TxHUB METADATA GUIDELINES

Version 1.0 (November, 2019)

These guidelines were drafted by the Texas Digital Library DPLA Metadata Working Group:

Albert Duran, Houston Public Library

Liz Gushee, University of Miami

Paloma Graciani Picardo, Harry Ransom Center

Kara Long (chair), Baylor University

Hannah Tarver, University of North Texas

Katie Tomfohrde, Houston Public Library

Anne Washington, University of Houston

Nicholas Woodward, Texas Digital Library

TABLE OF CONTENTS

INTRODUCTION

This document is intended to provide guidance for Texas institutions that want to have their digital collections aggregated into the TxHub which will be harvested by the Digital Public Library of America (DPLA). DPLA is an aggregated collection, describing the digital and digitized materials from libraries, archives, and museums across the United States. These rich resources are brought together under a single, searchable portal through DPLA.

The DPLA has a defined collection scope (DPLA Collection Development Guidelines). Some types of content require review by DPLA, such as scholarly materials (theses and dissertations, recent journal articles), finding aids (EADs, collection guides). TxHub is currently not harvesting collections of scholarly materials. In addition, TxHub will not harvest items that do not resolve to a publicly accessible URL in the institution's local repository.

Participating member institutions will work with TDL to evaluate their metadata and map collection fields to the hub aggregation metadata application profile. Section II describes technical requirements for this process and Section III describes the required metadata fields and suggestions to ensure that field values are findable and useful within the aggregated collections. These guidelines incorporate best practices for each metadata field based on the DPLA Metadata Application Profile, version 5.

ACKNOWLEDGEMENTS

The TDL DPLA Metadata working group would like to thank Texas Digital Library for recognizing the need to pursue this worthwhile project and giving us the opportunity to be a part of it.

The authors acknowledge that this endeavor would not have been possible without the continued support and encouragement of our home institutions: Texas Digital Library, Baylor University, University of North Texas, University of Texas at Austin, University of Houston, and Houston Public Library.

We would also like to convey sincere thanks and gratitude for the exemplary work of our community reviewers for their invaluable time and advice. Recognition and thanks go to Itza Carbajal, Rachel Appel, Elliot Williams, Melanie Cofield, Jeremy Bartczak, Carla Alvarez, and Laura Capell, who reviewed and provided useful feedback for this document. We have also tremendously benefitted from the guidance, suggestions, and feedback that our distinguished colleagues and anonymous peer reviewers have submitted via an online survey.

The TxHub Metadata Guidelines are deeply informed by the work of the DPLA and other DPLA hubs and service providers, including the following works:

- [PA Digital PA-DPLA Metadata Guidelines](#)
- [DPLA Metadata Application Profile (version 5.0, December 2017)](#)
- [UNT Digital Projects: Input Guidelines for Descriptive Metadata](#)
- [University of Houston Metadata Dictionary](#)
- [Metadata Best Practices for the Illinois Digital Heritage Hub](#)

We deeply thank Gretchen Gueguen in her role as DPLA Data Services Coordinator and Network Manager for her time and guidance on numerous DPLA metadata requirements and practices. We would also like to thank Michael Della Bitta at DPLA for his support as DPLA Director of Technology.

TECHNICAL REQUIREMENTS

In order to contribute through the TDL DPLA Service Hub, your digital collections must

- Be able to provide metadata through OAI-PMH
- Be available on a publicly accessible site (i.e., no paywalls, IP restrictions, or required log-ins)
- Be able to store and provide preview or thumbnail images of the objects in the collection (this is a common feature in most systems)
- Be able to link back to the record on your digital collection site via a static link
- Be willing to allow metadata to be made accessible under a [CC0 license](#) through DPLA (see the [DPLA Metadata Application Profile](#) for more information)

Commonly used repositories that are compatible with OAI-PMH include CONTENTdm, Bepress, Omeka, Islandora, DSpace and others. These guidelines reference mappings in Dublin Core and MODS. TxHub will also accept mapped METS metadata but the guidelines do not contain those mappings. Please contact Texas Digital Library for assistance mapping METS to TxHub.

Additionally, TxHub is pleased to work with member institutions that have implemented IIIF (International Image Interoperability Framework) as their image delivery technology so that the IIIF information can be mapped to the appropriate fields in the DPLA MAP v.5. More information about TxHub mappings for IIIF URIs and manifests will be provided by TDL.

METADATA GUIDELINES

In an aggregated environment, some metadata may lack the full context of the local digital collection. This can cause issues such as meaningless decontextualized values, incomplete fields, or confusing fields. DPLA may not display all of the metadata that your institution is able to provide, so it is important to consider your field mappings in the context of what DPLA can display, and what metadata is necessary to identify items in your collection outside of their original context. This profile focuses on the metadata elements that are aggregated and displayed through TxHub and DPLA.

Note also that while these metadata guidelines are largely based on the DPLA MAP v.5, TxHub has some additional metadata requirements that are incorporated into the guidelines below.

In the following sections, "Partner" refers to the institution/organization submitting their metadata to TxHub for inclusion in DPLA.

TxHUB METADATA ELEMENTS

TxHUB METADATA ELEMENTS – QUICK GUIDE
* Do not display in the DPLA portal.

