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**.................................................................................................18
- **Yale Location**...........................................................................................................19
- **Language**..................................................................................................................21
- **Yale Digital Collection**............................................................................................23
- **Museum or 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. A sample ingest spreadsheet using Ladybird field identifiers is also appended at the end of the document. Note that the creation of richer metadata beyond this minimum set will increase the accessibility and discoverability of your content.

Discovery Metadata Task Force
Dominique Bourassa
Robert Carlucci
Youn Noh
Chuck Riley
Kalee Sprague (Chair)
Creator

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

Required: No – use if applicable

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

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](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.

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.

Titles in non-Roman scripts
応永三十二年具注暦
太上感應篇圖說

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}
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.

<table>
<thead>
<tr>
<th>Format Type</th>
<th>Display Terms</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>still image</td>
<td>Image</td>
<td>Resources predominantly visual in nature – for example, a photograph, a painting, a plan. For moving images, see video.</td>
</tr>
<tr>
<td>sound recording</td>
<td>Audio</td>
<td>Resources consisting predominantly of musical and/or nonmusical sounds.</td>
</tr>
<tr>
<td>sound recording-</td>
<td>Audio</td>
<td>Used when a resource is predominately a musical sound recording.</td>
</tr>
<tr>
<td>musical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sound recording-</td>
<td>Audio</td>
<td>Used when the sound recording is nonmusical in nature.</td>
</tr>
<tr>
<td>nonmusical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>moving image</td>
<td>Video</td>
<td>Resources containing predominantly moving images – for example, a motion picture; a filmstrip; a video lecture.</td>
</tr>
<tr>
<td>text</td>
<td>Text</td>
<td>Resources that are expressed through a form of notation intended to be read – for example, a manuscript, a letter, a book of poetry.</td>
</tr>
<tr>
<td>cartographic</td>
<td>Maps</td>
<td>Includes maps, atlases, globes, digital maps, and other cartographic items.</td>
</tr>
<tr>
<td>notated music</td>
<td>Notated Music</td>
<td>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.</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>software, multimedia</td>
<td>Software</td>
<td>Electronic resources, such as computer programs, enabling a computer to execute specific tasks.</td>
</tr>
<tr>
<td>dataset</td>
<td>Dataset</td>
<td>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.</td>
</tr>
<tr>
<td>three dimensional object</td>
<td>Physical Objects</td>
<td>Three-dimensional artifacts or naturally-occurring objects – for example, a sculpture, a dress, or specimens mounted for viewing.</td>
</tr>
</tbody>
</table>

**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](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:**

<table>
<thead>
<tr>
<th>Paintings &amp; Drawings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prints &amp; Photographs</td>
</tr>
<tr>
<td>Sculptures, Models, &amp; Architecture</td>
</tr>
<tr>
<td>Animal &amp; Plant Specimens</td>
</tr>
<tr>
<td>Sound Recordings</td>
</tr>
<tr>
<td>Films, Video Testimonies &amp; Lectures</td>
</tr>
<tr>
<td>Books, Journals &amp; Pamphlets</td>
</tr>
<tr>
<td>Clothing &amp; Accessories</td>
</tr>
<tr>
<td>Currency, Coins &amp; Medals</td>
</tr>
<tr>
<td>Numeric Data</td>
</tr>
<tr>
<td>Fossil, Rock &amp; Mineral Specimens</td>
</tr>
<tr>
<td>Furnishings &amp; Decorative Arts</td>
</tr>
<tr>
<td>Archives or Manuscripts</td>
</tr>
</tbody>
</table>
Maps, Atlases & Globes
Music (Printed & Manuscript)
Tools, Equipment & Instruments

A mapping of detailed Genre terms to the broad categories above is available at: <URL>?

Example: An object consisting of a digitized photograph and a letter would have two associate genre elements:
Prints & Photographs
Manuscripts & Archives

Mods Mapping:
The Content Type element should be mapped to the MODS <genre> element with the attributes authority="yale" and displayLabel="ContentType". For example, <genre authority="yale" displayLabel="ContentType"> 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, if applicable. 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.

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

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, if applicable. 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.

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.

Guidelines:
The date of creation is for the original item that is being described and not for a surrogate of that item (for visual material, for example, record the date of the object depicted in a generic photograph). In practice, the creation date is a statement of fact or a knowledgeable deduction describing one component of the intellectual content specific to the item that is being described. The format of the date is flexible – it can be numeric and/or text, but should conform to established conventions for the character of the item and any related discipline-specific practices. The Date Created is used primarily for patron display.

