# CDL Guidelines for Digital Objects

# (CDL GDO)

**Maintained by the California Digital Library**

**Version 2.0**

**April 2009**

**Reviewed and Updated Semi-Annually**

**California Digital Library**

# Table of Contents

# 1. Introduction

## 1.1. Scope

The CDL Guidelines for Digital Objects (CDL GDO, this document) provides specifications for all *new* digital objects prepared by institutions for submission to CDL for access and preservation services. They are not intended to cover all of the administrative, operational, and technical issues surrounding the creation of digital object collections.

The guidelines seek to support the following objectives:

- Ensure a basic level of uniformity in the structure and encoding of non-licensed digital content managed by the CDL
- Advance interoperability among digital content from diverse institutions
- Promote efficient ingest procedures
- Support the orderly management of digital content
- Facilitate access to digital content by users
- Minimize costs

These guidelines do not set requirements for digital materials submitted to or collected by the CDL through other means:

- Metadata exposed to CDL harvesting systems via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
- Metadata targeted by federated search systems
- Web-crawled resources

In addition, these guidelines do not address requirements for collections delivered to the CDL/University of California Office of Scholarly Communication's (OSC) eScholarship Repository or for the UC Image Service. Institutions interested in submitting content to these repositories should consult the projects' websites.

## 1.2. Service Levels

Digital materials of ever-increasing variety and complexity are seen to be worth collecting and preserving by memory organizations such as libraries, archives, and museums. Materials include objects converted into digital form from existing collections of manuscripts, maps, visual images, and sound files, as well as born-digital materials such as web sites, videos, and data sets. Submitted objects consist of metadata, a set of content files, and something called a METS digital wrapper file.

In order to create coherent and cost-effective services for such diverse collections, the CDL and other digital libraries sometimes require certain common digital object

features that offer strategic points of leverage. This is a delicate undertaking, as it tends to involve a reduction in diversity that implies a loss of information, and every imposed requirement incurs the risk of rejecting valuable materials that fail to meet it. Simply meeting requirements is often hard because funding is unavailable or the original producer of the digital objects cannot be reached.

To mitigate these difficulties, the CDL adopts "sliding scale" guidelines: the more points at which a digital object can be made to conform, the more preservation and access services can be provided for it. The CDL GDO sets forth minimum submission requirements for digital objects submitted to the CDL.

At the lower end of the scale, given no information about the structure and semantics of a set of files comprising an object, the level of preservation that we can promise is limited to bit-level preservation and identifier-based retrieval -- the bits of a digital object that you submit will be the same bits that you are able to retrieve, and the only access is by known identifier or by any internal data that happens to be discoverable and indexable (e.g., content files in the form of text). In short, no metadata is required. However, providing metadata may allow access to additional preservation services while also enhancing value for future users.

At the mid-level of the scale, given a small (or "kernel") set of structured metadata encoded in a METS wrapper, the CDL will be better equipped to manage the objects and provide preservation and kernel metadata-based retrieval services. (For more information about kernel metadata, see the Dublin Core website).

Moving up the scale, generally speaking, the more metadata encoded in a METS wrapper that you supply, the better we will be able to provide you with preservation and access services. Our systems may not be able to take advantage of every distinct metadata element that you supply, but the ability to act on any element may be developed over time as our systems evolve. The higher end of the scale includes the ability to customize the formatting and grouping of collection objects, which depends on a combination of XSL style sheets and your provision of metadata elements that our systems can recognize.

The CDL GDO specifies requirements for two primary levels of services offered by the CDL:

- **Basic Service Level**: sufficient for the ingest of digital objects into the UC Libraries Digital Preservation Repository (DPR), this level is designed to support the orderly management of objects in the DPR, hence our ability to provide at least bit-level preservation without turning away valuable materials. It currently does not support the presentation of digital assets via CDL websites. This service level does not require any metadata, but strongly encourages kernel metadata. A range of content file formats is supported at this level.

- **Enhanced Service Level**: includes the presentation of digital assets via CDL websites. It is also sufficient for increased preservation services in the DPR. This level is a detailed extension of the Basic Service Level digital object specification, and therefore prescribes for additional metadata encoding. Particular content file formats are supported at this level.

## 1.3. Terminology

For an explanation of general terms used throughout these guidelines, see the CDL Glossary. For an explanation of concepts and terms pertaining to metadata in particular, consult the RLG Cultural Materials Descriptive Metadata Guidelines**.**

## 1.4. How to Use These Guidelines

Consult the appropriate section of the guidelines, based on the level of CDL service that your institution is interested in utilizing:

- **Basic Service Level**: consult Section 2 only
- **Enhanced Service Level**: consult Section 3 only

# 2. Basic Service Level Requirements

## 2.1. METS

### METS Profiles

CDL ingests content in the form of METS (Metadata Encoding and Transmission Standard) encoded digital objects. CDL depends upon METS Profiles to successfully process submitted objects.

METS profiles describe classes of METS digital objects that share common characteristics, such as content file formats (e.g., digital images, TEI texts) or metadata encoding formats (e.g., MODS or Dublin Core). Profiles should include enough details to enable METS creators and programmers to create and process METS-encoded digital objects conforming with a particular profile. A METS profile itself is an XML document that should adhere to the METS XML Profile Schema. For information about METS profiles, see the METS website.

METS files must conform to valid METS profiles, which must be declared during pre-submission discussions with CDL staff. The METS top-level <mets> element must have a PROFILE attribute that contains a URI or other identifier for the METS profile.

### Metadata and Encoding Transmission Standard <METS> Element

The METS top-level <mets> element must have an OBJID attribute containing an ARK identifier for the digital object. For more information about ARKs, visit the Archival Resource Key (ARK) page.

If an ARK is not supplied (within objects submitted to the CDL for the Basic Service Level only), a unique local identifier must be supplied as the OBJID. Under this scenario, CDL will generate an ARK when ingesting the object, and will use this ARK as the primary identifier and consider the supplied local identifier to be the equivalent of the <metsHdr><altRecordID> element.

### Content File <fileSec> Element

The METS Content File Section <fileSec> must contain links to network-exposed (i.e., online) content files using File Location <FLocat> elements. Each <FLocat> element must contain a xlink:href attribute that identifies a link to its associated content file.

The METS file and associated content files must be well formed and uncorrupted.

### File <file> Element

To support the orderly transmission and ingest of digital objects, the CDL strongly recommends submission of checksum (MD5, SHA-1, or CRC32) and byte size values in the METS File <file> element.

### Linking from Digital Objects to External Metadata: General Use of the <mdRef> Metadata Reference Element

Although METS allows for linking to external metadata using <mdRef>, the DPR ingest process will not capture this information. If you want to preserve external metadata, link to the file in the <fileSec> using <file><FLocat>.

## 2.2. Metadata

### 2.2.1. Descriptive Metadata

The Basic Service Level does not require any metadata, but strongly encourages that you supply the following kernel metadata:

| |
|---|
| *[NOTE: See Appendix A for detailed descriptions of each element. Element names below are also linked to those descriptions]* |
| Identifier |
| Title |
| Creator **(**or Contributor or Publisher**)** |
| Date |
| Description |
| Format/Physical Description |

The descriptive metadata mappings provided in Appendix A are for MODS and qualified Dublin Core. Other descriptive metadata schemas may be used, but must be defined as part of the pre-submission negotiation and will require either A) a mapping of the metadata to Dublin Core, or B) an XSL style sheet that performs the mapping.

