

Plains to Peak Collective Metadata Guidelines for DPLA Participation

Draft 8/2017

Note: This is currently a working document. It is expected to evolve and change during the testing phase and as the needs of our partners are identified. Please feel free to comment and ask questions within the document or send them to Leigh directly. This document is intended for use by those that have established digital collections and by those that are just getting starting.

Introduction

These guidelines include Digital Public Library of America (DPLA) and Plains to Peaks Collective (PPC) requirements, recommendations and best practices for preparing your collections metadata for participation in the PPC Service Hub and the DPLA. These guidelines were largely informed by the [DPLA Metadata Application Profile](#) (v4.0) and other active DPLA service Hub metadata guidelines. DPLA intends metadata fields describe the **original resource** (original item), such as a photograph or letter, not the digital representation (digital scan) of that item. However fields could describe a digital item if the item being described is born-digital. These guidelines are not meant to be a comprehensive cataloging guide but rather a guide for sharing data for participation in the DPLA. Your institution may have its own institutional cataloging guidelines. It is the PPC hope that institutional guideline can work alongside of or be crosswalked to these guidelines.

The PPC does not recommend a particular metadata schema, as every institution needs are different. For each field we do offer crosswalks to MARC, MODS, and Qualified Dublin Core. If you follow a local metadata schema or one the is not listed as a crosswalk we are happy to review your metadata schema to discover compatible fields. We highly recommend that within your chosen metadata schema, you are as consistent as is possible across the records supplied to the PPC and the DPLA. The DPLA has few requirements for the metadata that is shared. However, the more fields a content provider supplies the more discoverable their items will be in the DPLA platform.

Objects not Accepted

The DPLA does not accept the following records as digital objects:

- Finding aids

- Objects that do not resolve to a digital object, such as this [example](#)
- Records for individual pages of a book or component records
- Full text transcriptions
- Secondary products of education and scholarship, such as lecture notes, presentations, and related materials that are often found in institutional repositories
- They prefer that student thesis be historical in nature
- Datasets (neither small or large, including XSL, etc.)

Terminology used in this Document

- **Metadata:**
- **Digital object:** A descriptive metadata record of a unique item such as a photograph, manuscript material, artwork, born-digital item, etc. It has to be a single descriptive metadata record but it does not need to describe a single item. A record should describe a single book, for example, rather than all its pages. A single digital object might also be for a folder of objects that are presented together as a single multi-part item, such as a file folder of archival material. What doesn't work is a link to a record that then has links to many multiple digital objects with their own descriptions, such as a finding aid for a collection.
- **Required:** This information must be provided by the institution that owns the item in order for it to be sent to DPLA
- **Required When Applicable:** If this information is relevant to the item being described, the owning institution needs to supply this information in order for it to be sent to DPLA
- **Required When Available:** If the information is available to the owning institution, it needs to supply this information in order for it to be sent to DPLA
- **Strongly Recommended:** Adding this information will provide maximum discoverability.
- **Recommended:** While not strictly required for the item to be discovered, adding this information will help searchers find your information more easily and provide potentially critical information for them to figure out if it's something that would be useful in their research.
- **Partner Supplied:** Institution that owns the item provides this information, which is taken verbatim by DPLA and used in its platform *as is*.
- **Derived from Partner-supplied data:** Institution that owns the item provides an appropriate form of this information. DPLA and/or PPC takes that information and creates a standardized version of that for its platform. For more information see the relevant section under *PPC Required and Recommended Fields* below.

Suggested Controlled Vocabularies

The PPC does not endorse one vocabulary over another but rather offers the below list as a reference. The PPC understands that many institutions have their own local controlled vocabularies. With the metadata supplied to PPC, it is important to be as consistent as possible in the use of any controlled vocabulary. Please also note that DPLA requires that the metadata (not necessarily the collection item) be licensed as [CCO](https://creativecommons.org/licenses/by/4.0/), Public domain, No rights reserved (<https://creativecommons.org/share-your-work/public-domain/cc0/>).

