

Discovery Metadata for Digital Materials

Table of Contents

Introduction	3
Creator	4
Title	6
Format	8
Content Type	10
Subject	12
Date Created	14
Date Created (ISO-8601)	16
Rights Management	17
Access	18
Yale Collection	19
Language	21
Yale Digital Collection	23
Original Repository	24

Introduction

This document outlines the recommended elements for digital collection records in the Yale University Library. The elements are format agnostic – they are designed to work with any metadata schema. An initial mapping to two metadata schemas is provided; the MODS metadata schema, and the Ladybird spreadsheet ingest schema. Collection owners can use either format for ingesting records into Ladybird. The elements can be mapped to any number of schemas, however, and it is anticipated that additional mappings will be created in the future.

Why do we need to create a discovery metadata element set? Metadata is critical to making Digital Collections discoverable in the Library's new digital interface. In order for users to execute a successful search, limit by facets, or sort records accurately by Date, a minimum number of descriptive metadata elements must be present in digital records. Those elements, in turn, need to have content that is consistent across collections. In the following document, guidelines are provided for content creation within each element, and pointers to appropriate authorities are given. Note that the creation of richer metadata beyond this minimum set will increase the accessibility and discoverability of your content. For guides to richer metadata for specific formats, see the [Art Images Metadata](#) guide and the [Metadata for Digitized Archival Papers Final Recommendations](#) .

Discovery Metadata Task Force (2013-2014)

Dominique Bourassa
Mary Caldera
Maureen Callahan
Robert Carlucci
Arcadia Falcone
Kate Kellet
Youn Noh
Jenn Nolte
Chuck Riley
Kalee Sprague (Chair)
Keiko Suzuki

Creator

Definition: An entity primarily responsible for the intellectual or artistic content of the resource.

Required: Yes

Repeatable: Yes - if naming more than one creator in the record, use separate Creator elements for each creator

Exception: **Creator is not required if provenance is unknown.**

Guidelines:

- Use an established or authorized form of a name from the Library of Congress (LC) authorities (<http://authorities.loc.gov/>).
- If no form of a name is found in the LC authorities, use the name established in the Union List of Artist Names (ULAN) (<http://www.getty.edu/research/tools/vocabularies/ulan/>).
- If no form of a name is found in ULAN, use a form found in the Virtual Authority File (<http://viaf.org/>)
- If no form of the name is established for a personal name, enter a Western style name in inverted form; otherwise enter a name in the order appropriate to the language or culture.
- If no form of the name is established for a corporate name, enter a corporate name in a brief, distinct and direct form.

Examples:

Personal names

Curie, Marie, 1867-1934

Picasso, Pablo, 1881-1973

Molière, 1622-1673

Hildegard, Saint, 1098-1179

Corporate names

Musée du Louvre
Smithsonian Institution
Yale School of Drama
Yale University. Library

MODS mapping

The creator element should be mapped to the MODS **<name>** **<namePart>** element. Repeat the element for multiple creators. To use the refinement Role to record the specific role of the person or organization named, consult MODS Subelement **<role>**. More detailed information about the name element is available at <http://www.loc.gov/standards/mods/userguide/name.html>.

Ladybird mapping

Name {fdid=69}

Title

Definition: The name of the resource or the name given to the resource.

Required: Yes. If title is not present, the record will not be ingested into the Hydra / Fedora repository.

Repeatable: No

Guidelines:

- Transcribe the title from the resource itself.
- If necessary to ensure appropriate indexing, exclude initial articles from the title.
- If no title is found on the resource itself, assign a title by following the rules of the standards used for the project or use a title assigned by others, as in reference sources.
- When assigning a title, make the title descriptive, distinctive, and succinct.
- If both a transliterated title and an original script title are available, use the Title field for the transliterated title, and the Title, alternate field for the original script (see Mods and Ladybird mapping information below).
- If the title is only available in a non-Roman script, transliterate the title using an established scheme such as the ALA-LC Romanization Tables at <http://www.loc.gov/catdir/cpsol/roman.html>, and include non-Roman script title as Title, alternate, if desired.

Examples:

Transcribed titles

African journey

Internationale situationniste

Love's labour's lost

Northanger Abbey

Assigned titles

[Caricature of Franz Liszt]

Embarkation for Cythera

Log cabin with clay chimney and covered porch

Scream

Transliterated titles

Kitāb Sharḥ al-Mu‘allaqāt al-sab‘

Ōei sanjūninen guchūreki.

Sefer ha-sha‘ashu‘im

T'aesang kamūng p'yōn tosōl.

MODS mapping

The title element should be mapped to the MODS <titleInfo> element. To enter more than one title if necessary for access or if in doubt about what constitutes the title, use the MODS attribute “alternative”: <titleInfo type="alternative">. More information about the titleInfo element is available at

<http://www.loc.gov/standards/mods/userguide/titleinfo.html>

Ladybird mapping

Title {fdid=70}

Ladybird mapping for Title, alternate

Title, alternate {fdid 71}

Format

Definition: The nature of the original resource.