Field Label	Definition	Requirement	Simple / Qualified Dublin Core	MODS
TxHub managed fields				
Collection*	TxHub supplied collection name or OAI set name	Required	dc:relation / dcterms:isPartOf	<relatedItem type="series"> or <relatedItem type="host">
Data Provider	Name of the partner institution supplying data	Required	dcterms:provenance	<name> with type="corporate"
Provider	TxHub supplied when applicable	Required		
Preview	Thumbnail URL	Required		

Field Label	Definition	Requirement	Simple / Qualified Dublin Core	MODS
	Institution Partner managed fields			
Alternate Title*	A secondary title for the resource, includes translations	Optional	dc:alternative	<titleInfo><subTitle > <titleInfo><title type="alternative," "uniform," "abbreviated," or "translated">
Contributor	Entity responsible for secondary contributions	Optional	dc:contributor / dcterms:contributor	<name><namePart > where <name> also contains <role><roleTerm>c ontributor</roleTer m></role>
Creator	Entity responsible for the resource	Recommended	dc:creator	<name><namePart > where a value of the <role><roleTerm> is "author" or "creator"
Date	Creation date of the item	Recommended	dc:date / dcterms:created	<originInfo><dateCr eated keyDate="yes">
Description	Free text description of the resource	Recommended	dc:description	<note type="content"> Or
Extent*	Size, duration, or physical description of the original resource	Optional	dcterms:extent	<physicalDescriptio n><extent>
Format	Description of the physical medium or genre of the resource	Recommended	dc:format / dcterms:medium	<physicalDescriptio n><form>
Identifier*	A string, box number, call number, or other identifier for the resource	Optional	dc:identifier	<identifier> or <location><url>

Field Label	Definition	Requirement	Simple / Qualified Dublin Core	MODS
Is Shown At	URL to item in institution's local repository	Required	dc:identifier	<identifier> or <location><url>
Language	Language(s) used in resource	Required if available	dc:language	<language><languageTerm>
Place	Location associated with the resource - captures aboutness	Recommended	dc:coverage / dcterms:spatial	<subject><geographic>
Publisher	Entity responsible for making the item available, e.g., publisher of text (not data provider)	Recommended	dc:publisher	<originInfo><publisher>
Relation	A related resource	Optional	dc:relation	<relatedItem><location><url> and / or <relatedItem><titleInfo><title>
Rights	Statement of rights provided by the partner institution	Required if Standardized Rights Statement is not available	dc:rights	<accessCondition>
Standardized Rights Statement	Statement from rightsstatements.org	Required if available	dc:rights	<accessCondition>
Subject	Topic describing what the resource is about	Recommended	dc.subject	<subject><topic>
Title	Primary name of item	Required	dc:title	<titleInfo><title>
Type	Nature/genre of resource	Recommended	dc:type	<typeOfResource>

TxHUB METADATA ELEMENTS – DETAILED GUIDE

The following sections detail input guidelines for each metadata field. The elements are divided into two sections. The Collection-Level Metadata Fields covers information provided by the TxHub to DPLA that are not necessarily mapped in the local collection. The Item-Level metadata fields are mapped from the local collection to TxHub. Guidelines are listed alphabetically by field label and sections include introductory information that apply generally to all mapped values.

These sections are intended to provide recommendations, but may be negotiable depending on local metadata. If there are required values that are not available or not used at your institution or other differences in usage, discuss any concerns or questions about mapping local repositories to the TxHub with TDL staff.

TABLE KEY

FIELD LABEL		
Simple / Qualified Dublin Core		MODS
Requirement	Required, Recommended, or Optional for the TxHub	
DPLA Usage	How DPLA uses this field, whether or not it displays in DPLA, and the DPLA requirement	
Definition	Definition of the field	
Repeatable	Yes or No	
Vocab/Syntax	Recommended controlled vocabularies, naming conventions, or formatting	
Notes	Additional information or considerations for the use of the field	
Examples	Examples of how the field can be used	

COLLECTION-LEVEL METADATA FIELDS

The following fields will be applied by TxHub and are required by DPLA. These fields do not need to be mapped in the local collection metadata. Values for the collection and data provider names can be arranged with TxHub at the time that metadata mappings are set up from the host institution to the hub aggregation.

COLLECTION	
dc:relation / dcterms:isPartOf	\<relatedItem type="series"\> or \<relatedItem type ="host"\>
Requirement	Required by TxHub
DPLA Usage	Does not display in DPLA. Strongly recommended by DPLA.
Definition	The name or title of the collection of which the resource is a part.
Repeatable	Yes
Vocab/Syntax	
Notes	This field does not display in DPLA at present, but there is an interest in collecting this information and potentially using it for display in the future. This value can be taken from the OAI set name and does not necessarily need to be mapped in the record.
Examples	Collection: War of the Rebellion Collection Collection: Edgar Allan Poe Collection

DATA PROVIDER

dcterms:provenance	<name> with type="corporate"
Requirement	Required
DPLA Usage	Displays in DPLA, and is required by DPLA. This field displays as "Contributing Institution."
Definition	The name of the contributing institution that is the custodian of the digital collection.
Repeatable	Yes
Vocab/Syntax	Name of partner institution
Notes	
Examples	Data Provider: Baylor University - The Texas Collection Data Provider: Harry Ransom Center

PROVIDER

No mapping needed. Value provided by hub.	
Requirement	Required
DPLA Usage	Displays in DPLA, and is required by DPLA. This field displays as "Partner."
Definition	The name of the hub that is providing the aggregated metadata to DPLA.
Repeatable	No
Vocab/Syntax	Name of hub
Notes	
Examples	Provider: TxHub

ITEM-LEVEL METADATA FIELDS

Item-level metadata is aggregated from the host digital library, and mapped directly into the TxHub. Some values may need to be normalized by TxHub. To complete this normalization, TxHub may:

- Use semicolons to separate multiple entries in a single field
- Take advantage of controlled vocabularies wherever possible
- Concatenate values in metadata elements mapped to the same Dublin Core or MODS elements; values will be separated with a triple dash (---)

Some fields described below are designated as "Does not display in DPLA." This means that the data values in those fields will not be visible on the DPLA portal; however, these fields can still be downloaded through the API and may have research value. It is at the partner institution's discretion whether or not to include these fields.

