

GUIDELINES FOR THE CREATION OF DIGITAL COLLECTIONS

Best Practices for Shareable Metadata

This document sets forth best practices for creating shareable metadata for digital collections, with the goal of enabling distributed content to be used more effectively by aggregators and end users. These recommendations are aimed at a broad audience that includes users of the CARLI Digital Collections (CDC) platform, CONTENTdm, and other CARLI member libraries who may be using different platforms. They are not intended to dictate or replace local policy and apply to shareable metadata only. CONTENTdm users are encouraged to also refer to [Best Practices for CARLI Digital Collections Descriptive Metadata](#) for input guidelines.

This document was created and is maintained by the CARLI Created Content Committee (CCC).

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

Definition of Shareable Metadata

Shareable metadata is interoperable, meaning that the metadata used in one collection is also meaningful when combined with metadata from other collections. When shareable metadata is aggregated, end users can search across multiple digital collections at the same time. Creating shareable metadata enables you to expose your items to a much wider audience and can help end users to draw meaningful connections between your items and items in other collections.

Shareable metadata is best conceived of as a particular view of a resource and not necessarily the only view. Even if your metadata is complete, accurate, and coherent locally, that may not be the case when it is aggregated with metadata from other providers. In fact, good shareable metadata may not be appropriate for a local environment. It is often the case that shareable metadata is derived from a “master” local record.

Creating Shareable Metadata

The Digital Library Federation defines the following characteristics of shareable metadata:

- **Proper context:** Shareable metadata stands on its own to give the user context for the item and does not rely on outside resources or descriptions. It should completely describe the item, even if isolated from the collection.
- **Content coherence:** Shareable metadata should be able to stand on its own in a shared environment, and should not contain information that only makes sense in the local environment (e.g. local jargon, identifiers, or codes).



- **Standard vocabularies:** Shareable metadata uses standard controlled vocabularies to facilitate cross-collection searching.
- **Consistency:** Shareable metadata uses elements, controlled vocabularies, and encoding schemes consistently across a collection.
- **Technical Conformance:** Shareable metadata should conform to the specified schema or ontology and should be encoded correctly.

In addition to these general recommendations, it is also important to consider how your shareable metadata will actually be used outside of your local environment. Each aggregator or service will often have their own guidelines and expectations.

Digital Public Library of America (DPLA)

The Digital Public Library of America (DPLA)¹ aggregates metadata from archives, libraries, and museum from across the United States. In January 2015, they added the Illinois Digital Heritage Hub (IDHH) to their organization, a partnership between the Illinois State Library (ISL), Chicago Public Library (CPL), CARLI, and the University of Illinois at Urbana-Champaign (UIUC).

The DPLA's Metadata Application Profile (DPLA-MAP)² specifies how metadata should be structured and validated in DPLA. It primarily uses elements from Dublin Core³ and the Resource Description Framework (RDF)⁴, as well as a small number of local properties. The Metadata Best Practices for the Illinois Digital Heritage Hubs⁵, developed by UIUC, provides input guidelines, mappings, and best practices for institutions that contribute their metadata to IDHH. This document also includes equivalent fields from Dublin Core, MODS, and MARC.

OAI Service Providers

Metadata may be shared with Open Archives Initiative (OAI)⁶ data aggregators, also known as service providers. OCLC's OAIster⁷ is one of the largest, but the Open Archives Initiative maintains a list of other major providers.⁸

¹ <https://dp.la/>

² <https://pro.dp.la/hubs/metadata-application-profile>

³ <http://dublincore.org/documents/dcmi-terms/>

⁴ <https://www.w3.org/RDF/>

⁵ <http://tinyurl.com/zq3e4tp>

⁶ <https://www.openarchives.org/>

⁷ <http://oaister.worldcat.org/>

⁸ <https://www.openarchives.org/service/listproviders.html>



The OAI Protocol for Metadata Harvesting defines the mechanism for data providers to expose their metadata for harvest. This protocol has relatively few requirements. Data providers must supply a valid response to OAI PMH verbs, use oai_dc as a metadata format, create unique and persistent identifiers, and include a datestamp for each item. The use of multiple metadata formats is not only allowed, but strongly encouraged. Unqualified Dublin Core (oai_dc) is simply required as an absolute minimum for interoperability. Suggested formats include CDWA Lite, EAD, ETD-MS, MARCXML, MODS, Qualified Dublin Core, and VRA Core. Although standard metadata formats should always be used, the most appropriate format will often depend on the community that is using the metadata. The format that provides the richest, most robust description should always be preferred. Some service providers will often outline their requirements or recommendations. For example, see the Sheet Music Consortium's Metadata Guidelines.⁹

Additional recommendations for classes and data elements can be found in the DLF's Best Practices for Shareable Metadata.¹⁰

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 be mapped to Dublin Core in Web Administration as appropriate.

CARLI Digital Collections 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. OCLC's Best Practices for CONTENTdm and Other OAI-PMH Compliant Repositories: Creating Shareable Metadata¹¹ contains some additional recommendations. CARLI Digital Collections users are also strongly encouraged to refer to the [Best Practices for CARLI Digital Collections Descriptive Metadata](#), which specifies recommended controlled vocabularies and encoding schemes.

⁹ http://digital2.library.ucla.edu/sheetmusic/aboutProject.html#Metadata_Guidelines

¹⁰ http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/ShareableMetadataPublic#Best_Practices_for_Shareable_Metadata

¹¹ <https://www.oclc.org/content/dam/support/wcdigitalcollectiongateway/MetadataBestPractices.pdf>