Abbreviation	Full Name	URL
LCNAF	Library of Congress Name Authority File	http://id.loc.gov/authorities/names.html
TGN	Getty Thesaurus of Geographic Names	http://www.getty.edu/research/tools/vocabularies/tgn/index.html
GeoNames	GeoNames Geographical Database	http://www.geonames.org/
FAST	Faceted Application of Controlled Vocabulary	http://fast.oclc.org/searchfast/
VIAF	Virtual International Authority File	https://viaf.org/
ULAN	Getty Union List of Artist Names	http://www.getty.edu/research/tools/vocabularies/ulan/
ISO 639-2	Codes for the Representation of Names of Languages	https://www.loc.gov/standards/iso639-2/php/code_list.php
ATT	Getty Art & Architecture Thesaurus	http://www.getty.edu/research/tools/vocabularies/aat/
TGM	Library of Congress Thesaurus for Graphic Material	http://www.loc.gov/pictures/collection/tgm/
LCSH	Library of Congress Subject Headings	http://id.loc.gov/authorities/subjects.html
IANA	IANA Media Types	https://www.iana.org/assignments/media-types/media-types.xhtml

DCMI	DCMI Type Vocabulary	http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=dcmitype#H7
Rights	RightsStatements.org	http://rightsstatements.org/page/1.0/?language=en

PPC Fields at a Glance (listed by requirement)

See also: [PPC Required and Recommended Fields \(listed alphabetically\)](#)

Field Label	Required by DPLA?	Who supplies this data?	Standardized by the PPC?*	Displayed in DPLA?	Display Type
Data Provider	Required	Derived from Partner-supplied data		Yes	Short, long and Facet
Is Shown At (URL)	Required	Derived from Partner-supplied data		Yes	Long
Rights	Required	Partner Supplied		Yes	Long
Title	Required	Partner Supplied		Yes	Short and long
Language	Required When Applicable	Derived from partner-supplied data	Standardized	Yes	Long and facet
Preview	Required When Applicable	Partner Supplied		Yes	Image
Date Created	Strongly Recommended	Derived from partner-supplied data	Standardized	Yes	Short, long, facet and timeline
Place	Strongly Recommended	Partner Supplied		Yes	Long, facet and map

Subject	Strongly Recommended	Derived from partner-supplied data		Yes	Long and facet
Creator	Recommended	Partner Supplied		Yes	Short and long
Description	Recommended	Partner Supplied		Yes	Short and long
Format	Recommended	Partner Supplied		Yes	Long
Publisher	Recommended	Derived from partner-supplied data		Yes	Long
Type	Recommended	Derived from partner-supplied data	Standardized	Yes	Short, long, and facet
Alternative Title	Optional	Partner Supplied		No	
Contributor	Optional	Partner Supplied		No	
Extent	Optional	Partner Supplied		No	
Identifier	Optional	Partner Supplied		No	
Relation	Optional	Derived from partner-supplied value		No	

*For a limited number of fields the PPC will transform partner supplied data and standardized it to meet the DPLA requirements

PPC Required and Recommended Fields (listed alphabetically)

Field	Alternate Title
DPLA Status	Optional

Hub Status	Optional
Description	Any alternative title of the described resource including abbreviation and translation.
Qualified DC	dcterms:alternative
MARC	246 Varying Form of Title
MODS	mods:titleInfo (use type="alternative")
Repeatable	Yes
CV/Syntax	Natural Language
Notes and Best Practices	This is not meant to be repetitive of main title, however translation of foreign language titles are acceptable. Ensure that the alternative title is for the object, not the title of the series or the collection, or for other related objects. Avoid the use of explanatory or qualifying symbols (such as brackets).
Examples	<ul style="list-style-type: none"> Undergraduate course catalog, 1961-62

PPC Field	Contributor
DPLA Status	Optional
Hub Status	Optional
Description	An entity responsible for making contributions to the described resource.
Qualified DC	dcterms:contributor
MARC	Multiple fields are possible: 700; 710; 711; 720 when the relator term (subfield e) is not 'aut' or 'cre'
MODS	mods:name (with 'role' subelement)
Repeatable	Yes