FIELD-SPECIFIC EXAMPLES

Information below is organized alphabetically by field label.

ALTERNATE TITLE

dcterms:alternative	<titleInfo><subTitle> <titleInfo><title type="alternative," "uniform," "abbreviated," or "translated">
Requirement	Optional
DPLA Usage	Does not display in DPLA.
Definition	An alternate name for the resource, including subtitles, abbreviations, or translations.
Repeatable	Yes
Vocab/Syntax	Natural language
Notes	May be used for alternative spellings.
Examples	Alternate Title: Courtship of Little Red and the Pin, volume 1

CONTRIBUTOR

dc:contributor / dcterms:contributor	<name><namePart> where <name> also contains <role><roleTerm>contributor</roleTerm></role>
Requirement	Optional
DPLA Usage	Does not display in DPLA.
Definition	Entity who has made contributions to the resource but who is not the primary creator.
Repeatable	Yes
Vocab/Syntax	- Since DPLA does not currently have a way to separate names from roles, TxHub does not recommend including the role of a contributor or a relationship designator in this field unless it is included in the authorized form of their name. Including relationship designators may impact discovery in the DPLA portal - Recommended vocabularies: - Library of Congress Name Authority File (LCNAF) - Handbook of Texas (HOT) - Getty name authorities - Local authority file - If not using a controlled vocabulary, it is recommended to use similar syntax to LCNAF (Lastname, Firstname, YYYY-YYYY)
Notes	
Examples	Corporate name Contributor: Zion & Breen Associates, Inc., Site Planners, Landscape Architects Personal name (local) Contributor: Ypiña, Emilio R.

CREATOR	
dc:creator	\<name>\<namePart> where a value of the \<role>\<roleTerm> is "author" or "creator"
Requirement	Recommended
DPLA Usage	Displays in DPLA. Recommended by DPLA.
Definition	The name of the entity primarily responsible for making the resource - may be a person, institution, corporate body or government agency.
Repeatable	Yes

Vocab/Syntax	• Since DPLA does not currently have a way to separate names from roles, TxHub does not recommend including the role of a creator or a relationship designator in this field unless it is included in the authorized form of their name; including relationship designators may impact discovery in the DPLA portal.
	• Recommended vocabularies:
	  ○ Library of Congress Name Authority File (LCNAF)
	  ○ Handbook of Texas (HOT)
	  ○ Getty name authorities
	  ○ Local authority file
	• If not using a controlled vocabulary, it is recommended to use similar syntax to LCNAF (Lastname, Firstname, YYYY-YYYY)

Notes	
Examples	Personal name Creator: Schlueter, Frank L. Government Agency Creator: United States. War Department.

DATE

dc:date / dcterms:created	<originInfo><dateCreated keyDate="yes">

Requirement	Recommended
DPLA Usage	Displays in DPLA. Recommended by DPLA.
Definition	The date the original resource was created. If a resource is about a specific time period, put that date in the Subject field. Do not enter the date of digitization.
Repeatable	No. Date ranges are acceptable, but a single date is preferred, if possible.
Vocab/Syntax	- Prefer use of ISO 8601 (W3CDTF) format (YYYY-MM-DD) with optional EDTF extension - Date ranges are acceptable - DPLA is able to parse a range of date information, including date information contain language terms (e.g. circa, c., ca., etc.) and uncertain dates -- see the DPLA Geographical and Temporal Guidelines for MAP 3.1 for more information
Notes	The preference is to accept a single date field, containing a single date that conforms to the above recommended standards. Due to the high level of variation in date terms and formatting, variable date metadata is acceptable and may be normalized either at the hub-level or through DPLA.
Examples	Exact date Date: 1975-01-07 Month precision Date: 1945-10 Approximate Date: 1975~ Range Date: 1981/1985 The following non-EDTF values are acceptable and can be parsed by DPLA: 199- [1992] 1992? 1992.12 19xx

DESCRIPTION

dc:description	<note type="content"> Or
Requirement	Recommended
DPLA Usage	Displays in DPLA.
Definition	A free-text expression of what the resource is "about." This may include an abstract, table of contents, or reference to a graphical representation, but is not limited to these.
Repeatable	Yes. If multiple fields are mapped to dc:description they will be concatenated upon aggregation.
Vocab/Syntax	Natural language
Notes	- Certain values may need the fuller context of the local collection - Prefer not to map transcripts to dc:description - Prefer item-level content descriptions rather than collection-level descriptions
Examples	More context is usually better for the Description field. For example: Less context - Description: 112 More context - Description: Plate 112 Even more context - Description: Plate 112 showing maps of fortifications from the Atlas to Accompany the Official Records of the Union and Confederate Armies, 1861-1865 Concatenated description from multiple fields - Description: Little girl, unidentified, sitting in front of a fig tree laden with figs. --- South Houston

EXTENT

dcterms:extent	<physicalDescription><extent>
Requirement	Optional
DPLA Usage	Does not display in DPLA.
Definition	The physical description - size, length or duration of the resource. For example, the dimensions of an object.
Repeatable	Yes
Vocab/Syntax	Natural language
Notes	
Examples	Extent: 1 score (7 p.) 34 cm.