Required: Yes

Repeatable: Yes – create a new format element for each applicable format.

Guidelines:

Record the nature of the original resource by using terms from the first columns in the following list. These terms will be translated into the terms in the second column for display to users.

The broad genre or class of the resource is recorded in the next element, Content Type.

Format Type	Display Terms	Definition
still image	Image	Resources predominantly visual in nature – for example, a photograph, a painting, a plan. For moving images, see video.
sound recording	Audio	Resources consisting predominantly of musical and/or nonmusical sounds.
sound recording-musical	Audio	Used when a resource is predominately a musical sound recording.
sound recording-nonmusical	Audio	Used when the sound recording is nonmusical in nature.
moving image	Video	Resources containing predominantly moving images – for example, a motion picture; a filmstrip; a video lecture.
Text	Text	Resources that are expressed through a form of notation intended to be read – for example, a manuscript, a letter, a book of poetry.
cartographic	Maps	Includes maps, atlases, globes, digital maps, and other cartographic items.

notated music	Notated Music	Graphic, non-realized representations of musical works, both in printed and digitized manifestations that represent the four components of musical sound: pitch, duration, timbre, and loudness.
software, multimedia	Software	Electronic resources, such as computer programs, enabling a computer to execute specific tasks.
dataset	Dataset	Resources in which Information is encoded in a defined structure such as lists and tables intended to be useful for direct machine processing – for example, a statistical dataset.
three dimensional object	Physical Objects	Three-dimensional artifacts or naturally-occurring objects – for example, a sculpture, a dress, or specimens mounted for viewing.

MODS mapping

The format element should be mapped to the MODS **<typeOfResource>** element.

Repeat the element for multiple formats. More detailed information about the typeOfResource element is available at

<http://www.loc.gov/standards/mods/userguide/typeofresource.html>

Ladybird mapping

Type of Resource {fdid=99}

Content Type

Definition: The broad class or genre of the resource, used to support faceting and collocation in the Yale digital interface.

Required: No – however, the absence of this field will exclude records from the Content Type facet.

Repeatable: Yes – if two or more categories apply, create a new Content Type element for each category.

Guidelines:

Assign the broad class or genre of the resource by using terms from the following list.

Specific genre terms from a variety of authorities can be assigned using a Genre element. Genre itself is not part of the minimum element set, but is recommended for richer description of the resource.

Content Type values:
Paintings & Drawings
Prints & Photographs
Sculptures, Models, & Architecture
Animal & Plant Specimens
Sound Recordings
Films, Video Testimonies & Lectures
Books, Journals & Pamphlets
Clothing & Accessories
Currency, Coins & Medals
Numeric Data
Fossil, Rock & Mineral Specimens
Furnishings & Decorative Arts
Archives or Manuscripts

Maps, Atlases & Globes
Music (Printed & Manuscript)
Tools, Equipment & Instruments

A mapping of detailed Genre terms to the broad categories above is available at:
<http://guides.library.yale.edu/loader.php?type=d&id=912714>

Example: an object consisting of a digitized photograph and a letter would have two associate genre elements:

Prints & Photographs

Archives or Manuscripts

Mods Mapping:

The Content Type element should be mapped to the MODS **<genre>** element with the attributes `authority="yale"` and `displayLabel="Content Type"`. For example, **<genre authority="yale" displayLabel="Content Type">** Paintings & Drawings **</genre>**. Repeat the element for multiple Content Types.

In addition to the Content Type, the `<genre>` element can also contain a richer level of detail using terms from authorities like AAT, etc. More detailed information about the Genre element is available at <http://www.loc.gov/standards/mods/userguide/genre.html>.

Ladybird Mapping:

Content Type {fdid 288}

Subject

Definition: A term or phrase describing the content found in or represented by a work

Required: Yes.

Repeatable: Yes – if two or more distinct subjects apply, create a new Subject element for each.

Exception: Subject may not be applicable to works that do not purport to be about anything, e.g., abstract art, or to works that are not expository in nature, e.g., belles lettres.

Guidelines:

Subject typically answers questions such as who, what, when, and where. Use as many Subject elements as needed to capture the scope of the work. Use Subject for terms from authorities.

Use a Local Subject element for keywords, tags, or other uncontrolled terms. Use specific subject fields for geographic subjects, names, titles, etc., if different types of subject can be differentiated; otherwise, use Subject for any subject term.

Possible values: The Library of Congress provides a list of authorities appropriate for use with subject at <http://www.loc.gov/standards/sourcelist/subject.html>.