The following data are generated by the CDL during the DPR ingest process, and can identify and provide access to digital objects submitted with no descriptive metadata. Only the most basic and fundamental of DPR services will be available for such objects. CDL-generated data:

- Object ID
- altObject ID
- Access Group ID
- Inventory (Collection) ID

- Date Ingested

### 2.2.2. Technical Metadata

The CDL generates the technical metadata required to support the orderly management of digital objects in its repositories. Currently, the CDL utilizes the JSTOR/Harvard Object Validation Environment (JHOVE) tool to derive technical metadata for accepted content file types.

You are encouraged to submit any additional technical metadata associated with a particular digital object (such as checksum [MD5, SHA-1, or CRC32] and byte size values in the METS <file> element, or information based on NISO's Data Dictionary: Technical Metadata for Still Images), but are not required to do so. CDL preservation services will store any supplied additional metadata with the object.

Note that all supplied technical metadata should be encoded using valid XML extension schemas as specified by CDL-supported METS profiles (such as in the NISO Metadata for Images in XML Schema (MIX**)** format). If a given set of metadata does not conform to a valid XML extension schema, then you should create a schema to embed the metadata and facilitate validation of the METS file. Otherwise, the metadata should be stored independently of the METS file and referred to using the METS <mdRef> Metadata Reference from within the METS file.

## 2.3. Content Files

The following content file formats are currently supported by the DPR:

- **Images:** GIF, JPG, JPG-2000, TIFF, MrSid, PDF
- **Texts:** HTML, XML, PDF, UTF-8, ASCII
- **Audio:** AIFF, WAVE
- **Containers**: GZIP, ZIP

New or unknown file formats may be submitted to the DPR, but must be established as part of the pre-submission negotiation. In addition, DPR administrators will not necessarily guarantee that all of the DPR services will be available for unknown file formats (i.e. migration or transformation processes) and will only guarantee preservation of the original bitstream.

All content files must be online or exposed over a network for the DPR software to be able to retrieve them during the ingest process. The exception is when content files are embedded within the METS wrapper using the <FContent> File Content element.

Each content file should have a file name that is unique to your institution (i.e., not necessarily globally unique); often the unique identifier is used to name the content file itself.

*Examples:*

- cacupchc_0423.tiff
- kt2g502035_fig05.gif

# 3. Enhanced Service Level Requirements

## 3.1. METS

### METS Profiles

CDL ingests content in the form of METS (Metadata Encoding and Transmission Standard) encoded digital objects. CDL depends upon METS Profiles to successfully process submitted objects.

METS profiles describe classes of METS digital objects that share common characteristics, such as content file formats (e.g., digital images, TEI texts) or metadata encoding formats (e.g., MODS or Dublin Core). Profiles should include enough details to enable METS creators and programmers to create and process METS-encoded digital objects conforming with a particular profile. A METS profile itself is an XML document that must adhere to the METS XML Profile Schema. For information about METS profiles, see the METS website.

METS files must conform to valid METS profiles, which must be declared during pre-submission discussions with CDL staff. The METS top-level <mets> element must have a PROFILE attribute that contains a URI or other identifier for the METS profile.

### Metadata and Encoding Transmission Standard <METS> Element

The METS top-level <mets> element must have an OBJID attribute containing an ARK identifier for the digital object (see bolded example). For more information about ARKs, visit the **Archival Resource Key (ARK)** page.

*Example:*
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
xmlns:mods="http://www.loc.gov/mods/v3"
xmlns:mix="http://www.loc.gov/mix/"
xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
xmlns:xlink="http://www.w3.org/TR/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/METS/
http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-0.xsd http://www.loc.gov/mix/
http://www.loc.gov/standards/mix/mix.xsd
http://cosimo.stanford.edu/sdr/metsrights/
http://cosimo.stanford.edu/sdr/metsrights.xsd"

OBJID="**ark:/13030/kt9g50158w**" TYPE="still image" LABEL="[Pablo de la Guerra (1833-1874), son of José de la Guerra y Noriega]" PROFILE="http://www.loc.gov/mets/profiles/00000001.xml">

## Content File Section <fileSec> Element

The METS Content File Section <fileSec> element must contain links to network-exposed (i.e., online) content files using File Location <FLocat> elements (see bolded example). Each <FLocat> element must contain a XLINK:HREF attribute that identifies a link to its associated content file.

*Example:*
<mets:file ID="FID8" MIMETYPE="image/jpeg" SEQ="2" CREATED="1999-06-28T00:00:00" ADMID="ADM1A" GROUPID="GID2">
<mets:FLocat
xlink:href="**http://sunsite.berkeley.edu/moa2/images/bkm00002774a_c.jpg**"
LOCTYPE="URL" />
</mets:file>

The METS file and associated content files must be well formed and uncorrupted

## File <file> Element and Checksum Values

To support the orderly transmission and ingest of digital objects, the CDL strongly recommends submission of checksum (MD5, SHA-1, or CRC32) and byte size values in the METS File <file> element.

## File <file> Element MIMETYPE Attribute

In addition to conforming to CDL-supported METS profiles, all digital objects must explicitly state content file format MIME types (Multipurpose Internet Mail Extensions) for each <file> File Element tag in the METS document (see bolded example).

*Example:*
<mets:file ID="FID1" MIMETYPE="**image/tiff**" SEQ="1" CREATED="1999-06-17T00:00:00" ADMID="ADM1A" GROUPID="GID1">

For a list of MIME type content type and subtype values, see the MIME Media Types from the Internet Assigned Numbers Authority.

## Institution/Repository Information: Specialized Use of the <mdRef> Metadata Reference Element

In order for the CDL to uniquely identify and manage digital objects by contributing institution, the CDL strongly recommends the use of a <mdRef> element with a MDTYPE attribute set to "other" and a OTHERMDTYPE attribute set to "contributing-institution-code". Additionally, use a XLINK:HREF attribute to reference the normalized version of the MARC Organization Code for the contributing institution.

The code should be listed at the end of the following URI string: "http://id.loc.gov/organizations/" (see bolded example).

*Example:*
<mets:dmdSec>
<mets:mdRef LOCTYPE="URL" MDTYPE="**other**"
OTHERMDTYPE="**contributing-institution-code**"
xlink:href="**http://id.loc.gov/organizations/cub**" />
</mets:dmdSec>

### Linking from Digital Objects to Collection Descriptions: Specialized Use of the <mdRef> Metadata Reference Element

For guidelines on linking digital objects to associated, parent-level collection descriptions (represented either in the form of a MARC record or an EAD finding aid), see Appendix C.

## 3.2. Metadata

### 3.2.1 Using Metadata Schemas

Metadata mappings are for extant XML extension metadata schemas such as MODS and qualified Dublin Core.

Encode metadata consistently based on the specific usage guidelines established for the schema. For example, if encoding in Dublin Core, follow the Dublin Core usage guidelines for each element.

Do not include HTML markup within metadata encoding, in cases where a metadata schema does not support it.

### Granularity

Whenever possible, provide the most granular and richest metadata possible. For example, if encoding in Dublin Core, encode your metadata in qualified Dublin Core.

### Repeatability of Elements and Data Values

Elements may be used repeatedly. Note that it may be necessary to supply multiple elements for the same piece of information, e.g., a general form of the date of creation of a resource (“January 1, 1999”) in addition to an ISO8601 normalized form of that date (“1999-01-01”).