CV/Syntax	LCNAF, ULAN, etc. If controlled vocabulary term is not available prefer/recommend the use of a syntax such as: Lastname, Firstname, birthyear-deathyear (if known). This assures that like names are sorted together
Notes and Best Practices	This field is used to note contribution to the original work. Examples of a Contributor include a person, an organization, or a service. This field should not be used to note the name of individuals who have cataloged or scanned the resource. Avoid the use of placeholders such as "Unknown."
Examples	<ul style="list-style-type: none"> • Illustrators • chapter authors • Oral history interviewer • United States. Army Map Service

PPC Field	Creator
DPLA Status	Recommended
Hub Status	Strongly Recommended
Description	An entity primarily responsible for making the resource.
Qualified DC	dcterms:creator
MARC	Same as above
MODS	mods:name (with 'role' subelement)
Repeatable	Yes
CV/Syntax	LCNAF, ULAN, etc if available. If controlled vocabulary term is not available prefer/recommend the use of a syntax such as: Lastname, Firstname, birthyear-deathyear (if known). This assures that like names are sorted together.

Notes and Best Practices	Examples of a Creator include a person, an organization, or a service. Can be used to indicate Maker role. Avoid the use of placeholder values such as “unknown.” For oral histories, Creator is interviewee.
Examples	<ul style="list-style-type: none"> • Beam, George L. (George Lytle), 1868-1935 • Lee, Herschel • United States. Geological Survey

PPC Field	Data Provider and Intermediate Provider
DPLA Status	Required
Hub Status	Required
Description	The organization or entity that supplies data to DPLA through the PPC.
Qualified DC	N/A See Notes and Best practices
MARC	N/A See Notes and Best practices
MODS	N/A See Notes and Best practices
Repeatable	No
CV/Syntax	Natural Language
Notes and Best Practices	This will be displayed in DPLA as “Contributing Institution.” When supplying content please provide the PPC with how you would like your institutional name displayed. In instances where one institution is hosting another institution’s content, the hosting institution will be mapped as the “intermediate” provider and the other will be the “data provider.” In these cases, when submitting records, please let the PPC know what local metadata field you are storing information that would be relevant to DPLA’s “Data Provider” field in instances where there is both a data provider and intermediate provider and we will map the information.

Examples	<ul style="list-style-type: none"> • Colorado College • History Colorado • Denver Public Library • DPLA example of both intermediate provider and data provider
-----------------	---

PPC Field	Date Created (Original)
DPLA Status	Strongly Recommended
Hub Status	Strongly Recommended
Description	Date of creation of the original resource.
Qualified DC	dcterms:created
MARC	260##\$c
MODS	mods:originInfo (with 'dateCreated' subelement)
Repeatable	No
CV/Syntax	EDTF (extended date/time format) YYYY-MM-DD or YYYY-YYYY is preferred.
Notes and Best Practices	This is not the day the item was digitized. Use of the EDTF schema is recommended to avoid ambiguity and to normalize the date format. You can give an exact date or a date range but each of those instances should be created in a consistent format. There are many ways to express uncertainty about dates. Recommend the use of the standard ways compatible with EDTF found in DPLA's Geographic and Temporal Guidelines (http://bit.ly/dpla-geo-styleguide). If those are not possible the use of internally consistent methods is recommended and the PPC will normalize metadata to EDTF . For date aboutness use subject field.
Examples	<ul style="list-style-type: none"> • 1973-05-22 • 1730-1750 • Date unknown, N/A or n.d. is not acceptable