Example:

Title: *Amish doll quilts, dolls, and other playthings*

Subject: Quilts, Amish

Subject: Doll quilts

Subject: Soft toys

MODS Mapping:

All terms entered in Subject are mapped to **<subject><topic>**. Terms entered in refinements of Subject other than topic are mapped to the specific subject fields, for example **<subject><geographic>**. More detailed information about the **<subject>** element is available at www.loc.gov/standards/mods/userguide/subject.html. For examples of mappings, see:

<http://www.loc.gov/standards/mods/userguide/subject.html#mappings>.

Ladybird mapping

Subject, topic {fdid=90}

If applicable, subjects can also be mapped to Subject, name {fdid=88}, Subject, geographic {fdid=91}, or Subject, geographic code {fdid=92}

Date Created

Definition: The date the item that is being described was created or modified.

Required: Yes.

Repeatable: Yes, but only repeat this field under the direction of a collection owner or cataloger. There will be cases where a collection date is described as a range, for example 1890-1915. But selected items within the collection may have known individual dates. In this case, a repeated date field containing the known date of the selected item, such as 1897, may be entered.

Exception: Date Created is strongly recommended, but in some cases accurate date information may not be available. Under the direction of a collection owner, date information may be determined using a range of years to encompass all or part of a collection.

Guidelines:

The date of creation is for the original item that is being described. It is not for a generic surrogate of that item (for visual material, for example, record the date of the object depicted in a generic photograph). The format of the date is flexible – it can be numeric and/or text, but should conform to established conventions for any related discipline-specific practices. The Date Created is used primarily for patron display.

Each department or unit may work in concert with a Metadata Services Librarian to normalize date of creation data within individual collections to provide reliable and consistent date information and to prepare them for conversion to index-able ISO8601 formatted dates (See next element, Date Created (ISO-8601)).

A detailed policy document describing Date Created and its companion Date Created (ISO-8601) field can be found at

<http://guides.library.yale.edu/loader.php?type=d&id=912716>.

Mods Mapping

The Date Created element maps to the MODS element <originInfo><dateCreated keyDate="yes" qualifier="approximate">. More detail about mapping to the MODS dateCreated element can be found at

<http://www.loc.gov/standards/mods/userguide/origininfo.html#datecreated>

Ladybird Mapping

Ladybird Mapping

Date, created {fdid=79}

Optional: If a record describes a physical object depicted in a photograph, the date the photograph was taken may be recorded in Date, depicted {fdid=80}. For example, if a record describes the Parthenon, built 447-438 BCE, and uses a photograph of the structure taken in 1956, enter 447-438 BCE in Date, created {fdid=79} and, optionally, 1956 in Date, depicted {fdid=80}.

Date Created (ISO-8601)

Definition: The ISO-8601 version of the Date Created.

Required: Yes, if applicable. Date Created (ISO-8601) is strongly recommended, but in some cases accurate date information may not be available. Note: if a Date Created is included in a record, an equivalent ISO-8601 version *must* be included.

Repeatable: No.

Guidelines:

The ISO-8601 date must be in the general format 'YYYY-MM-DDT00:00:00Z', which is the format of the ISO-8601 date standard supported by SOLR.

The ISO-8601 date is used for sorting results in SOLR, and must be a single date. If the Date Created element contains a range, then it is recommended that you use the begin date in the range as the ISO Date element.

Additional information about this standard can be found at <http://wiki.apache.org/solr/SolrQuerySyntax>. Refer to ISO8601 documentation for a complete description of the ISO standard. http://dotat.at/tmp/ISO_8601-2004_E.pdf

Examples:

A single year date: 2013-00-00T00:00:00Z

The date May 22, 1783: 1783-05-22T00:00:00Z

Ladybird Mapping

Date Created, ISO {fdid 280}

Rights Management

Definition: Information governing the use of materials after access has been provided, e.g., copyrights, licenses, et

Required: Yes.

Repeatable: Yes.

Guidelines: Yale University Library has adopted a generic Rights Management statement approved by the Office of General Counsel. A Rights Management element containing this statement is required (see below).

Yale Rights Statement: The use of this image may be subject to the copyright law of the United States (Title 17, United States Code) or to site license or other rights management terms and conditions. The person using the image is liable for any infringement.

Additional rights information can be included as additional, repeating Rights Management elements. These statements should be as free of legalese and technical jargon as possible.

Use the Access element for information on who has access to a resource and what they are authorized to do with it.

MODS Mapping:

Rights Management is mapped to the MODS **<accessCondition type="useAndReproduction">** element. For additional examples, see: <http://www.loc.gov/standards/mods/userguide/accesscondition.html#mappings>

Ladybird Mapping:

Restriction {fdid=103} – Yale Rights Statement

Access

Definition: Information about restrictions imposed on access.

Required: No

Repeatable: Yes.