However, avoid combining different kinds of data values or repeating the same type of data values within a single element; use separate elements for each data value. For example, avoid encoding multiple subject terms (“Municipal government; City Council

members") in a single element. Instead, encode the two different terms within their own elements.

## Character Encoding

Use UTF-8 or UTF-16 standard character sets or encodings. The CDL recommends using standardized forms of names for character sets, as documented by the Internet Assigned Numbers Authority (e.g., use "UTF-8" and not "UTF8").

If using the UTF-8 character set in particular, encode directly in Unicode or use Unicode decimal or hexadecimal character references. All decimal character references should begin with an ampersand and pound sign, and end with a semicolon (use the syntax "&#D;" where D is a decimal number). All hexadecimal character references should begin with an ampersand, pound sign, and lower- or uppercase "x", and end with a semicolon (use the syntax "&#xH;" or "&#XH;" where H is a hexadecimal number); see the Unicode Code Charts for hexadecimal character reference codes.

For more detailed information about UTF-8 Unicode, see the W3C/Unicode Consortium document Unicode in XML and other Markup Languages.

> *Example using UTF-8 Unicode hexadecimal character references to encode the letter "é" in the term "émigrés":*
>
> ... The papers also document trends in high school and university education among Russian émigrés...

Characters reserved for XML markup delimiters (ampersand, left angle bracket, and right angle bracket) need to be replaced with the character entities in the following table.

| Reserved Characters | | |
|---|---|---|
| **Character** | **Character Name** | **Character Entity** |
| & | Ampersand | &amp; |
| < | Left angle bracket | < |
| > | Right angle bracket | > |
| ‘ | Single quote | ' |
| “ | Double quote | " |

## Headings, Labels, Punctuation, and Formatting

Do not include line breaks, list formatting or other any formatting controls within the body of elements. Headings and labels should not appear within the body of elements (except for certain cases; see Section 3.2.3).

Some XML extension schemas (e.g., MODS) provide label attributes on particular elements. In these cases, institutions may encode data values (e.g., text comprising concise headings or descriptions) within those label attributes as permitted by those schemas.

Note that the CDL GDO supports the creation of digital objects that are largely independent of a particular online presentation. The encoding can be manipulated and repurposed through the application of customized style sheets to meet custom display needs and formatting preferences. This includes the special formatting of text, the ordering and positioning of text, the addition of headings and labels, and punctuation.

In order to provide a consistent user experience, CDL style sheets support a standard presentation that may not accommodate local preferences. Your institution may devise and implement local style sheets for presenting customized views of its digital objects .

## 3.2.2. Descriptive Metadata

### Using Descriptive Metadata Schemas

The CDL strongly supports the assertion that Dublin Core does not provide enough encoding granularity. The CDL therefore prefers that descriptive metadata is encoded in a richer format, such as MODS. Institutions should use qualified Dublin Core only in cases where MODS is not locally supported.

### Object Description

Descriptive metadata can be used to describe different expressions of a given resource. In the case of analog objects that have been digitized, the descriptive metadata may apply to the source analog object or the digital surrogate. For example, the “creator” of a resource may apply to an illustrator of a graphic book or the name of the technician responsible for scanning an image from that book. Likewise, the “date of creation” of a resource may apply to the date of printing for a graphic book or the date of scanning an image from that book. In the case of born-digital objects, the descriptive metadata pertains to the born-digital object itself.

Some descriptive metadata schemas do not allow encoders to clearly disambiguate between uses of a given element to apply to source analog objects versus digital surrogates. Therefore, when creating descriptive metadata for an analog object that has been digitized, we suggest that you consider the following two points:

- Be consistent in your use of descriptive metadata elements: emphasize the description of *either* the source analog object *or* the digital surrogate.
- Provide descriptive metadata that supports user access to and discovery of the digital object. Information about the source analog object may be more relevant to users.

| **Descriptive Metadata Guidelines (Summary)**<br><br>**[NOTE: See Appendix A for detailed descriptions of each element. Element names below are also linked to those descriptions]** | |
|---|---|
| **Element** | **Status** |
| Identifier | Required element |
| Title | Required element |
| Creator | Required element (NOTE: if no name can be supplied, provide a name in **Contributor**, **Institution/Repository**, and/or **Publisher**) |
| Date | Required element |
| Description | Recommended element |
| Language | Recommended element |
| Subject (Name) | Recommended element |
| Subject (Title) | Recommended element |
| Subject (Place) | Recommended element |
| Subject (Topic, Function, or Occupation) | Recommended element |
| Genre | Recommended element |
| Type | Required element |
| Format/Physical Description | Recommended element |
| Related Collection/Project | Recommended element |
| Institution/Repository | Required element |
| Contributor | Recommended element |
| Publisher | Recommended element |

### 3.2.3. Rights Management Administrative Metadata

CDL's Rights Management Group (RMG) has developed a Rights Management Framework that may assist institutions contributing content to CDL preservation and access services in thinking about copyright and fair use issues for digital objects. The CDL strongly encourages contributors to provide rights information whenever possible, using one of the following methods:

- Use rights-related elements in the schema chosen for supplying descriptive metadata (e.g., <dc:rights> in Dublin Core, <accessCondition> in MODS). Elements in these schemas are repeatable, so if more than one rights-related element is used, contributors should provide clarifying information about each piece of rights information either using a label attribute (MODS) or by providing a label as part of the element's content (Dublin Core).
- Supply rights information using METSRights, an approved extension schema for METS.

| **Rights Management Administrative Metadata Guidelines (Summary)**<br><br>**[NOTE: See Appendix B for detailed descriptions of each element. Element names below are also linked to those descriptions]** | |
|---|---|
| **Element** | **Status** |
| Copyright Status | Recommended element |
| Copyright Statement | Recommended element |
| Copyright Date | Recommended element |
| Copyright Owner Name | Recommended element |
| Copyright Owner Contact Information | Recommended element |

### 3.2.4. Structural Metadata

Structural metadata must be encoded in the METS format: structural metadata is represented in the <structMap> Structural Map section of a METS document. This section defines a structure that allows users of the digital object to navigate through its hierarchical organization. Guidelines for preparing Structural Maps are documented in CDL-supported METS profiles.

### 3.2.5. Technical Metadata

The CDL generates the technical metadata required to support the orderly management of digital objects in its repositories. Currently, the CDL utilizes the JSTOR/Harvard Object Validation Environment (JHOVE) tool to derive technical metadata for accepted content file types.

You are encouraged to submit any additional technical metadata associated with a particular digital object (such as checksum [MD5, SHA-1, or CRC32] and byte size values in the METS <file> element, or information based on NISO's Data Dictionary: Technical Metadata for Still Images), but are not required to do so. CDL preservation services will store any supplied additional metadata with the object.

Note that all supplied technical metadata should be encoded using valid XML extension schemas as specified by CDL-supported METS profiles (such as in the NISO Metadata for Images in XML Schema (MIX) format). If a given set of metadata does not conform to a valid XML extension schema, then you should create a schema to embed the metadata and facilitate validation of the METS file. Otherwise, the metadata should be stored independently of the METS file and referred to using the METS <mdRef> Metadata Reference from within the METS file.

