

CLIR Hidden Collections: Pioneers of Wilderness Conservation
Reference Guide for Digital Surrogate Metadata

I. [Introduction](#)

- a. [Definitions](#)

II. [Fields](#)

- a. [Collection ID](#)
- b. [Contact Information](#)
- c. [Creator](#)
- d. [Date of Digital](#)
- e. [Date of Original](#)
- f. [Description](#)
- g. [Digital Collection](#)
- h. [File Name](#)
- i. [Format of Digital](#)
- j. [Hidden Date](#)
- k. [Holding Institution](#)
- l. [Identifier](#)
- m. [Library Council](#)
- n. [Link to Object](#)
- o. [Local Location](#)
- p. [Location](#)
- q. [NYH Topic](#)
- r. [Physical Description](#)
- s. [Physical Format](#)
- t. [Publisher of Digital](#)
- u. [Rights](#)
- v. [Subjects](#)
- w. [Technical Data](#)
- x. [Title](#)
- y. [Type](#)

III. [Resources](#)

I. Introduction

The metadata describing a digital item is one of its most important components. Without accurate and appropriate metadata even the most compelling item will go undiscovered. The metadata schema developed for the digital surrogates of the CLIR Hidden Collections: Pioneers of Wilderness Conservation project at Union College were devised from New York Heritage's Best Practices Guide, 2014, with deviations made in order to increase discoverability of these collections' unique needs and components in DPLA and other online portals as well as to suit this collection's copyright needs.

This guide is broken into two sections. The first section is an alphabetical reference guide of the fields included in our digital surrogate metadata. Each field is given a definition, followed by comments discussing both the purpose of this field and why it was selected, the controlled vocabulary selected for this field, as well as an example of the field. In the next section one will find a list of resources used to develop the digital surrogate metadata.

Definitions:

Controlled Vocabulary: A limited set of terms and phrases used as headings in indexes and as access points in catalogs. Terms in a controlled vocabulary are selected so that only one term represents a concept, allowing all material relating to that concept to be retrieved using that term even if the term does not appear in the text. The term used for indexing purposes is the preferred term. The vocabulary may also include cross-references from non-preferred terms to preferred terms.ⁱ

Metadata: Metadata is commonly defined as "data about data." Metadata is frequently used to locate or manage information resources by abstracting or classifying those resources or by capturing information not inherent in the resource. Typically metadata is organized into distinct categories and relies on conventions to establish the values for each category. For example, administrative metadata may include the date and source of acquisition, disposal date, and disposal method. Descriptive metadata may include

information about the content and form of the materials. Preservation metadata may record activities to protect or extend the life of the resource, such as reformatting. Structural metadata may indicate the interrelationships between discrete information resources, such as page numbers.ⁱⁱ

II. Fields

Collection ID

Default: Yes.

Definition: This is an alphanumeric identifier, based on OLCL symbol, provided by CDLC to Schaffer Library designating the order in which the collection is added to NY Heritage.

Comments: This field separates designated items into digital collections.

Vocabulary: Assigned by CDLC.

Example:

VYD001

VYD is the OCLC symbol for CDLC, and 001 would be the first collection added to NYH.

Contact Information

Default: Yes.

Definition: The contact information for the holding institution of the collection.

Comments: This may include mailing address, e-mail address, phone number, or URL and this field responds to HTML.

Vocabulary: None.

Example:

Union College
 Kelly Adirondack Center
 Adirondack Research Library

 897 St. David's Lane
 Niskayuna, NY 12309
 (518) 280-5951
ADK@union.edu
 <http://libguides.union.edu/arl>

Creator

Default: No

Definition: The entity primarily responsible for making the content of the original item.

Comments: A creator can be a person, organization, service, author, artist, photographer, composer, etc. Multiple creators may be entered by separating the items with a semicolon and one space.

Vocabulary: [LCNAF](#)

Example:

Wollstonecraft, Mary, 1759-1797

General Electric; Steinmetz, Charles Proteus, 1865-1923

Date of Digital

Default: No

Definition: The date of the creation of the digital object.

Comments: This date is typically the creation of the master TIFF file.