Each department or unit may work in concert with the Chief Catalog 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 [...].

**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](http://www.loc.gov/standards/mods/userguide/origininfo.html#datecreated)

**Ladybird Mapping**
The current recommendation is to map the Date Created field to Ladybird field `{fdid=80}`. (*For a generic photograph taken in 1956 of the Parthenon, record 447-438 B.C. here*).

Optional: Use Ladybird field `{fdid=79}`-to record the date of the resource (media) that was digitized, not the thing depicted (*for a generic photograph taken in 1956 of the Parthenon, record 1956 here*)
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 format for an open-ended range of dates is
[ YYYY-MM-DDT00:00:00Z to *]

The format for a closed range of dates is
[ YYYY-MM-DDT00:00:00Z] to [YYYY-MM-DDT00:00:00Z]


Examples:

A single year date:  2013-00-00T00:00:00Z  
The date May 22, 1783:  1783-05-22T00:00:00Z  
An open-ended range:  [1976-03-06T23:59:59.999Z TO *]  
A closed range:  [1995-12-31T23:59:59.999Z TO 2007-03-06T00:00:00Z]
Ladybird Mapping
Date Created, ISO {fdid 280}
Rights Management

Definition: A statement that indicates rights relating to access and use of a resource.

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).

Additional rights information intended for end users can be included as additional, repeating Rights Management elements. These statements should be as free of legalese and technical jargon as possible. Rights Management may include information on who has access to a resource and what they are authorized to do with it (a resource may be available only to the Yale community, for example).

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.

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](http://www.loc.gov/standards/mods/userguide/accesscondition.html#mappings)

Ladybird Mapping:
Access condition {fdid=103} – general counsel statement
Restriction {fdid=102} – statement about restrictions on use
Yale Location

**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 […]

<table>
<thead>
<tr>
<th>Yale Locations:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bass Library</td>
</tr>
<tr>
<td>Beinecke Library</td>
</tr>
<tr>
<td>Center for Science and Social Science Information</td>
</tr>
<tr>
<td>Classics Library</td>
</tr>
<tr>
<td>Cushing/Whitney Medical Library</td>
</tr>
<tr>
<td>Divinity Library</td>
</tr>
<tr>
<td>Haas Arts Library</td>
</tr>
<tr>
<td>Film Study Center</td>
</tr>
<tr>
<td>Law Library</td>
</tr>
<tr>
<td>Lewis Walpole Library</td>
</tr>
<tr>
<td>Manuscripts and Archives</td>
</tr>
<tr>
<td>Math Library</td>
</tr>
<tr>
<td>Music Library</td>
</tr>
<tr>
<td>Oral History of American Music</td>
</tr>
<tr>
<td>Sterling Memorial Library</td>
</tr>
<tr>
<td>Visual Resources Collection</td>
</tr>
</tbody>
</table>
Yale Center for British Art
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 'zxx' should be used.

Examples:

<table>
<thead>
<tr>
<th>Language</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>akk</td>
<td>Cuneiform tablet in Akkadian.</td>
</tr>
<tr>
<td>fre, ara</td>
<td>Book with sections in French and Arabic.</td>
</tr>
<tr>
<td>kwz, por</td>
<td>Audio recording of Kwadzi with associated field notes in Portuguese.</td>
</tr>
<tr>
<td>tss</td>
<td>Instructional video in Taiwan Sign Language.</td>
</tr>
<tr>
<td>zxx</td>
<td>Anscombe’s quartet (dataset).</td>
</tr>
</tbody>
</table>
MODS mapping
The Language element should be mapped to the MODS element <language>
<languageterm type="code" authority="iso639-3">. For more detail, see

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:

<table>
<thead>
<tr>
<th>Digital Collection</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maurice Durand Hán Nôm Collection</td>
<td>Digitized collection of Maurice Durand’s papers.</td>
</tr>
<tr>
<td>Yale Indian Papers Project</td>
<td>Database of a virtual collection of primary sources from Native American history.</td>
</tr>
</tbody>
</table>

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}
Museum or Repository

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

**Required:** Yes, if applicable

**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/](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/](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](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. Diplomatic Correspondence of Thomas Bodley

Field Museum of Natural History
Mods mapping:
The Museum or 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}`