FORMAT

dc:format / dcterms:medium	<physicalDescription><form>
Requirement	Recommended
DPLA Usage	Displays in DPLA. Recommended by DPLA.
Definition	Describes the genre, medium, or materials of the resource. This is a more granular description than the Type field.
Repeatable	Yes
Vocab/Syntax	Recommended vocabularies: ● Getty Art and Architecture Thesaurus (AAT) ● Library of Congress Thesaurus for Graphic Materials (LCTGM)
Notes	
Examples	Format: Color photographs Format: Correspondence Format: Sheet music

IDENTIFIER

dc:identifier		<identifier> or <location><url>
Requirement	Optional	
DPLA Usage	Does not display in DPLA.	
Definition	A reference to the resource within a given context.	
Repeatable	Yes	
Vocab/Syntax	● Recommend use of alphanumeric string without special characters (@, $, %, etc.) ● Recommend that the identifier is unique to other identifiers within an institution's collection	
Notes	This field can also be used to store OCLC numbers, accession numbers, call numbers, and other types of identifiers.	
Examples	Identifier: RGD0033-r0325 Identifier: Box 1, Folder 3 Identifier: PFORZ 417 Identifier: 12092904	

IS SHOWN AT

dc:identifier		<identifier> or <location><url>
Requirement	Required	
DPLA Usage	Displays in DPLA.	
Definition	Unambiguous URL reference to the item in its local repository.	
Repeatable	No	
Vocab/Syntax	URL must resolve to the digital object in its local context	
Notes	For the purposes of updating records and re-harvest, this URL may be used as a functional record identifier and should be a persistent URL.	
Examples	Is Shown at: https://texashistory.unt.edu/ark:/67531/metapth49953/	

LANGUAGE

dc:language	<language><languageTerm>
Requirement	Recommended
DPLA Usage	Displays in DPLA. Strongly recommended by DPLA for text materials.
Definition	The language or languages used in the resource. This may not be the same as the language of the record.
Repeatable	Yes
Vocab/Syntax	- ISO 639-3 - Alternatively, provide the English name of the language as provided in the ISO standard (http://www.loc.gov/standards/iso639-2/php/code_list.php) - Do not use for non-linguistic content - The use of placeholder values (e.g., "Unknown" or "no language") is discouraged
Notes	
Examples	Language: eng (English) Language: ger (German)

PLACE	
dc:coverage / dcterms:spatial	<subject><geographic>
Requirement	Recommended
DPLA Usage	Displays in DPLA. Recommended by DPLA.
Definition	Geographic location associated with the original item. This should be the location or place the resource is about, not necessarily where it was published or disseminated.
Repeatable	Yes
Vocab/Syntax	• Library of Congress Name Authority File (LCNAF) • LCSH • Getty Thesaurus of Geographic Names (TGN) • GeoNames • Handbook of Texas (HOT) • FAST • Local controlled vocabulary
Notes	• Leave the field blank if not applicable or if a place cannot be accurately determined • Addresses, latitude/longitude, or other forms of location markers may also be mapped to dcterms:spatial • DPLA recommends being consistent if using hierarchical values in a single instance of the term (e.g., "Erie, Pennsylvania, United States of America") • For latitude/longitude, DPLA prefers latitude first, with the two coordinates separated by a comma, limited to whole numbers and decimals as opposed to degree-minutes-seconds (e.g., "6.703, 46.627"), and, if used, compass directions should appear after the number. • Full names of places are preferred over abbreviations • If abbreviations are used, preferred use of standard postal code abbreviations (PA, TX, etc.) • There is additional information about how to best format place names for MODS and Dublin Core in the DPLA Geographical and Temporal Guidelines for MAP 3.1 • DPLA prefers comma separated strings of geographic places in ascending order within a single element, as opposed to atomized terms in separate elements

	PLACE
Examples	LCSH Place: Houston Ship Channel (Tex.) FAST Place: Virginia--Dinwiddie County Local: Place: Third Ward, Houston Geocode: northlimit=29.125; eastlimit=-98.625; southlimit=29; westlimit=-98.75;

PUBLISHER

dc:publisher	<originInfo><publisher>

Requirement	Recommended
DPLA Usage	Displays in DPLA. Recommended by DPLA.
Definition	Entity responsible for making the original resource available. Typically, this is a publishing house, a corporate body, a government agency, a university, or other organization.
Repeatable	Yes
Vocab/Syntax	• Recommended to not include the role of the person, or a relationship designator, unless it is included in the authorized form of their name • Recommended vocabularies: ○ Library of Congress Name Authority File (LCNAF) ○ Handbook of Texas (HOT) ○ Getty name authorities ○ Local authority file • If not using a controlled vocabulary, it is recommended to use similar syntax to LCNAF (Lastname, Firstname, YYYY-YYYY)
Notes	• Do not use this field for the entity making the digital version available • Use this field for published materials, such as books, magazines, journals, etc. • Leave the field blank if a publisher cannot be accurately determined or if there is no publisher • Avoid the use of placeholder values (e.g., Unknown, s.n., etc.)
Examples	LCNAF Publisher: Broadway Music Corp. Local: Publisher: Architectural Book Publishing Company