### 3.2.6. Other Metadata (Digital Provenance Administrative Metadata, Source Administrative Metadata, and Behaviors Metadata)

You may submit any additional metadata associated with a particular digital object, but are not required to do so. CDL preservation services will store any additional metadata with the object. CDL access services (OAC, Calisphere) will not necessarily display supplemental metadata to users.

Note that all supplied metadata should be encoded using valid XML extension schemas as specified by CDL-supported METS profiles. If a given set of metadata does not conform to a valid XML extension schema, then you should create a schema to embed the metadata and facilitate validation of the METS file. Otherwise, the metadata should be stored independently of the METS file and referred to using the METS <mdRef> Metadata Reference from within the METS file.

## 3.3. Content Files

The following content file types are currently supported by the CDL for the Enhanced Service Level. Consult the appropriate guidelines for preparing these content file types:

| Content File Type | Content File Guidelines |
|---|---|
| • Images | Image files should comply with the CDL Guidelines for Digital Images. |

| | |
|---|---|
| | Institutions are strongly encouraged to submit at least one copy of a digital master file for each digital object. Institutions must submit, at minimum, at least two derivative file types for each digital object:<br><br>• An access image (a service or reference image for more detailed viewing).<br>• A thumbnail image (for the fastest access during the search, browse, and retrieval process). |
| • PDF texts | All PDF file formats are supported. The CDL prefers PDF files with embedded text transcriptions.<br><br>Institutions must submit one PDF file per digital object. |
| • TEI texts | TEI text files should comply with the CDL Structured Text Working Group TEI Encoding Guidelines.<br><br>Institutions must submit one TEI file per digital object. |

Each content file should have a file name that is unique to your institution (i.e., not necessarily globally unique); often the unique identifier is used to name the content file itself.

# 4. Revision History

This is the second version of the CDL GDO. This version is based upon and supersedes the *CDL Digital Object Standard, Version 1.0* (May 2001) and the *OAC Best Practice Guidelines for Digital Objects, Version 1.1* (January 2004). These guidelines were prepared by the CDL Digital Object Working Group from the fall of 2004 through the winter of 2005.

**June 2007**

- Modified Sections 2.1 and 3.1. The METS top-level <mets> element must have an OBJID attribute containing an ARK identifier for the digital object. Additionally, the METS top-level <mets> element must have a PROFILE attribute that contains a URI or other identifier for the METS profile.

**September 2007**

- Modified subheadings and reorganized content within Sections 2.1 and 3.1: subheadings are now consistently based on METS element names.
- Added METS File <file> element specifications to Sections 2.1, 2.2.2, 3.1, and 3.2.4: Technical metadata associated with a particular digital object (such as checksum [MD5, SHA-1, or CRC32] and byte size values may be supplied in the METS <file> element, but is not required.

**April 2009**

- Included encoding examples in Section 3.1 and Appendix C.
- Added recommendation in Section 3.1 and Appendix A ("Institution/Repository" element) for encoding unique identifiers for contributing institution. Recommendation specifies use of <mdRef> with a MDTYPE attribute set to "other" and a OTHERMDTYPE attribute set to "contributing-institution-code".

# Appendix A. Descriptive Metadata Guidelines (Detailed)

The following conventions are used to express guidelines for each metadata element:

- **Definition:** A definition of the element.
- **Recommended data values:** Recommended data values for the element. May include references to appropriate content standard, authority file, thesaurus, encoding standard, etc. to guide data value entry.
- **Crosswalks:** The crosswalks provide encoding analogs between elements in Dublin Core and MODS, two schemas for descriptive metadata that are commonly used with METS.
- **Examples:** Provides examples of preferred data values within elements.

## Identifier

**Definition:**A unique identifier for the resource.

**Recommended data values:** Identify the resource by means of a unique string or number conforming to a formal or locally-derived identification system. Example formal identification systems include:

- Uniform Resource Identifier (URI), including the Uniform Resource Locator (URL)
- Digital Object Identifier (DOI)
- International Standard Book Number (ISBN)

The METS top-level <mets> element must have an OBJID attribute containing an ARK identifier for the digital object. For more information, see **Section 3.1**.

**Crosswalks:**

- **Dublin Core:**
  - <dc:identifier>
- **MODS:**
  - <mods:identifier type=“”>
  - < mods:location>
    < mods:url>
- **METS:**

- <mets:mets OBJID="">

**Examples:**

calb_p3353 *[Note: locally-derived unique identifier]*
0609609718 *[Note: ISBN]*

## Title

**Definition:** A succinct, identifying name for the resource.

**Recommended data values:** Transcribe the formal title of the resource or supply a title, if necessary, using an appropriate content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), Describing Archives: a Content Standard (DACS), or Graphic Materials (GIHC).

**Crosswalks:**

- **Dublin Core:**
  - <dc:title>
- **MODS:**
  - <mods:titleInfo>
    <mods:title>
  - <mods:titleInfo>
    <mods:title>
    <mods:subtitle>

**Examples:**

Formal titles
Two dancers on a stage / Frasher Foto *[Note: transcribed according to AACR2]*
The Rocky Mountains, emigrants crossing the plains [graphic] / F.F. Palmer, del. *[Note: transcribed according to Graphic Materials]*

Supplied titles
[Photograph of musicians performing at a cultural program] *[Note: derived according to AACR2]*
Mitchell Bonner photograph of musicians performing at a cultural program *[Note: derived according to DACS]*
[Phoenix] / Ben Shahn *[Note: derived according to Graphic Materials]*

## Creator

**Definition:** The name of the person, institution, agent, or group primarily responsible for the creation of the resource.

**Recommended data values:** The form of the name should be taken from a standard naming authority file, such as the Library of Congress Name Authority File (LCNAF) or Union List of Artists' Names (ULAN). If a name does not appear in an authority file, establish the name according to a content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), or Describing Archives: a Content Standard (DACS).

Additionally, if possible, indicate the code for a standard naming authority file from which the name is taken. Use "lcnaf" for the LCNAF or "ulan" for ULAN. For all others, use the appropriate code for the source (see the Library of Congress' Term, Name, and Title Sources Code List).

If the name is not found in a standard naming authority file, indicate the content standard by which the name is established, e.g., "aacr" for AACR2, "dacs" for DACS, and "gihc" for *Graphic Materials (GIHC)* (see the Library of Congress' Descriptive Conventions Code List). If a content standard is not used, use "local".

**Crosswalks:**

- **Dublin Core:**
  - <dc:creator>
- **MODS:**
  - <mods:name type="personal | corporate | conference" authority=""><br><mods:namePart>
  - <mods:name type="personal | corporate | conference" authority=""><br><mods:namePart><br><mods:role><br><mods:roleTerm type="text">

**Examples:**

Personal name entry
Yamada, Mitsuye *[Note: determined from local cataloging authority or LCNAF]*
Chase, Alexander W. (Alexander Wells), 1843-1888 *[Note: derived according to AACR2]*
White, Ira Johnson *[Note: determined from ULAN]*
Robinson family *[Note: derived according to DACS]*

Corporate name entry
American Philosophical Society *[Note: determined from local cataloging authority or LCNAF]*
Frasher Foto (Firm) *[Note: derived according to AACR2]*

# Date

**Definition:** A single date or inclusive dates indicating when the resource was created.

**Recommended data values:** Construct dates using an appropriate content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), Describing Archives: a Content Standard (DACS), or Graphic Materials (GIHC).