PPC Field	Description
DPLA Status	Optional
Hub Status	Strongly Recommended
Description	A free text account of the resource.
Qualified DC	dcterms:description
MARC	520##\$a (Summary) ; 505#0\$a (Table of Contents)
MODS	mods:abstract
Repeatable	No
CV/Syntax	Natural Language
Notes and Best Practices	Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource. Not full text description or OCR output is allowed. Ensure that the description is of the object being described and not a collection to which it belongs or any other parent or child entity.
Examples	<ul style="list-style-type: none"> • White cotton batiste baby dress believed to have been worn by Greta Puckett as a baby in Nebraska circa 1902. The dress has embroidery and openwork in three rectangles and openwork across the front. It has long sleeves with a band, and two buttons in the back. • Photograph of an unidentified young man, possibly a graduation photograph that appears to have been taken in Nebraska, circa 1900-1910. The boy appears to be holding a diploma in one hand. The photograph was taken by the Lesmeister studio in Shelton, Nebraska.

PPC Field	Extent
DPLA Status	Optional

Hub Status	Recommended
Description	The size or duration of the original resource.
Qualified DC	dcterms:extent
MARC	300 ; 306
MODS	mods:physicalDescription
Repeatable	Yes
CV/Syntax	Natural language
Notes and Best Practices	Examples include a number of pages (letter), dimensions (Object), period of time in hours, minutes and sec (recording). Recommend use of consistency in handling measurement terminology for maximum understandability. Include the digital file size only when the resource is born digital.
Examples	<ul style="list-style-type: none"> ● 4 7/8 x 8 3/16 inches ● 1 map on 13 sheets ● 00:14:21 ● 406 pages ● Height x width x depth

PPC Field	Format
DPLA Status	Recommended
Hub Status	Recommended
Description	Physical medium or dimensions of described resource.
Qualified DC	dcterms:format

MARC	008/23 ; 338
MODS	mods:physicalDescription (with subelement 'internetMediaType' for born-digital materials, or 'extent' to describe the physical original from which the digital surrogate was created)
Repeatable	Yes
CV/Syntax	Recommend use of a controlled vocabulary (TGM, AAT, etc.) is highly recommended. The information can be relevant to determine the equipment needed to display or operate a born-digital resource (e.g. if the described resource has format pdf you need a pdf reader to use it). For that purpose you can use IANA type.
Notes and Best Practices	Format is a more granular description of the type of object described than the simple vocabulary used in the Type field. It can encompass description of the medium, materials, genre, or other similar terms.
Examples	<ul style="list-style-type: none"> ● application/pdf ● audio/mpeg ● image/tiff ● video/mpeg ● videocassette ● gelatin silver negatives ● Broadsides

PPC Field	Identifier
DPLA Status	Optional
Hub Status	Optional
Description	ID of described resource within a given context.
Qualified DC	dcterms:identifier
MARC	020 (ISBN) ; 022 (ISSN) ; 024 (Other identifier)

MODS	mods:identifier (with type="uri") -- Persistent identifiers For other uses, include a relevant 'type' value and add accession number, call number, etc.
Repeatable	Yes
CV/Syntax	Natural language
Notes and Best Practices	An institution could have more than one instance of an identifier, for example an accession number and a call number.
Examples	<ul style="list-style-type: none"> ● P441110B ● HPHWPZ201404000165 ● 1999-002_006 ● Accession number ● Object ID

PPC Field	Is Shown at (URL)
DPLA Status	Required (can be derived from OAI feed)
Hub Status	Required
Description	Unambiguous URL reference to digital objects in its full information context.
Qualified DC	N/A
MODS	Mods:location (with subelement 'url')
Repeatable	No
CV/Syntax	Must be a URL
Notes and Best Practices	This field is used so that DPLA visitors can link back to the object record at the home institution that displays the full metadata associated with the object.