RELATION	
dc:relation	\<relatedItem>\<location>\<url> and/or \<relatedItem>\<titleInfo>\<title>
Requirement	Optional
DPLA Usage	Does not display in DPLA.
Definition	A related resource, such as the name of a collection, series, or group of thematic works. May be suitable for items that are part of an archival collection. This may be a literal value or a reference.
Repeatable	Yes
Vocab/Syntax	Natural language or reference, such as URL or URI
Notes	
Examples	Relation: Baguio, Camp John Hay, and Asin Hot Springs: https://id.lib.uh.edu/ark:/84475/do924949714

RIGHTS

dc:rights		\<accessCondition\>
Requirement	Required	
DPLA Usage	Displays in DPLA. Required by DPLA if no Standardized Rights Statement is available.	
Definition	Information about the rights held over the resource, generally a statement about copyright or intellectual property rights. Statement may also include information about reuse or applicable licenses.	
Repeatable	Yes. Partners may provide a local statement in natural language. Either a Rights Statement or a Standardized Rights Statement may be provided (see below). At least one of these fields must be provided.	
Vocab/Syntax	• Natural language • Local rights statement	
Notes	• An institutional rights statement in natural language is acceptable • If supplying a natural language rights statement and a URI, these statements should be in alignment and in two separate fields • If supplying a value in the Standardized Rights Statement field, it is preferred not to include a local statement unless it is necessary for clarification • For compound objects where there may be varying rights statements for different portions of the work, select the more restrictive rights statement -- additional guidance on this issue can be found in Aggregating and Representing Collections in the Digital Public Library of America [white paper] - http://bit.ly/dplaCollections • See Appendix A for additional resources	
Examples	Local statement Rights: Materials in this collection may be licensed for reproduction or publication, according to the policies of the UNT Libraries' Special Collections department. Restrictions may apply.	

STANDARDIZED RIGHTS STATEMENT

dc:rights	\<accessCondition\>
Requirement	Required if available
DPLA Usage	Displays in DPLA. Strongly preferred by DPLA.
Definition	Information about the rights held over the resource, generally a statement about copyright or intellectual property rights. Statement may also include information about reuse or applicable licenses.
Repeatable	Yes. Partners may provide a URI to a standardized rights statement and a local statement in natural language. At least one of these is required.
Vocab/Syntax	• URI for a standardized statement from RightsStatements.org • URI for a Creative Commons license
Notes	• Strongly prefer use of a persistent URI, corresponding to a standardized rights statement from RightsStatements.org • Fields containing a URI should contain no other text. • For compound objects where there may be varying rights statements for different portions of the work, select the more restrictive rights statement -- additional guidance on this issue can be found in Aggregating and Representing Collections in the Digital Public Library of America [white paper] - http://bit.ly/dplaCollections • For materials are under a Creative Commons (CC) license, the license should be included in the record and the item will have no standardized rights statement • See Appendix A for additional resources
Examples	Creative Commons License Rights: https://creativecommons.org/licenses/by/4.0/ Standardized statement Rights: http://rightsstatements.org/vocab/NoC-US/1.0/

SUBJECT

dc:subject	<subject><topic>
Requirement	Recommended
DPLA Usage	Displays in DPLA. Recommended by DPLA. Some subject formatting may be normalized in DPLA.
Definition	Topic(s) of the item or what the item is about, expressed in keywords, phrases, names, or subject headings.
Repeatable	Yes
Vocab/Syntax	Recommended vocabularies: • Library of Congress Name Authority File (LCNAF) • LCSH • Handbook of Texas (HOT) • FAST • MeSH • Getty Art and Architecture Thesaurus (AAT) • Local subjects and keywords
Notes	Be as consistent as possible in the application of subject terms.
Examples	LCSH Subject: Municipal government--Records and correspondence Subject: Travis County (Tex.)--Maps Subject: Kennedy, John F. (John Fitzgerald), 1917-1963 FAST Subject: Bee culture Subject: Railroads--Freight MeSH Subject: Pathological Conditions, Signs and Symptoms Subject: Eye Diseases Art & Architecture Thesaurus (AAT) Subject: elementary schools (buildings) Subject: filet lace Chenhall's Nomenclature Subject: Communication Artifacts Subject: Metalworking T&E Keywords Subject: horseback riding

TITLE

dc:title	<titleInfo><title>
Requirement	Required
DPLA Usage	Displays in DPLA. Required by DPLA.
Definition	A name given to the resource.
Repeatable	No
Vocab/Syntax	Natural language
Notes	Give preference to original titles provided by the creator of the resource. If no original title is present, use a cataloger devised title. Descriptive and informative titles are preferred. Avoid the use of titles containing filenames, identifiers, or archival box and folder information.
Examples	Title: Admiral T. C Hart with the U. S. High Commissioner

TYPE

dc:type	<typeOfResource>
Requirement	Recommended
DPLA Usage	Displays in DPLA. Strongly recommended by DPLA.
Definition	A broad description of the nature of the resource.
Repeatable	Yes
Vocab/Syntax	Recommended vocabularies: • DCMI Type Vocabulary • Local controlled vocabulary
Notes	Values that are not in the DCMI Type Vocabulary or from Appendix C: List of Preferred Subtype Values from the DPLA MAP v. 5 may be mapped to the Format field in DPLA.
Examples	Type: Image Type: Text

APPENDIX A
RIGHTS AND RIGHTS STATEMENTS.ORG

Selected documentation that may be helpful in applying standardized rights statements to materials in a digital collection. Resources are grouped by content: (1) information about the technical specifications and implementations of the standardized statements, (2) examples from institutions about how they have approached standardized statements, and (3) information about U.S. copyright law and determining rights.