At least one form of the date should be normalized (note that the Date element is repeatable) using one of the following standards:

- Temporal Enumerated Ranges (TEMPER) standard. TEMPER is a simple date and time syntax for representing points, lists, and ranges of time stamps.The syntax is designed to be machine-parseable and human-reader-friendly, and to support simple lexical sorting algorithms. TEMPER consists of four-, eight-, and 14-digit points, point ranges, and lists of points and ranges.
- International Standard Organization (ISO) 8601 standard, using a modified version of the W3C date and time formats profile.

**Crosswalks:**

- **Dublin Core:**
  - <dc:date>
  - <dcterms:created>
- **MODS**:
  - <mods:originInfo>
    <mods:dateCreated encoding="temper | w3cdtf" qualifier="">
    *[Note: do not use <dateCaptured> when describing date of creation for a born-digital resource]*
  - <mods:originInfo>
    <mods:dateOther encoding="temper | w3cdtf" qualifier="">
  - <mods:publicationInfo>
    <mods:dateIssued encoding="temper | w3cdtf" qualifier="">

**Examples:**

TEMPER encoding

***Single dates***

1901 = 1901
January 1901 = 19010100
1901 January 3 = 19010103

***Date spans***

1900-1950 = 1900-1950
1956 January-July = 19560100-19560700
1980s = 1980-1989 *[Note: use an interval to indicate every year of the decade]*
19th century = 1801-1900

***Broken date spans***

1924, 1956-1975 = 1924, 1956-1975 *[Note: separate by a comma]*

***Open date spans***
1911- = 1911-
-1911 = -1911

***Approximate dates***
circa 1950 = 1950~

***Undated material***

undated: circa mid 20th century = 1935~-1965~ *[Note: if a resource is undated this can be stated but provide an estimate if possible; normalize as an interval, perhaps using the dates of the life of creator, etc.]*

International Standard Organization (ISO) 8601 encoding (using a modified version of the W3C date and time formats profile)

***Single dates***

1901 = 1901
January 1901 = 1901-01
1901 January 3 = 1901-01-03

***Date spans***

1900-1950 = 1900/1950
1956 January-July = 1956-01/1956-07
1980s = 1980/1989 *[Note: use an interval to indicate every year of the decade]*
19th century = 1801/1900

***Broken date spans***

1924, 1956-1975 = 1924, 1956/1975 *[Note: separate by a comma]*

***Open date spans***
1911- = 1911/9999 *[Note: use an interval and set the end date to 9999]*

***Approximate dates***
circa 1950 = 1945/1955 *[Note: normalize as an interval to express an appropriate date range]*

***Undated material***

undated: circa mid 20th century = 1935/1965 *[Note: if a resource is undated this can be stated but provide an estimate if possible; normalize as an interval, perhaps using the dates of the life of creator, etc.]*

## Description

**Definition:** A brief free-text note, abstract, table of contents listing, or descriptive statement that characterizes more fully than the title does the scope or content of the resource.

**Recommended data values:** Use when the intellectual content of the item is not sufficiently captured in the title and other descriptors. Construct a note using an appropriate content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), Describing Archives: a Content Standard (DACS), or Graphic Materials (GIHC).

**Crosswalks:**

- **Dublin Core:**
  - <dc:description>
  - <dcterms:abstract>
- **MODS:**
  - <mods:abstract>
  - <mods:tableOfContents>
  - <mods:note>
  - <mods:note type=“”> *[Note: use for scope and content notes that are equivalent to MARC 520 element]*

**Examples:**

Depicts unknown automobile driver stopping at roadside to add water to engine on all-day drive from Chico to Sacramento. Exact location unknown. Verso stamped with 596; manuscript note indicates car owned by “N.E.R.” *[Note: derived according to AACR2]*
View of the Alaskan King Ice Cream Parlor, with horse-drawn delivery wagon in foreground and City Hall in background, Eugene, OR. *[Note: derived according to DACS]*

Signed in red ink. Edition of 59; Library has 14/59. *[Note: derived according to Graphic Materials]*

## Language

**Definition:** Term that indicates the language that is an integral part of the resource, such as a caption that is part of a photograph or a title that is part of a painting.

**Recommended data values:** At least one form of the language term should be normalized in coded form using a three-letter code from the from the International Organization for Standardization (ISO) 639-2 Codes for the Representation of Names of Languages (note that the Language element is repeatable, for representing the language term in textual form).

**Crosswalks:**

- **Dublin Core:**
  - <dc:language>
- **MODS:**
  - <mods:languageTerm authority="iso639-2b" type="code">

**Examples:**

eng *[Note: use for English]*
vie *[Note: use for Vietnamese]*
ger *[Note: use for German]*

## Subject (Name)

**Definition:** Significant names (personal, corporate, family, meeting, etc.) represented in or by the resource.

**Recommended data values:** The form of the name should be taken from a standard naming authority file, such as the Library of Congress Name Authority File (LCNAF) or Union List of Artists' Names (ULAN). If a name does not appear in an authority file, establish the name according to a content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), or Describing Archives: a Content Standard (DACS).

Additionally, if possible, indicate the code for a standard naming authority file from which the name is taken. Use "lcnaf" for the LCNAF or "ulan" for ULAN. For all

others, use the appropriate code for the source (see the Library of Congress' Term, Name, and Title Sources Code List).

If the name is not found in a standard naming authority file, indicate the content standard by which the name is established, e.g., "aacr" for AACR2, "dacs" for DACS, and "gihc" for *Graphic Materials (GIHC)* (see the Library of Congress' Descriptive Conventions Code List). If a content standard is not used, use "local".

**Crosswalks:**

- **Dublin Core:**
  - <dc:subject>
- **MODS:**
  - <mods:subject authority=""><br><mods:name type="personal | corporate | conference" authority=""><br><mods:namePart>

**Examples:**

Personal name entry
Yamada, Mitsuye *[Note: determined from local cataloging authority or LCNAF]*
Chase, Alexander W. (Alexander Wells), 1843-1888 *[Note: derived according to AACR2]*
White, Ira Johnson *[Note: determined from ULAN]*
Robinson family *[Note: derived according to DACS]*

Corporate name entry
American Philosophical Society *[Note: determined from local cataloging authority or LCNAF]*
Frasher Foto (Firm) *[Note: derived according to AACR2]*

## Subject (Title)

**Definition:** Significant titles of other resources (e.g., works, expressions of those works, individual items, etc.) represented in or by the resource.

**Recommended data values:** The form of the title should be taken from a standard naming authority file, such as the Library of Congress Title Authority File (LCTAF). If a title does not appear in an authority file, establish the title according to a content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), Describing Archives: a Content Standard (DACS), or Graphic Materials (GIHC).

Additionally, if possible, indicate the code for a standard naming authority file from which the title is taken. Use "lctah" when the name is established in the LCTAF.

If the title does not appear in the authority file, indicate the content standard by which the title is established, e.g., “aacr” for AACR2, “dacs” for DACS, and “gihc” for *Graphic Materials (GIHC)* (see the Library of Congress’ Descriptive Conventions Code List). If a content standard is not used, use “local”.