Guidelines: Use this element where access to a resource is restricted:

- Classified documents are restricted until declassified by the government.
- Access restricted to campus users.

MODS Mapping:

Access is mapped to the MODS `<accessCondition type="restrictionOnAccess">` element. For additional examples, see:

<http://www.loc.gov/standards/mods/userguide/accesscondition.html#mappings>

Ladybird Mapping:

Access condition {fdid=102} – statement about restrictions on use

Yale Collection

Definition: The organization or unit at Yale that is responsible for providing access to the resource or that holds the source of a reproduction.

Required: Yes.

Repeatable: No.

Guidelines: Use a value from one of the following list of locations. Materials at the LSF should be listed under the owning library. If no owning library is apparent, use the location 'Yale University Library'.

If your location is not on this list, and you think it should be, contact digital_metadata@mailman.yale.edu

Yale Locations:
Bass Library
Beinecke Library
Center for Science and Social Science Information
Classics Library
Cushing/Whitney Medical Library
Divinity Library
Haas Arts Library
Film Study Center
Law Library
Lewis Walpole Library
Manuscripts and Archives
Math Library
Music Library
Oral History of American Music
Sterling Memorial Library

Visual Resources Collection
Yale Center for British Art
Yale Peabody Museum of Natural History
Yale University Art Gallery
Yale University Library

Example:

Haas Arts Library

MODS Mapping:

Yale Location is mapped to the MODS <location><physicalLocation> element.

Ladybird Mapping:

Location, YUL {fdid=100}

Language

Definition: The system of written, spoken, or signed communication used in a resource.

Required: Yes, if applicable. The Language element is required if a resource contains any textual content.

Repeatable: Yes. If more than one language is widely used in the resource, create a new language element for each language in the resource.

Authority: ISO 639-3 Registration Authority (<http://www-01.sil.org/iso639-3/codes.asp>)

Guidelines:

- Use an established or authorized three-letter language code from the *Codes for the representation of names of languages - Part 3: Alpha-3 code for comprehensive coverage of languages*, for each language known to be substantially represented in the resource.
- If a language of the resource cannot be determined with accuracy, use the code 'und' ('undetermined').
- If there is no linguistic content in the resource being described, the code 'zzz' should be used.
- Note that the three letter language code generates a display of the full language name in both the Ladybird and Blacklight public interfaces, e.g. 'English'.

Examples:

Language Code	Resource
akk	Cuneiform tablet in Akkadian.
Fre ara	Book with sections in French and Arabic.
Kwz por	Audio recording of Kwadzi with associated field notes in Portuguese.

Tss Instructional video in Taiwan Sign Language.
Zxx Anscombe's quartet (dataset).

MODS mapping

The Language element should be mapped to the MODS element <language>
<languageTerm type="code" authority="iso639-3">. For more detail, see
<http://www.loc.gov/standards/mods/userguide/language.html>.

Ladybird mapping

Language {fdid=84}

Yale Digital Collection

Definition: Any combination of digital objects brought together as a set under the responsibility of a Yale entity.

Required: Yes.

Repeatable: Yes

Guidelines:

- Enter a name for the collection of digital materials in a brief, distinct, direct and human-readable form.
- If Yale is part of the complete name conventionally used for the collection, it may be applied, but does not need to be present as part of a value entered into this field.

Examples:

Digital Collection	Resource
Maurice Durand Hán Nôm Collection	Digitized collection of Maurice Durand's papers.
Yale Indian Papers Project	Database of a virtual collection of primary sources from Native American history.

MODS mapping

The Digital Collection element should be mapped to the MODS element `<location>` `<physicalLocation displayLabel="Yale Collection">`. For more details, see:

<http://www.loc.gov/standards/mods/userguide/location.html>

Ladybird mapping

Digital Collection {fdid=275}

Original Repository

Definition: Holding institution of the physical object or collection, if different from the Yale Collection element.

Required: No

Repeatable: Yes

Guidelines:

- Use established or authorized forms of the name of the collection from the Library of Congress (LC) authorities (<http://authorities.loc.gov/>), if present.
- If no form of an institution's name is found in the LC authorities, use the name established in the Union List of Artist Names (ULAN) (<http://www.getty.edu/research/tools/vocabularies/ulan/>).
- If no form of the holding institution's name is in ULAN, use established or authorized forms of the name from the Virtual International Authority File (<http://viaf.org>).
- If no form of the holding institution's name is established, enter a corporate name in a brief, distinct and direct form.

Examples:

Bibliothèque nationale de France

University College, London. Center for Editing Lives and Letters.

Field Museum of Natural History

Mods mapping:

The Original Repository element should be mapped to the MODS element <location><physicalLocation displayLabel="Collection of Origin">. For more detail see <http://www.loc.gov/standards/mods/userguide/location.html#holdingSimple>

Ladybird mapping:

Location, not YUL {fdid=101}