The following resources are subject to periodic updates. Wherever possible, we have linked to the most recent version as of the writing of these guidelines.

A.1 INFORMATION ABOUT THE STANDARDIZED STATEMENTS

RESOURCE	LINK	SUMMARY / USAGE
Rightsstatements.org list of statements	http://rightsstatements.org/page/1.0/?language=en	Official page listing the standardized statements including: • Descriptions and usage for each statement • Statement URIs • Documentation and policies
Recommendations for the Standardized International Rights Statements [white paper]	http://rightsstatements.org/en/documentation/	Full-text recommendation about the creation and adoption of standardized rights statements including: • Rationale with background on the need for standardized rights and the development process • Specifications for each of the statements with detailed information
Rightsstatements.org White Paper: Requirements for the Technical Infrastructure for Standardized International Rights Statements	http://rightsstatements.org/files/170106requirements_for_the_technical_infrastructure_for_standardized_international_rights_statements_v1.2.pdf	Description of technical specifications for implementing standardized statements including: • URI design and usage • Data modeling • Expected behaviors of services using standardized statements

RESOURCE	LINK	SUMMARY / USAGE
DPLA Standardized Rights Statements Implementation Guidelines [white paper]	https://docs.google.com/document/d/1aInokOIIsgf-B4iMTXU33qYN5B2jA3s91KgWoh7DZ7Q/edit	"This document contains a description of DPLA's implementation of standardized rights statements and recommendations for the use of statements in records that will be shared with DPLA." It includes: ● Guidance about how to include standardized statements ● How the statements will be integrated into records by DPLA ● Suggestions by metadata standard for storing URIs in XML

A.2 DOCUMENTATION FROM OTHER INSTITUTIONS

RESOURCE	LINK	SUMMARY / USAGE
Implementing RightsStatements.org at the University of Miami Libraries (Laura Capell & Elliot Williams)	https://jlsc-pub.org/articles/abstract/10.7710/2162-3309.2254/	Paper describing the implementation of standardized statements at the University of Miami including: ● Steps they took for legacy collections ● Technical implementation ● Challenges & risk assessment ● Future plans
PA Digital Rights Resources	https://padigital.org/for-contributors/rights-resources/	Resources compiled by PA Digital for participant institutions, including: ● An interactive tool for selecting standardized statements ● Three webinar video modules: 　○ Copyright 101 　○ What is a Rights Statement 　○ Implementing Rightsstatements.org ● A list of references (also documented in this appendix)

RESOURCE	LINK	SUMMARY / USAGE
Providing Quality Rights Metadata for Digital Collections Through RightsStatements.org (Linda Ballinger, Brandy Karl, and Anastasia Chiu)	http://palrap.pitt.edu/ojs/index.php/palrap/article/view/157	"This article describes the rights statements, provides guidance and examples on how to apply the most common ones, and addresses common pitfalls." It includes: ● Descriptions of the statements ● Recommendations and examples for each statement ● The workflow that Penn State used to implement statements ● Copyright determinations and recording them ● Providing information in addition to the standardized statements
Guide to Implementing Rights Statements from RightsStatements.org (Society of American Archivists)	https://www2.archivists.org/standards/guide-to-implementing-rights-statements-from-rightsstatementsorg	Recommendations from SSA including: ● General rights information ● Rights by material types ● A flowchart to help assess rights for archival materials
TRLN Rights Statements Report. A Roadmap for Implementing RightsStatements.org Statements (Triangle Research Library Network)	https://docs.google.com/document/d/1KgPXjG4RcHmK1Dy-hZ8JBNsG8CJlRUGvCJ2Nec-5Fj4/edit	Document identifying roles and responsibilities of those who might implement the statements, guidance for planning and documentation, as well as a few example workflows.
Rights Statements Resources (Minnesota Digital Library)	https://mndigital.org/standards-best-practices/rights-statements-resources	Rights statements resources and videos from the Minnesota Digital Library

A.3 GUIDANCE ABOUT COPYRIGHT IN THE U.S.

RESOURCE	LINK	SUMMARY / USAGE
Copyright Term and the Public Domain in the United States (Peter Hirtle)	https://copyright.cornell.edu/publicdomain	Tables listing current copyright terms for materials based on publication date, location of publication, and material types
Copyright and Cultural Institutions: Guidelines for Digitization for U.S. Libraries, Archives, and Museums (Peter Hirtle, Emily Hudson, and Andrew Kenyon)	https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1495365	Article describing copyright concerns specific to cultural heritage institutions, especially in regards to digitization, such as compliance, rights of copyright owners, exemptions for institutions, and risk assessment for projects.
The Online Books Page: Copyright Registration and Renewal Records (University of Pennsylvania)	http://onlinebooks.library.upenn.edu/cce/	Database version of the U.S. Copyright Office's Catalog of Copyright Entries with a list of other summaries and indexes to find registered copyrights for materials
Copyright Renewal Database (Stanford University)	https://exhibits.stanford.edu/copyrightrenewals	"[A] searchable index of the copyright renewal records for books published in the US between 1923 and 1963. Note that the database includes only renewal records, not original registrations, and only Class A (book) renewals received by the US Copyright Office between 1950 and 1992."

APPENDIX B
RESOURCES FOR METADATA ASSESSMENT AND CLEAN-UP

B.1 GENERAL RESOURCES / TOOLS

These resources include lists/databases of information about metadata assessment and generic applications.