**Crosswalks:**

- **Dublin Core:**
  - <dc:subject>
- **MODS:**
  - <mods:subject authority=““>
    <mods:titleInfo authority=““>
    <mods:title>

**Examples:**

Kim Hà, 1950-. Qua con bao du : hoi ky vuot bien bang duong bo. *[Note: manuscript material documenting creation of a monograph titled “Qua con bao du”; entry derived according to AACR2]*

## Subject (Place)

**Definition:** Significant names of geographic locations represented in or by the resource.

**Recommended data values:** The form of the name should be taken from a standard naming authority file, such as the Library of Congress Name Authority File (LCNAF) or Thesaurus of Geographic Names (TGN). If a name does not appear in an authority file, establish the name according to a content standard such as Anglo-American Cataloging Rules (AACR2) or Cataloging Cultural Objects (CCO).

Additionally, if possible, indicate the code for a standard naming authority file from which the name is taken. Use “lcnaf” for the LCNAF or “ulan” for ULAN. For all others, use the appropriate code for the source (see the Library of Congress’ Term, Name, and Title Sources Code List).

If the name is not found in a standard naming authority file, indicate the content standard by which the name is established, e.g., “aacr” for AACR2, “dacs” for DACS, and “gihc” for *Graphic Materials (GIHC)* (see the Library of Congress’ Descriptive Conventions Code List). If a content standard is not used, use “local”.

**Crosswalks:**

- **Dublin Core:**

- <dc:coverage>
- <dcterms:spatial>

- **MODS:**
  - <mods:subject authority=“”>
    <mods:geographic>
  - <mods:subject authority=“”>
    <mods:hierarchicalGeographic>
  - <mods:subject authority=“”>
    <mods:cartographics>

**Examples:**

Santa Cruz (Calif.) *[Note: determined from local cataloging authority or LCNAF]*
Santa Cruz *[Note: determined from TGN]*
Rancho Boca de la Playa (Calif.) *[Note: established according to AACR2]*

## Subject (Topic, Function, or Occupation)

**Definition:** Significant topics or subjects (including concepts, events, etc.), functions, or occupations represented in or by the resource.

**Recommended data values:** The form of the heading should be taken from a standard or local thesaurus, such as the Library of Congress Subject Headings (LCSH), Art and Architecture Thesaurus (AAT), or Thesaurus of Graphic Materials I (TGM I).

If a heading does not appear in a thesaurus, establish the heading according to standard thesaurus rules (such as the Library of Congress’ *Subject Cataloging Manual*, AAT rules, or TGM I rules), or local thesaurus rules.

Additionally, if possible, indicate the code for a standard naming authority file from which the heading is taken. Use “lcsh” for LCSH, “aat” for AAT, or “gmgpc” for TGM I (see the Library of Congress’ Term, Name, and Title Sources Code List).

If the heading does not appear in a standard thesaurus, indicate the thesaurus rules by which the term is established, e.g., “lcsh” for LCSH, “aat” for AAT, or “gmgpc” for TGM I (see the Library of Congress’ Term, Name, and Title Sources Code List). If standard thesaurus rules are not used, use “local”.

**Crosswalks:**

- **Dublin Core:**
  - <dc:subject>
- **MODS:**

- <mods:subject authority=“”>
  <mods:topic>
- <mods:subject authority=“”>
  <mods:occupation>

**Examples:**

Viticulture -- California -- Sonoma County *[Note: determined from LCSH]*
Surveyors--California--Orange County *[Note: determined from LCSH]*
Street railroads *[Note: determined from AAT]*
Agricultural laborers--Italian Americans--California--Salinas *[Note: determined from TGM I]*

## Genre

**Definition:** Primary genre(s) represented in or by the resource.

**Recommended data values:** The form of the heading should be taken from a standard or local thesaurus, such as the Library of Congress Subject Headings (LCSH), Art and Architecture Thesaurus (AAT), Genre Terms (RBGENR), or Thesaurus of Graphic Materials II (TGM II).

If a heading does not appear in a thesaurus, establish the heading according to standard thesaurus rules (such as the Library of Congress’ *Subject Cataloging Manual*, AAT rules, or TGM II rules), or local thesaurus rules.

Additionally, if possible, indicate the code for a standard naming authority file from which the heading is taken. Use “lcsh” for LCSH, “aat” for AAT, or “gmgpc” for TGM II (see the Library of Congress’ Term, Name, and Title Sources Code List).

If the heading does not appear in a standard thesaurus, indicate the thesaurus rules by which the term is established, e.g., “lcsh”, “aat”, or “gmgpc”. Use the appropriate code for the thesaurus by which the term is established (see the Library of Congress’ Term, Name, and Title Sources Code List). If standard thesaurus rules are not used, use “local”.

- **Crosswalks:**
- **Dublin Core:**
  - <dc:type>
- **MODS:**
  - <mods:genre authority=“ ”>

**Examples:**

Photographs *[Note: determined from LCSH]*
Photographic prints *[Note: determined from AAT]*
Photographic prints *[Note: determined from TGM II]*

## Type

**Definition:** A high-level type data value that generally characterizes the resource represented by the digital object. This high-level data value may also be repeated, or more specific genre data values may also be encoded as part of the descriptive metadata (see Genre).

**Recommended data values:** Choose data values from one of the following lists, based on the descriptive metadata scheme being utilized. Select data values from the MODS type vocabulary if in doubt:

- Dublin Core type vocabulary
- MODS type vocabulary (see values listed under <typeOfResource>)

**Crosswalks:**

- **Dublin Core:**
  - <dc:type>
- **MODS:**
  - <mods:typeOfResource>
- **METS:**
  - <mets:mets TYPE="">

**Examples:**

image *[Note: determined from Dublin Core type vocabulary]*
still image *[Note: determined MODS type vocabulary]*

## Format/Physical Description

**Definition:** The physical or digital manifestation of the resource. Typically, this may include the media-type or dimensions of the resource. Examples of dimensions include size and duration.

**Recommended data values:** Construct a statement using an appropriate content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), Describing Archives: a Content Standard (DACS), or Graphic Materials (GIHC).

**Crosswalks:**

- **Dublin Core:**
  - <dc:format>
  - <dcterms:extent>

- **MODS:**
  - <mods:physicalDescription>
    <mods:extent>
  - <mods:physicalDescription>
    <mods:form>
  - <mods:physicalDescription>
    <mods:internetMediaType>

**Examples:**

1 photographic print ; 9 x 14 cm. *[Note: derived according to AACR2]*
14 letters *[Note: derived according to DACS]*
1 leaflet : ill. ; 21.5 x 38.5 cm., folded to 21.5 x 10 cm. *[Note: derived according to Graphic Materials]*

## Related Collection/Project

**Definition:** A machine access-oriented identifier for a collection or project that the resource is a member of or related to in some manner.

**Recommended data values:** If the resource is a member of or related to a collection or project, at least one **Related Collection/Project** element must refer to a unique identifier or to the title for the collection or project. Data producers may use a unique local identifier to refer to a related collection or project, such as a URL or other unique identifier. Unique identifiers managed by the CDL are preferred. To obtain a unique CDL identifier for a particular collection or project, contact the CDL.

Alternatively, indicate the title for a collection or project.

For guidelines on linking digital objects to associated, parent-level collection descriptions (represented either in the form of a MARC record or an EAD finding aid),

see **Appendix C** . Use a METS <mdRef> element with a MDTYPE attribute set to either "MARC" or "EAD".