Vocabulary: YYYY-MM-DD

Date of Original

Default: No

Definition: The date of creation of the original item.

Comments: This date is associated with the creation or date of availability of the original item.

Vocabulary: YYYY-MM-DD

Description

Default: No

Definition: An account of the content of the original item.

Comments: This field is the opportunity to include words to catch in keyword searching that do not apply to the title or that do not exist in current controlled subject headings.

This field may also include the transcript of an annotation of an item. This description should accurately define the contents while also contextualizing the item.

Vocabulary: None.

Example:

View of Fort Plain, the Erie Canal, and the Mohawk River mid 1800's. Pictured are bridges spanning the Erie Canal from present day Canal and State Streets. Canal boats are shown navigating the canal. To the far left of the picture, a covered bridge spans the Mohawk River connecting Fort Plain to Nelliston, NY.

Digital Collection

Default: Yes.

Definition: The name of the digital collection contributed to NY Heritage.

Comments: This name may be different than the name of the physical collection.

Vocabulary: None.

Example:

Ramée Architectural Plans for Union College
Hiram S. Wilson Civil War Letters

File Name

Default: No

Definition: The name of the digital file with the file extension.

Comments: This field can be automatically generated by ContentDM by using the Project Client metadata template.

Vocabulary: [Apperson and Schaefer Metadata Standards for Filenames](#)

Example:

UCAKappyB13F05_001.jpeg

Format of Digital

Default: Yes.

Definition: The digital format of the object.

Comments: The media type of the digital object followed by the file type.

Vocabulary: [MIME](#)

Example:

image/jpeg

image/tiff

Hidden Date

Default: No

Definition: The definition of the original item in ISO-8601 format.

Comments: This is typically the same as the creation date. This field is not required for upload, but it is recommended for best practice.

Vocabulary: YYYY-MM-DD

Holding Institution

Default: Yes.

Definition: The institution where the item is physically located.

Comments: Include the name of the institution as well as the organization or department. Maintain a standard for each institution to allow for better access.

Vocabulary: Local.

Example:

Union College, Kelly Adirondack Center, Adirondack Research Library

Union College, Schaffer Library, Special Collections and Archives

Identifier

Default: No

Definition: A unique alpha-numeric identifier for the item.

Comments: Each digital object must have a unique identifier. ContentDM recommends 8 or fewer characters, no special characters, and the use of a leading zero. The identifiers for the Schaefer and Apperson collections are the filename with the removal of the local location and the addition of the object's "digital surrogate" number.

Vocabulary: [Apperson and Schaefer Metadata Standards for Filenames](#)

Example:

UCAKappy114

UCAKpauls002

Library Council

Default: Yes.

Definition: The membership Library Council of the contributing institution.

Comments: Union College is a member of the Capital District Library Council. Use the full name of the council in the field.

Vocabulary: [New York Heritage Controlled Vocabulary List](#)

Example:

Capital District Library Council

Link to Object

Default: No

Definition: A link to the full-sized JPEG file created of the object.

Comments: This digital object is stored on the library web server.

Local Location

Default: No

Definition: The location of the original item of the digital object within the physical collection.

Example:

Location

Default: No

Definition: The geographic location of the content or subject of the object.

Comments: The location depicted in a photograph, the subject of a document, etc., NOT the location in which the item is held. Geonames has been selected for this project instead of LCSH because it affords the coordinates of smaller locations within the Adirondacks and thus better suits the individual needs of this project.

Vocabulary: [Geonames](#)

NYH Topic

Default: No

Definition: Broad topic categories defined by NY Heritage.

Comments: Up to three topic may be added. Separate multiple topics by a semi-colon and one space.

Vocabulary: [New York Heritage Controlled Vocabulary List](#)

Example:

Ethnic Groups

People; Recreation & Sports; Community & Events

Physical Description

Default: No

Definition: A description of the physical manifestation of the original item.

Comments: This field provides the user of the digital object with a better understanding of the physical manifestation of the digital object. This field includes the physical dimensions in both inches and centimeters, terms from a controlled vocabulary to

describe the item as appropriate, and terms describing the physical characteristics or aspects of the original item.

