St. Clair can be identified.

Format

CARLI status	Recommended
Notes/description/details	The physical or digital manifestation of the resource. May include the media-type, dimensions of the resource (such as size and duration), or used to determine software/hardware/equipment needed to operate the resource.
Common vocabularies or encoding schemes	Internet Media Type (IMT) http://www.iana.org/assignments/media-types/index.xhtml
Additional information for CARLI's CONTENTdm environment	N/A
Relationship to MARC	Similar to some 3xx fields.
Usage examples	image/jpeg application/pdf 300 dpi; b&w; 4 x 4 in.; 344.7 KB

Identifier

CARLI status	Recommended
Notes/description/details	An unambiguous reference to the resource within a given context.
Common vocabularies or encoding schemes	URI
Additional information for CARLI's CONTENTdm environment	Identifiers that could be used for this element include, but are not limited to, call number, bar code, file name of image, or other unique numbering scheme.
Relationship to MARC	N/A
Usage examples	001_cumae_acropolis_euboean_shore.jpg doi:10.1088/1475-7516/2009/02/007 ULUA.008.007

Language

CARLI status	**Required if applicable**. (see: https://www.carli.illinois.edu/ products-services/contentdm/mdatafields)
Notes/description/details	A language of the intellectual content of the resource. Language field is required for text and spoken word items and recommended for images containing text.
Common vocabularies or encoding schemes	ISO639-2 http://www.loc.gov/standards/iso639-2/
Additional information for CARLI's CONTENTdm environment	For any text item or item with a spoken element, such as audio or video, language information must be included in the metadata.
Relationship to MARC	Similar to 008/35-37 and 041 fields.
Usage examples	eng spa

Publisher

CARLI status	Optional
Notes/description/details	An entity responsible for making the resource available.
Common vocabularies or encoding schemes	Library of Congress Name Authority File (LCNAF)
Additional information for CARLI's CONTENTdm environment	N/A
Relationship to MARC	Similar to 260 subfield b.
Usage examples	Leo Feist, Inc. R. Knaplock [etc.] U.S. Dept. of Agriculture Education Coordinating Council University of Chicago

Relation

CARLI status	**Required for collection name**. Optional for other information. (see: https://www.carli.illinois.edu/products-services/contentdm/mdatafields)
Notes/description/details	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.
Common vocabularies or encoding schemes	N/A
Additional information for CARLI's CONTENTdm environment	CARLI requires a field named "Collection" and requires that the Collection field be mapped to Relation - Is Part Of. Include the entire name of the collection in this field for every item in the collection. The relation field may be repeated for additional information other than collection information.
Relationship to MARC	Can be similar to 77x and 78x fields.
Usage examples for Collection field	Great Lakes Digital Collection (Newberry Library) Image of Research (University of Illinois at Chicago)

Rights

CARLI status	**Required** (see: https://www.carli.illinois.edu/products-services/contentdm/mdatafields)
Notes/description/details	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.
Common vocabularies or encoding schemes	N/A
Additional information for CARLI's CONTENTdm environment	CARLI requires that rights information be provided for all digital content; contact information--either email or phone number--for digital collections is also required.
Relationship to MARC	Similar to 506 or 540 fields.
Usage examples	Copy permitted for study or investigation purposes, citing the source Illinois Mathematics and Science Academy. Any other use requires permission. For more information please contact xxx@imsa.edu. All rights held by William Rainey Harper College Archives. For permission to reproduce, distribute, or otherwise use this image, please contact Firstname Lastname at xxx@harpercollege.edu.

Source

CARLI status	Optional
Notes/description/details	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.
Common vocabularies or encoding schemes	URI
Additional information for CARLI's CONTENTdm environment	Source information that could be used for this element includes, but is not limited to, ISBN, original URL, and accession number.
Relationship to MARC	Similar to some 0xx fields or 856 (with second indicator 0).
Usage examples	978-0399254123 http://yourlibrary.youruniversity.edu/images/buildings/1888-2.jpg RC607.A26W574 1996

Subject

CARLI status	Recommended
Notes/description/details	The topic of the content of the resource.
Common vocabularies or encoding schemes	Library of Congress Subject Headings (LCSH) Art and Architecture Thesaurus (AAT) Thesaurus for Graphic Materials (TGM) Medical Subject Headings (MeSH) RBMS Controlled Vocabularies (Rare Books and Manuscripts) Library of Congress Name Authority File (LCNAF)
Additional information for CARLI's CONTENTdm environment	CONTENTdm offers several vocabularies, such as TGM, MeSH, AAT and more. Other vocabularies, including custom institution-generated term lists, can be used. Vocabularies may also be shared between collections.
Relationship to MARC	Similar to 6xx fields.
Usage examples	Chautauqua (N.Y.) -- Buildings, structures, etc.; Chautauqua (N.Y.) --History. cameo prints Urban beautification Public Health Nursing Festival books Coard, Frederick McDermott, 1893-

Title

CARLI status	**Required** (see: https://www.carli.illinois.edu/products-services/contentdm/mdatafields)
Notes/description/details	A name given to the resource.
Common vocabularies or encoding schemes	N/A
Additional information for CARLI's CONTENTdm environment	Title should be descriptive of the object. It should not be a file name or other non-descriptive text. CARLI requires the title to be descriptive. Collections using file names, accession numbers, etc. will not be published in the CDC.
Relationship to MARC	Similar to 245 field.
Usage examples	Baker Dormitory, Evanston IL

Acupuncture treatments

Canal Wall 1930 |

Type

CARLI status	**Required** (see: https://www.carli.illinois.edu/products-services/contentdm/mdatafields)
Notes/description/details	The nature or genre of the content of the resource.
Common vocabularies or encoding schemes	DCMI Type vocabulary http://dublincore.org/documents/dcmi-type-vocabulary/
Additional information for CARLI's CONTENTdm environment	The DCMI Type vocabulary is automatically set as the controlled vocabulary for the Type field for all new collections.
Relationship to MARC	Similar to 655 field.
Usage examples	Text StillImage MovingImage