**Crosswalks:**

- **Dublin Core:**
  - <dc:relation>
  - <dcterms:isPartOf>
- **MODS:**
  - <mods:relatedItem>
    <mods:url>
  - <mods:relatedItem>
    <mods:identifier>

**Examples:**

CDL-supplied unique identifier
http://jarda.cdlib.org

Title
Silicon Valley History Online

## Institution/Repository

**Definition:** The name of the owning or contributing institution of the resource.

**Recommended data values:** The form of the name should be taken from a standard naming authority file, such as the Library of Congress Name Authority File (LCNAF). If the name does not appear in an authority file, establish the name according to a content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), or Describing Archives: a Content Standard (DACS**)**.

In order for the CDL to uniquely identify and manage digital objects by contributing institution, the CDL strongly recommends the use of a METS <mdRef> element with a MDTYPE attribute set to "other" and a OTHERMDTYPE attribute set to "contributing-institution-code". Additionally, use a XLINK:HREF attribute to reference the normalized version of the MARC Organization Code for the contributing institution. The code should be listed at the end of the following URI string: "http://id.loc.gov/organizations/". For more information, see **Section 3.1**.

**Crosswalks:**

- **Dublin Core:**

- <dc:publisher>

- **MODS:**
  - <mods:location>
    <mods:physicalLocation authority="">
    <mods:physicalLocation>

- **METS:**
  - <mets:mdRef LOCTYPE="URL" MDTYPE="other" OTHERMDTYPE="contributing-institution-code" xlink:href="http://id.loc.gov/organizations/" />

**Examples:**

Fowler Museum of Cultural History *[Note: determined from local cataloging authority or LCNAF]*
Orange Public Library *[Note: derived according to AACR2]*
University of California, San Francisco. Library. Archives and Special Collections *[Note: derived according to AACR2]*

## Contributor

**Definition:** The name of the person, institution, agent, or group responsible for contributing to the resource in some significant manner, such as a illustrator, designer, autographer, etc.

**Recommended data values:** The form of the name should be taken from a standard naming authority file, such as the Library of Congress Name Authority File (LCNAF) or Union List of Artists' Names (ULAN). If a name does not appear in an authority file, establish the name according to a content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), or Describing Archives: a Content Standard (DACS).

Additionally, if possible, indicate the code for a standard naming authority file from which the name is taken. Use "lcnaf" for the LCNAF or "ulan" for ULAN. For all others, use the appropriate code for the source (see the Library of Congress' Term, Name, and Title Sources Code List).

If the name is not found in a standard naming authority file, indicate the content standard by which the name is established, e.g., "aacr" for AACR2, "dacs" for DACS, and "gihc" for *Graphic Materials (GIHC)* (see the Library of Congress' Descriptive Conventions Code List). If a content standard is not used, use "local".

**Crosswalks:**

- **Dublin Core:**
  - <dc:contributor>

- **MODS:**
  - <mods:name type="personal | corporate | conference" authority=" ">
    <mods:namePart>
    <mods:role>
    <mods:roleTerm type="text">

**Examples:**

Personal name entry
Yamada, Mitsuye *[Note: determined from local cataloging authority or LCNAF]*
Chase, Alexander W. (Alexander Wells), 1843-1888 *[Note: derived according to AACR2]*
White, Ira Johnson *[Note: determined from ULAN]*
Robinson family *[Note: derived according to DACS]*

Corporate name entry
American Philosophical Society *[Note: determined from local cataloging authority or LCNAF]*
Frasher Foto (Firm) *[Note: derived according to AACR2]*

## Publisher

**Definition:** The name of the publisher of a formally published resource. This element may not be relevant for unpublished materials.

**Recommended data values:** The form of the name should be taken from a standard naming authority file, such as the Library of Congress Name Authority File (LCNAF) or Union List of Artists' Names (ULAN). If a name does not appear in an authority file, establish the name according to a content standard such as Anglo-American Cataloging Rules (AACR2), Cataloging Cultural Objects (CCO), or Describing Archives: a Content Standard (DACS).

**Crosswalks:**

- **Dublin Core:**
  - <dc:publisher>
- **MODS:**
  - <mods:originInfo>
    <mods:publisher>

**Examples:**

Simon & Schuster *[Note: determined from local cataloging authority or LCNAF]*
New Albion Records *[Note: derived according to AACR2]*

# Appendix B. Rights Management Administrative Metadata Guidelines (Detailed)

## Copyright Status

**Definition:** Indicates general type of copyright status for the resource.

**Recommended data values:** If using Dublin Core or MODS, enter one of the following data values: enter “Public domain” if in the public domain, “Copyrighted” if copyrighted, or “Unknown” of copyright status is unknown.

If using METSRights: enter “PUBLIC DOMAIN” if in the public domain, or “COPYRIGHTED” if copyrighted . If the copyright status is unknown, enter “OTHER” and add the following additional attribute: OTHERCATEGORYTYPE=“UNKNOWN”. Note that all attributes and attribute values must be in upper case when using the METSRights schema.

**Crosswalks:**

- **METSRights:**
  - <RightsDeclarationMD RIGHTSCATEGORY=“ ”>
  - <RightsDeclarationMD RIGHTSCATEGORY=“OTHER” OTHERCATEGORYTYPE=“UNKNOWN”>
- **Dublin Core:**
  - <dc:rights>
- **MODS:**
  - <accessCondition type=“useAndReproduction”>

## Copyright Statement

**Definition:** A free-text note that describes copyright restrictions pertaining to the resource.

**Recommended data values:** Usage of one of the following copyright statements is recommended, based on the data value assigned in Copyright Status:

- When the status is “unknown” :

  Some materials in these collections may be protected by the U.S. Copyright Law (Title 17, U.S.C.). In addition, the reproduction, and/or

commercial use, of some materials may be restricted by gift or purchase agreements, donor restrictions, privacy and publicity rights, licensing agreement(s), and/or trademark rights. Distribution or reproduction of materials protected by copyright beyond that allowed by fair use requires the written permission of the copyright owners. To the extent other restrictions apply, permission for distribution or reproduction from the applicable rights holder is also required. Responsibility for obtaining permissions, and for any use rests exclusively with the user.

- When the status is “public domain” :

  Material in the public domain. No restrictions on use.

- When the status is “copyrighted” :

  Transmission or reproduction of materials protected by copyright beyond that allowed by fair use requires the written permission of the copyright owners. Works not in the public domain cannot be commercially exploited without permission of the copyright owner. Responsibility for any use rests exclusively with the user.

**Crosswalks:**

- **METSRights:**
  - <Context CONTEXTCLASS=“GENERAL PUBLIC”>
    - <Constraints>
      - <ConstraintDescription>
- **Dublin Core:**
  - <dc:rights>
- **MODS:**
  - <accessCondition type=“useAndReproduction”>

## Copyright Date

**Definition:** The year the resource was copyrighted. Use only if **Copyright Status** is “Copyrighted”.

**Recommended data values:** Supply the year the resource was copyrighted, typically based on a copyright notice on the resource itself. The resource does not have to have been registered with the copyright office. Do not approximate copyright year if it does not appear on the work, or in some reliable alternate source for this information. Use the standardized form of YYYY, and do not include month or day information.