Vocabulary: [Thesaurus of Graphic Materials](#)

Example:

photograph; albumen; b&w; 4 x 2 1/2 in. (10 x 6 cm.)

Physical Format

Default: No

Definition: This field further specifies and defines the NY Heritage “type” field.

Comments: This field must be in the sub-category of the item’s “type” field.

Vocabulary: [New York Heritage Controlled Vocabulary List](#)

Example:

Furniture

Maps

Publisher of Digital

Default: Yes.

Definition: The institution of responsible for making the digital item available.

Comments: In this collection there is a separation between the rights and ownership of the physical object and the rights and ownership of the digital object. Thus, additional fields concerning the digital publication are included.

Vocabulary: None.

Example:

Union College, Schaffer Library

Rights

Default: Yes.

Definition: Information about the rights held in and over the item.

Comments: The rights of this collection are defined in the loan agreement.

Vocabulary: None. However, a controlled vocabulary for rights is in development by the NYPL.

Example:

Rights reserved; address reproduction inquiries to Special Collections, Schaffer Library, Union College, Schenectady, NY, 12308

Subjects

Default: No

Definition: A topic of the content of the item.

Comments: This is used to describe both what an item *is* as well as what it is *about*.

Separate multiple entries with a semi-colon and one space. Use up to five subjects, and use subjects consistently. We are using two controlled vocabularies.

AAT

Definition: AAT is a structured vocabulary, including terms, descriptions, and other information for generic concepts related to art, architecture, conservation, archaeology, and other cultural heritage.

Comments: Use to define items when LCSH is not appropriate.

LCSH

Definition: The Library of Congress Subject Headings are a massive controlled vocabulary list used by the majority of libraries.

Comments: This is the “go-to” controlled vocabulary for subjects.

Technical Data

Default: Yes.

Definition: Information about the digitization of the item.

Comments: Information about the hardware and software used to create the digital object.

Vocabulary: None.

Example:

JPEG made from archival TIFF - Epson Perfection V700 Photo scanner

Title

Default: No

Definition: The name given to the digital item.

Comments: The title should be descriptive enough to describe the item without context. It is alright to have a long title, include details. Do not use simple titles such as “group photo”. Include location and dates. This is what people will see in a search result, and it will be read out of context with only a small thumbnail of the digital object. The title must tell people *why* they want to view or use the item.

Vocabulary: None.

Example:

Men and women performing rippapping in bathing suits on the shore of Lake George in Bolton Landing, NY, 1917.

Type

Default: No

Definition: The nature or genre of the content of the original item.

Comments: More than one term may be applied, for example, if the item has both text and illustrations or photographs. Separate multiples with a semi-colon and one space.

Vocabulary: DCMI, but, see Appendix B (pg. 48) of [New York Heritage Metadata Dictionary](#) for a chart.

Example:

Text; Still Image

Physical Object

http://www2.archivists.org/sites/all/files/best-practice-guide-contentdm_0.pdf

III. **Resources**

- [Apperson and Schaefer Metadata Standards for Filenames](#)
- [Art and Architecture Thesaurus \(AAT\)](#)
- [Best Practices for CONTENTdm and other OAI-PMH compliant repositories: creating sharable metadata](#)
- [Geonames](#)
- [Library of Congress Name Authority File \(LCNAF\)](#)
- [Library of Congress Subject Headings \(LCSH\)](#)
- [Multipurpose Internet Mail Extensions \(MIME\)](#)
- [New York Heritage Metadata Dictionary and Best Practices](#)
- [Spreadsheets](#)
- [Thesaurus for Graphic Materials \(TGM\)](#)
- [Worksheets](#)

ⁱ SAA. "Controlled Vocabulary." Controlled Vocabulary. January 1, 2015. Accessed April 1, 2015.
<http://www2.archivists.org/glossary/terms/c/controlled-vocabulary>.

ⁱⁱ SAA. "Metadata." Metadata. January 1, 2015. Accessed April 1, 2015.
<http://www2.archivists.org/glossary/terms/m/metadata>.