RESOURCE	LINK	DESCRIPTION / USAGE
DLF AIG Metadata Application Profile Clearinghouse	https://dlfmetadataassessment.github.io/MetadataSpecsClearinghouse/	Collection of documentation submitted by institutions including: ● Application profiles ● Mappings ● Code ● Best practices
DLF AIG Metadata Working Group Metadata Assessment Toolkit	http://dlfmetadataassessment.github.io	Documentation collected by the Working Group including: ● Literature related to metadata quality ● Suggestions for applying metrics ● Tools for assessment with specifications and reviews
OpenRefine	http://openrefine.org/	Used to analyze/clean up data in spreadsheets ● May not work for datasets larger than 250K
MarcEdit	https://marcedit.reeset.net/	Editor for working with MARC records
AutoHotKey	https://autohotkey.com/	Windows-based automation scripting

B.2 SELECTED IMPLEMENTATIONS / TOOLS

This is a non-exhaustive list of tools that may be helpful for metadata assessment and/or corrections, roughly grouped by usage.

RESOURCE	LINK	DESCRIPTION / USAGE
Reconcile CSV (Open Knowledge Labs)	http://okfnlabs.org/reconcile-csv/	Use for: CSV files ● Allows OpenRefine to use fuzzy matching to reconcile data against a local CSV file
XML Starlet	http://xmlstar.sourceforge.net/	Use for: XML files ● Command-line tools that can transform, query, validate, and edit XML files using shell commands
Python Metadata Breakers (Mark Phillips)	https://github.com/vphill/metadata_breakers	Use for: XML files ● Break or atomize OAI-PMH repositories into simpler text formats ● Described further in this article - Metadata Analysis at the Command-Line: https://digital.library.unt.edu/ark:/67531/metadc157309/
Metadata Quality Analysis (Christina Harlow)	https://github.com/cmh2166/metadataQA	Use for: XML files ● Additional information and usage about harvesting/analyzing records using the metadata breakers
AutoHotKey script (Andrew Weidner)	https://github.com/drewhop/AutoHotkey/blob/master/TDNP_Metadata.ahk	Use for: files in Windows ● Sample script to do basic quality control on files in a Windows file browser ● Described further in this article - Workflow Tools for Digital Curation: http://journal.code4lib.org/articles/8419
Hunting	https://github.com/uhlibraries-digital/hunting	Use for: CONTENTdm ● "Hunting is a Ruby wrapper for the CONTENTdm API. Quickly 'Scout' for collections and objects in your Repository, 'Hunt' for metadata in your Collections, and 'Trap' individual Digital Objects."
CONTENTdm Catcher	https://www.oclc.org/support/services/contentdm/help/add-ons-help/contentdm-catcher.en.html	Use for: CONTENTdm ● "[A] Web service that supports batch requests to add metadata-only records to a CONTENTdm collection, edit existing metadata fields in a record, or delete a record"

APPENDIX C
SUGGESTED CONTROLLED VOCABULARIES AND STANDARDS

The following is an annotated list of controlled vocabularies suggested on these guidelines. The vocabularies are organized by entity type, however the same vocabulary could be used to describe different type of entities.

ENTITY TYPE	VOCABULARY	LINK	DESCRIPTION / USAGE
Concept	Medical Subject Headings (MeSH)	https://meshb.nlm.nih.gov/search	Hierarchically-organized terminology for indexing and cataloging of biomedical information such as MEDLINE/PUBmed and other NLM databases. MeSH is also published as Linked Open Data (LOD): https://id.nlm.nih.gov/mesh/.
Concept & Place	Library of Congress Subject Headings (LCSH)	http://id.loc.gov/authorities/subjects.html	One of the library domain's most widely used subject terminology schemas. Many of the Library of Congress controlled vocabularies have been published as Linked Open Data (LOD): http://id.loc.gov/.
Concept, Name & Place	Faceted Application of Subject Terminology (FAST)	https://fast.oclc.org/searchfast/	A simplified syntax of LCSH developed in collaboration of OCLC Research and Library of Congress with the intention of making the schema easier to understand, control, apply, and use. FAST has an available Linked Data Service: https://experimental.worldcat.org/fast/
Date	W3CDTF (W3C Date and Time Formats)	https://www.w3.org/TR/NOTE-datetime	Profile of the International Standard for the Representation of Dates and Times (ISO 8601), developed by the W3C.
Date	EDTF (Extended Date Time Format)	https://www.loc.gov/standards/datetime/edtf.html	Extension developed by the Library of Congress to include a number of date and time format conventions in common use that are not supported on the ISO 8601. An undergoing revision of ISO 8601 to be published mid 2019, is expected to include the EDTF specification as a profile.