**Crosswalks:**

- **METSRights:**
  - <RightsDeclarationMD RIGHTSCATEGORY="COPYRIGHTED">
    <RightsDeclaration>
- **Dublin Core:**
  - <dcterms:dateCopyrighted>
- **MODS:**
  - <originInfo>
    <copyrightDate encoding="temper | w3cdtf" qualifier=" ">

## Copyright Owner Name

**Definition:** The name(s) of the copyright holders of the resource. Use only if **Copyright Status** is "Copyrighted".

**Recommended data values:** Specify the most common form of the name in natural or direct order.

**Crosswalks:**

- **METSRights:**
  - <RightsHolder>
    - <RightsHolderName>
- **Dublin Core:**
  - <dc:rightsHolder>
- **MODS:**
  - <accessCondition type="useAndReproduction">

## Copyright Owner Contact Information

**Definition:** Publicly accessible contact information for the copyright owner(s) of the resource. Use only if **Copyright Status** is "Copyrighted".

**Recommended data values:** Provide as much contact information as possible that can be made available to the public. Otherwise, use the phrase "Consult contributing institution" or a similar note.

**Crosswalks:**

- **METSRights:**
  - <RightsHolder>

        - <RightsHolderContact>
    - <RightsHolderContactAddress>
- **Dublin Core:**
    - <dc:rightsHolder>
- **MODS:**
    - <accessCondition type="useAndReproduction">

# Appendix C. Linking from Digital Objects to Collection Descriptions

The following guidelines apply to specialized use of the METS <mdRef> Metadata Reference element to create links from digital objects to associated, parent-level collection descriptions (represented either in the form of a MARC record or in the form of an EAD finding aid content file). Note that particular METS profiles may provide for more specific guidelines on use of the <mdRef> element for this or other purposes.

If the collection description is encoded in a MARC record, then encode the entire URL for the MARC record in a <mdRef> Metadata Reference HREF attribute of the object's METS wrapper. Include a MDTYPE attribute (see the bolded examples). Note that it may be difficult to generate a static and durable URL for particular MARC records, depending on local OPACs:

> *Example:*
> <METS:dmdSec ID='DMR1'>
> <METS:mdRef LOCTYPE='URL' MDTYPE='MARC'
> xlink:href='http://antpac.lib.uci.edu/search/tkim+ha/tkim+ha/1%2C3%2C3%2CB/frameset&FF=tkim+ha+papers+1983+1999&1%2C1%2C'/>
> </METS:dmdSec>

If the collection-level description is encoded as an EAD finding aid, then use the following procedures:

1. Obtain the Archival Resource Key (ARK) URL for the finding aid that you would like to link the object to. (For finding aids already submitted to the CDL, the finding aid ARK URL can be obtained by viewing the finding aid online in CDL access systems).

   The ARK is a machine-readable unique identifier scheme for persistent access to digital resources managed by the CDL. Because ARKs are specially constructed and globally unique identifiers, their production and management is controlled by the CDL. For more information on the use of ARKs at the CDL and obtaining new ARKs, please see the CDL ARK service description.

   In some cases, you may be creating objects that will link to a finding aid that has not yet been created or completed (and will be submitted to the CDL at a later time). To obtain a new “placeholder” ARK for the finding aid, contact the CDL or mint an ARK locally using a CDL tool. At the point that you create the finding aid, encode the ARK within the <eadid> identifier attribute (see the bolded example). Only encode the portion of the ARK beginning with “ark:/...”, and not the entire ARK URL:

*Example of EAD finding aid with ARK encoding:*
<eadid countrycode="us" identifier="**ark:/13030/kt4w10133d**" mainagencycode="CU-SC" publicid="PUBLIC "-//University of California, Santa Cruz::University Library::Special Collections//TEXT (US::CU-SC::MS 74::John Cage Mycology Collection)//EN" "ms74.sgm">ms74.xml</eadid>

2. Encode the entire ARK URL for the finding aid in a <mdRef> metadata reference HREF attribute of the object's METS wrapper. There are two possible methods for doing this. [Note: the second method will require additional and more intensive finding aid encoding].
   - Link the digital object to the beginning of the EAD finding aid. Include the ARK URL for the finding aid in the HREF attribute as well as a LABEL value (see the bolded examples).

     *Example of object with link to the beginning of an EAD finding aid:*
     <METS:dmdSec ID="DMR1">
     <METS:mdRef LOCTYPE="URL" MDTYPE="**EAD**" LABEL="**Arnold Rubin Papers**" xlink:href="**http://www.oac.cdlib.org/findaid/ark:/13030/kt7489n8zj**" />
     </METS:dmdSec>

   - Link the digital object to the place in the EAD finding aid where the child-level digital object is described. Include the ARK URL for the finding aid in the HREF attribute, MDTYPE and LABEL attributes, and a unique identifier ID value (see the bolded examples).

     *Example of object with link to specific section of an EAD finding aid:*
     <METS:dmdSec ID="DMR1">
     <METS:mdRef LOCTYPE="URL" MDTYPE="EAD" LABEL="**Arnold Rubin Papers**" xlink:href="**http://www.oac.cdlib.org/findaid/ark:/13030/kt7489n8zj**" XTPR="xpointer(id('**xyzj0098**'))"/>
     </METS:dmdSec>

The LABEL value should be the title for the collection that the object is related to (e.g., "Arnold Rubin Papers"). This should be the same title used for the collection in the associated finding aid.

The unique identifier serves as an anchor point that links the object to a specific point in the finding aid, and vice versa. Therefore, it must be present in the object in addition to the specific point in the finding aid. The value of the unique identifier must be unique for the finding aid. The unique identifier must also comply with rules for IDs specified in the W3C's XML Schema Part 2: Datatypes. Namespaces must start with a letter and can include letters, digits, periods, hyphens, and underscores. Colons are not allowed.

The same unique identifier ID value also needs to be encoded at the appropriate <archdesc> Archival Description or <c0x> Component level where the object(s) is described or implied in the finding aid. Encode the unique identifier (see bolded example) within an ID attribute for that <c0x> Component tag.

*Example of EAD encoding:*
<c02 id="**xyzj0098**" level="item">
<did>
<container type="box" label="Flat file">7</container>
<unittitle>Ushiwaka and Benkei duelling on Gojo Bridge or Gojo Bridge, an episode from the Life of Yoshitsune, Chronicles of Yoshitsune, <unitdate>1881</unitdate>.
</unittitle>
<dao role="http://oac.cdlib.org/arcrole/link/image" href="http://ark.cdlib.org/ark:/13030/kt4p3005qx/"></dao>
</did>
</c02>

3. Create or update your finding aids to link to their associated objects, following the specifications outlined in the Online Archive of California Best Practice Guidelines for Encoded Archival Description (OAC BPG EAD), Sections 4.4.5 and 4.5.

   Do not add linking information to a finding aid prior to submitting the associated objects.

   The link from a finding aid to an object or group of objects can be made at any level (i.e., collection, series, subseries, file, or item level) in the finding aid, but it should be made at the level where the object(s) is described or implied in the finding aid.

   Creating links from finding aids to objects can be done two ways:

   - Link from a specific level in the finding aid (e.g., collection, series, etc.) to a group of objects, by linking to a search query for the group of objects in CDL access systems. This is most practical when linking from the collection level or major subdivision to the entire group of objects represented by that level of description. Note that you will be creating a link to a "canned" search for a group of objects, not to an individual or specific object. For an example, see the Redwood City Public Library's Morrish Collection.
   - Link from a specific level in the finding aid (e.g., file- or item-level) to a specific, individual object. This is most practical when linking from the file or item level to an individual object.