Examples	<ul style="list-style-type: none"> • http://5008.sydneyplus.com/HistoryColorado_ArgusNet_Final/ViewRecord.aspx?template=Object&record=f5f22708-7016-46b2-9dc0-e09950c02d42&displayFields=Attachment&lang=en-US
-----------------	---

PPC Field	Language
DPLA Status	Required when applicable
Hub Status	Required when applicable
Description	A language of the resource.
Qualified DC	dcterms:language
MARC	041
MODS	mods:language
Repeatable	Yes
CV/Syntax	Controlled vocabulary, ISO639-3, RFC4646, Lexvo (URL) is preferred
Notes and Best Practices	Strongly recommended for text materials. List multiple entries separated with a semicolon. PPC will normalize data
Examples	<ul style="list-style-type: none"> • German • eng • fre • http://www.lexvo.org/page/iso639-3/dan

PPC Field	Place
DPLA Status	Strongly recommended

Hub Status	Strongly recommended
Description	Spatial characteristics of the described resource, such as a country, city, region, address or other geographical term. Captures aboutness. Geographic location relevant to the original item.
Qualified DC (Recommended)	dcterms:spatial
MARC	522##\$a
MODS	mods:subject (with subelement 'geographic')
Repeatable	Yes
CV/Syntax	Recommend the use of a controlled vocabulary such as LCNAF, TGN, GeoNames or consistent local vocabulary.
Notes and Best Practices	Please see DPLA's Temporal and Geographic Guidelines http://bit.ly/dpla-geo-styleguide . Use only for spatial topics that a resource is about.
Examples	<ul style="list-style-type: none"> ● Laramie (Wyo.) ● Ouray County (Colo.) ● Denver, Colorado ● http://www.geonames.org/maps/google_39.739_-104.985.html ● DPLA Temporal and Geographic Guidelines has several other examples.

PPC Field	Preview
DPLA Status	Required when applicable
Hub Status	Required when applicable

Description	The URL of a thumbnail, extract, preview or other type of resource representing the digital object for the purposes of providing a preview.
Qualified DC	N/A
MARC	N/A
MODS	N/A
Repeatable	No
CV/Syntax	Must be a URL
Notes and Best Practices	<p>Each platform may supply this information differently. In some cases this information can be embedded into the feed or derived from the feed. During the ingest process, the PPC will work with each of its partners to determine the process for supplying this information.</p> <p>The URL must resolve to a digital object. It must be a URL to a thumbnail, not a landing page. A preview is highly recommended but not required by DPLA for text, video or audio. In those cases do not supply a generic image as DPLA prefers to supply their own generic icon. Any image supplied will displayed on the front end at 300px on longest side.</p>
Examples	

PPC Field	Publisher (original)
DPLA Status	Recommended
Hub Status	Recommended
Description	Entity responsible for making the described resource available, typically the publisher of a text.
Qualified DC	dcterms:publisher

MARC	260##\$b ; 264##\$b
MODS	mods:originInfo (with subelement 'publisher')
Repeatable	Yes
CV/Syntax	LCNAF or VIAF
Describing	Original Resource
Notes and Best Practices	The field is intended to contain the publisher of the original item, not institutions involved in its digitization or sharing. Use this field for published materials such as books, magazines and journals. Avoid placeholder values like "unknown".
Examples	<ul style="list-style-type: none"> • Rand McNally and Company

PPC Field	Relation
DPLA Status	Optional
Hub Status	Optional
Description	A related resource.
Qualified DC	dcterms:relation
MARC	Many options, including 856 (all-purpose) and 555 (finding aids/indexes)
MODS	mods:relatedItem
Repeatable	Yes
CV/Syntax	Free text. Recommended use of a local controlled vocabulary. Can include a URL.
Notes and Best	Relation is intended for use with other items that have some relationship with the content. May be used to

Practices	indicate that items are related based on accession, series, collection, provenance or theme. Recommend only using information that is intelligible outside of the original institution context. For example, call numbers or identifiers that do not make sense out of context should not be included.
Examples	<ul style="list-style-type: none"> • Colorado Italian Americans Collection • George Lytle Beam Photograph Collection

