

DPLA in Ohio Metadata Final Report—April, 2016

Table of Contents

[Table of Contents](#)

[Executive Summary](#)

[Introduction](#)

[Summary of Recommendations](#)

[Working Group Members](#)

[Working Group Charge](#)

[Environmental Scan](#)

[Key Survey Findings on Metadata in Ohio](#)

[Primary Challenges and Proposed Solutions](#)

[Challenge 1: Ensuring the quality and consistency of metadata](#)

[Challenge 2: CC0 License for Metadata](#)

[Challenge 3: Required Rights Field](#)

[Challenge 4: Variety of metadata expertise across libraries in Ohio](#)

[Challenge 5: Familiarity/capability related to OAI-PMH](#)

[Challenge 6: Thumbnail/preview images](#)

[Challenge 7: Limited Scope of DPLA Collections](#)

[Challenge 8: Future Enhancements and/or Revisions to the DPLA's Metadata](#)

[Application Profile \(MAP\)](#)

[Additional Challenges](#)

[Proposed Metadata Education & Advocacy Activities](#)

[Proposed Metadata Application Profile](#)

[Introduction](#)

[Described Resource Elements](#)

[Derived Elements](#)

[Element Details](#)

[Appendix A: Controlled Vocabularies and Syntax](#)

[Appendix B: Metadata Best Practices Resources](#)

Executive Summary

Introduction

As the DPLA Ohio project embarks on planning a statewide DPLA hub - a partnership that will bring unprecedented access to Ohio's online cultural heritage collections - metadata concerns represent a primary set of challenges to the success of the project. There are myriad issues related to metadata in digital collections, issues which are further complicated by the need to coordinate multiple stakeholders with thousands of records to provide a single feed to the DPLA.

The Metadata Working Group was charged with identifying these issues and recommending approaches for DPLA Ohio and contributing partners. The group began meeting in October 2015 and discussing existing information and standards from both DPLA and other state Service Hubs. The group then identified sections of information to provide to the Steering Committee in this report, and divided up the work so that all members of the group participated in writing this report.

Summary of Recommendations

1. Because Ohio does not have a set of statewide metadata guidelines in place, the Metadata Working Group recommends adapting the Pennsylvania Digital Collections Project (PDCP) PA-DPLA Metadata Guidelines for use in Ohio. See [Proposed Metadata Best Practices](#) and [Proposed Metadata Application Profile](#) sections for more details.
2. The DPLA Ohio project must provide metadata best practices and guidelines to participating institutions and they should be presented as clearly as possible. Some additional materials may need to be created to assist specialized institutions (museums, historical societies) with metadata standards for DPLA Ohio. In addition to a Metadata Application Profile, a Metadata Best Practices document should be created. See Appendix B for a possible starting place for this document.
3. The required metadata fields for aggregation in DPLA Ohio should be kept minimal, in line with the requirements of DPLA.
 - a. Required: Title, Rights
 - b. Required When Available: Collection*, Language, Type
 - c. Strongly Recommended: Date, Place, Subject

*NOTE: The Working Group is considering designating Collection as a Required field.

4. Because of the complexity of metadata issues, compounded by the number of potential partners in a statewide DPLA Service Hub, a dedicated individual in a metadata specialist position will be critical throughout the pilot, and therefore should be accounted for in any proposed budget for the pilot phase of DPLA Ohio project. If the pilot is successful, it will be important to review the value of this position and how it might need to change in the post-pilot DPLA Ohio program.
5. The proposed Community Engagement Centers should facilitate professional development related to metadata remediation for institutions. This should be coordinated with the State Library and the DPLA Ohio Project Manager.
6. Metadata remediation should occur at the participating institutions, remain the responsibility of the institution's staff, and be supported by DPLA Ohio's best practices and guidelines. This remediation may be considered an iterative process, with updates and improvements occurring over time, and with repeated metadata harvesting.
7. A standing Metadata Working Group should be established for DPLA Ohio and charged with creating documentation, sharing metadata best practices and monitoring metadata issues and potential changes to the DPLA MAP.

Working Group Members

Co-chairs: Meghan Frazer, OhioLINK; Katrina Marshall, Public Library of Cincinnati and Hamilton County

Members: Lily Birkhimer, Ohio History Connection; Stephanie Bricking, Public Library of Cincinnati and Hamilton County; Damon DeBorde, Ohio University Libraries; Marsha Miles, Cleveland State University; Aaron O'Donovan, Columbus Metropolitan Library; Amanda Raab, Rock and Roll Hall of Fame + Museum Library and Archives; Maureen Walsh, The Ohio State University

Working Group Charge

Identify shareable metadata best practices that can serve as models for Ohio; determine if it's possible to adopt or adapt one of those best practices for DPLA Ohio; compare DPLA Metadata API to Ohio metadata best practices/standards; develop and recommend a shareable metadata best practice for the DPLA Ohio project, focusing on the content within the elements; initiate a metadata advocacy/education program; and identify metadata barriers to contribution at the digitization hubs, Ohio Memory, and other major metadata content contributors.