ENTITY TYPE	VOCABULARY	LINK	DESCRIPTION / USAGE
Language	ISO 639-2	https://www.loc.gov/standards/iso639-2/php/code_list.php	Codes for the representation of Names of Languages maintained by the Library of Congress as the registration authority. This part 2 of the ISO 639 includes 3 letter codes referred to as "Alpha-3" codes.
Medium & Concept	Art and Architecture Thesaurus (AAT)	http://www.getty.edu/research/tools/vocabularies/aat/	Polyhierarchical thesaurus developed by the Getty to describe material culture and archival materials. In addition to the hierarchical relationships, the AAT has equivalence and associative relationships. All the Getty Vocabularies have been published as Linked Open Data (LOD): http://www.getty.edu/research/tools/vocabularies/lod/.
Medium	Library of Congress Thesaurus for Graphic Materials (LCTGM)	http://id.loc.gov/vocabulary/graphicMaterials.html	Tool for indexing visual materials by subject and by genre/format. Many of the Library of Congress controlled vocabularies have been published as Linked Open Data (LOD): http://id.loc.gov/.
Name	Union List of Artist Names (ULAN)	http://www.getty.edu/research/tools/vocabularies/ulan/	Vocabulary developed by the Getty to include proper names or anonymous appellations, biographies, related people or corporate bodies, and other information about artists, architects, firms, studios, museums, etc. All the Getty Vocabularies have been published as Linked Open Data (LOD): http://www.getty.edu/research/tools/vocabularies/lod/.
Name & Place	Virtual International Authority File (VIAF)	https://viaf.org/	Service hosted by OCLC that aggregates and clusters Name Authority data from major National Libraries and international projects such as ISNI and Wikidata.
Name & Place	Library of Congress Name Authority File (LCNAF)	http://id.loc.gov/authorities/names.html	Authoritative data for names of persons, organizations, events, places, and titles. Its purpose is the identification of these entities and, through the use of such controlled vocabulary, to provide uniform access to bibliographic resources. Many of the Library of Congress controlled vocabularies have been published as Linked Open Data (LOD): http://id.loc.gov/.

ENTITY TYPE	VOCABULARY	LINK	DESCRIPTION / USAGE
Name & Place	Handbook of Texas (HOT)	https://tshaonline.org/handbook/online	Digital encyclopedia developed by the Texas State Historical Association (TSHA) to cover the entire history of Texas from the indigenous Native Americans and the Prehistoric Era to the state's diverse population and the Modern Age.
Place	Getty Thesaurus of Geographic Names (TGN)	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	Thesaurus of place names with emphasis on historical places. Although it includes coordinates, TGN is not a GIS. All the Getty Vocabularies have been published as Linked Open Data (LOD): http://www.getty.edu/research/tools/vocabularies/lod/.
Place	GeoNames	https://www.geonames.org/	Native LOD ontology that integrates geographical data such as names of places (in various languages), elevation, population and other information from NGA (National Geospatial-intelligence Agency), GNIS (US Geological Survey Geographic Names Information System), and other gazetteers. The GeoNames Webservices allow users to interact with the data in multiple ways: http://www.geonames.org/export/web-services.html.
Rights	RightsStatements.org	https://rightsstatements.org/en/	Set of standardized rights statements that can be used to communicate the copyright and re-use status of digital objects to the public. See Appendix A for further information regarding the implementation RightsStatments.org.
Rights	Creative Commons License	https://creativecommons.org/share-your-work/	Set of standardized copyright licenses maintained by the nonprofit legal organization Creative Commons.
Type	DCMI Type Vocabulary	http://dublincore.org/documents/dcmi-type-vocabulary/	High level controlled vocabulary developed by the Dublin Core Metadata Initiative to describe the nature or genre of the resource.

REFERENCES AND FURTHER READING

Digital Projects: Input Guidelines for Digital Collections. University of North Texas, http://www.library.unt.edu/digital-projects-unit/input-guidelines-descriptive-metadata. Accessed October 11, 2018.

DPLA, DPLA Geographic and Temporal Guidelines for MAP 3.1, December 2017, http://bit.ly/dpla-geo-styleguide-3_1. Accessed October 11, 2018.

DPLA, [Software Repositories], GitHub, https://github.com/dpla. Accessed October 30, 2018.

DPLA Archival Description Working Group, Aggregating and Representing Collections in the Digital Public Library of America, November 2016, http://bit.ly/dplaCollections

DPLA MAP Working Group, Introduction to the DPLA Metadata Application Profile, version 5.0, December 2017, http://dp.la/info/map. Accessed October 11, 2018.

DPLA MAP Working Group, Metadata Application Profile, version 5.0, December 2017, http://dp.la/info/map. Accessed October 11, 2018.

PA Digital, PA Digital Metadata Guidelines: Requirements, Recommendations, and Best Practices for Preparing Metadata for PA Digital's DPLA Aggregator, version 1.2, August 2018, https://drive.google.com/file/d/1-6P9QtdWuxQGLKkHpcOqGU1_nMZoW20V/view

Phillips, Mark, "Metadata analysis at the Command-Line." Code4Lib Journal, issue 19, 2013-01-15, https://journal.code4lib.org/articles/7818

Ronallo, Jason, "IIIF Technical Workshop." http://ronallo.com/iiif-workshop-new/. Accessed December 21, 2018

Stitzlein, Hannah E., Patricia Lampron, and Myung-Ja K. Han. Metadata Best Practices for the Illinois Digital Heritage Hub, January 2017, https://docs.google.com/document/d/12NYtWUO9WNzewoXcFuKgpLsCvFA4fj2VCx9fAiiP1PE/edit#. Accessed October 11, 2018.

Tarver, Hannah, Oksana Zavalina, and Mark Phillips, "An Exploratory Study of the Description Field in the Digital Public Library of America." Proceedings International Conference on Dublin Core and Metadata

Applications 2016. October 14, 2016. https://digital.library.unt.edu/ark:/67531/metadc910346/. Accessed October 11, 2018.

University of Houston Libraries, BCDAMS MAP Metadata Input Guidelines, https://vocab.lib.uh.edu/bcdams-map/guidelines. Accessed October 12, 2018.

University of North Texas Libraries, Digital Projects: Minimally-Viable Records, https://www.library.unt.edu/digital-projects-unit/minimally-viable-records. Accessed November 20, 2018.