PPC Field	Rights
DPLA Status	Required
Hub Status	Required
Description	Information about rights held in and over the resource. Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.
Qualified DC	dcterms:Rights
MARC	540##\$a
MODS	mods:accessCondition (with type="use and reproduction" xlink:href="[URI of rightsstatements statement]")
Repeatable	No
CV/Syntax	Must be a URL. http://rightsstatements.org
Notes and Best Practices	Use the provided URL on the rightsstatements.org website that best matches the rights associated with that digital object. DPLA will then display the label, description and icon on dp.la, for example https://dp.la/item/723513064603690b2c9a28ffc6fd15a5 . The use of rightsstatements.org is only required for the metadata supplied to DPLA. You can use your own local rights statement at your home institution.
Examples	<ul style="list-style-type: none"> • http://rightsstatements.org/page/InC/1.0/

PPC Field	Subject
DPLA Status	Strongly Recommended
Hub Status	Strongly Recommended
Description	The topic of the resource. Typically, the subject will be represented using keywords, key phrases, or classification codes.
Qualified DC	dcterms:subject
MARC	6XX
MODS	mods:subject
Repeatable	Yes
CV/Syntax	Use of a controlled vocabulary (LCSH, TGM, FAST, etc.) or name authority (LCNAF, VIAF, etc.) is highly recommended.
Notes and Best Practices	<p>If you are using a controlled vocabulary, recommend the use of the URI in addition to the string value depending on the schema used. Recommend uncoordinated subject heading if headings are being newly created.</p> <ul style="list-style-type: none"> For example, in a Dublin Core record you might use: <pre><dc:subject>Civil rights movements</dc:subject> <dc:subject>Mississippi</dc:subject> <dc:subject>Jackson</dc:subject></pre> <p>Instead of</p> <pre><dc:subject>Civil rights movements -- Mississippi -- Jackson</dc:subject></pre>

	This suggestion is made to increase matching of terms in the aggregated data set. Not all providers will use the same controlled vocabulary lists. Even among those who do, the granular nuance of the coordinated subject headings makes it impossible to bring together records based on the larger concepts present in the heading.
Examples	<ul style="list-style-type: none"> ● Women ● Sweetland, Henry Hale, 1848-1938 ● Coal Miners

PPC Field	Title
DPLA Status	Required
Hub Status	Required
Description	A name given to the resource. Typically, a Title will be a name by which the resource is formally known.
Qualified DC	dcterms:title
MARC	245 & 246
MODS	mods:titleInfo
Repeatable	No
CV/Syntax	Natural Language
Notes and Best Practices	When titles are created for works, they should be concise. The description field should be used for more detail. Descriptive and informative titles are preferred whenever possible (as opposed to things like "unknown" or an id number). Not all materials can or should be titled uniquely. This recommendation exists to encourage data creators to create unique and informative titles when they can. Recommend minimal but appropriate use of punctuation. DPLA prefers that titles not have unnecessary quotation marks, brackets or ending periods.

Examples	<ul style="list-style-type: none"> • General View of Mesa at Tsankawi Ruin, Bandelier National Monument, N. M. • Golden Jubilee Program • Battenburg Lace Doily
-----------------	--

PPC Field	Type
DPLA Status	Required when available
Hub Status	Required when Available
Description	The nature or genre of the resource. Ref (Strongly Recommended)
Qualified DC	dcterms:type
MARC	336
MODS	<p>Depends on institutional practice:</p> <ul style="list-style-type: none"> • mods:typeOfResource -- if using MODS-based Type vocabulary) • mods:genre (with type="dct") -- if using DCMI CV
Repeatable	Yes
CV/Syntax	Recommend use of DCMI Type Vocabulary URLs
Notes and Best Practices	<p>Type is intended as a broad categorization, not a more granular term/field like format or genre. Types need to be distinguishable from these other terms. Recommended best practice is to assign the type Text to images of textual materials. Use the DCMI type vocabulary if possible, or an internal standard that can be mapped and is consistent and the PPC will normalize the metadata to DCMI terms. Some examples of other vocabularies that PPC could easily transforms are the MODS type of resource values (http://www.loc.gov/standards/mods/mods-outline-3-6.html#typeOfResource) and the Library of Congress's Content Types list (http://id.loc.gov/vocabulary/contentTypes.html).</p>

Examples

- Text
- Image
- Physical Object
- Sound
- Moving Image