Environmental Scan

Before making recommendations for the DPLA Ohio hub, the Metadata Working Group reviewed six existing metadata models. In this research, we made the following observations:

- **The Minnesota Digital Library Metadata Entry Guidelines** are some of the most up-to-date, having been revised in July 2015. The guidelines are very thorough and have ample examples for each metadata field.
- **Ohio Memory Metadata Best Practices** are more user-friendly than the other models and may be more accessible to a wider audience. They include tips and recommend best practices when there is no metadata to be documented in a given field (ex. to leave a field blank rather than using “unknown” or “N/A”).
- **Mountain West Digital Library Dublin Core Profile** includes Dublin Core and MARC mapping and lists recommended thesauri or vocabularies for each field. These types of examples may be useful when creating Ohio’s model.
- **South Carolina Digital Library Metadata Schema & Guidelines** have similar content to other models, however, the formatting is not as clearly presented.
- **The Florida Digital Action Plan** is brief compared to some of the other models, but includes necessary sections. There are also sections on training, related resources, standards, and best practices and guidelines which would be useful for Ohio’s model.
- **The Pennsylvania Digital Collections Project (PDCP) PA-DPLA Metadata Guidelines** follows the centralized aggregation model and closely aligns with the DPLA recommendations. Mapping recommendations are needed; however, the CONTENTdm mapping included in the PA-DPLA guidelines is not suited for all Ohio institutions.

Because Ohio does not have a set of statewide metadata guidelines currently in place, the Metadata Working Group recommends adapting the Pennsylvania Digital Collections Project (PDCP) PA-DPLA Metadata Guidelines for use in Ohio. See [Proposed Metadata Best Practices](#) and [Proposed Metadata Application Profile](#) sections for more details.

Key Survey Findings on Metadata in Ohio

The October, 2015 Ohio DPLA survey provided insight into the current environment among potential participants in the state. The Metadata Working Group focused on several sections of the survey:

- **Digital Asset Management Systems (DAMS)** - Results from the survey showed that CONTENTdm is the most widely used DAMS, but 25% of those surveyed did not currently have a DAMS. Other digital asset management systems used in Ohio are PastPerfect, Omeka, BePress, D-Space and Hydra.
- **Metadata Issues** - Three separate issues fell into this category:
 - **Schema**—Over 40% of those completing the survey indicated that they were using Dublin Core. Another schema that was indicated as widely used was MARC. Further investigation is needed to determine how respondents are using MARC to describe their digital objects.
 - **Cataloging Standards/Controlled Vocabulary**—Library of Congress Subject Headings and AACR2/RDA were the two most popular replies when asking about cataloging standards and controlled vocabularies. However, over 25% of those who responded had no standards in place, while 17% did not know what standards existed at their organization. Although we can likely assume that some respondents are using local rules for their metadata, more information is needed.
 - **Metadata Harvesting**—A large group of survey participants either did not support metadata harvesting capabilities or did not know if these were currently in place. Of those who do support harvesting, responses included Excel, OAI-PMH, FTP and Z39.50. Surprisingly, only 15% of those who took the survey replied with OAI-PMH, despite CONTENTdm's support of OAI-PMH and its wide use among survey participants. Most likely, the actual number of respondents with OAI-PMH capabilities is higher than what is indicated in the survey.
- **Copyright and Other Intellectual Property Issues** - The survey asked respondents to read several statements on copyright and rate how accurate the statements were in relation to their institutions. For the statement "we record and maintain rights metadata to limit delivery of collections to authorized users," approximately 18% of those responding indicated this was an accurate statement. This area will require outreach to educate project participants on the importance of rights statements and to emphasize DPLA and Europeana's rightsstatements.org project.

After evaluating the results of the October, 2015 DPLA Ohio Survey, it was determined that the Metadata Working Group needed further clarification on several points. Questions were submitted for another survey that was sent to a select group of participants in February 2016.

Analysis of Follow-up Survey:

Q6: What best practices/guidelines is your organization using for the creation of the metadata for digital resources?

79% of those surveyed use locally developed metadata guidelines. The Metadata Working Group's best practices and guidelines will offer guidance in this area.

Q7: Based on what you heard at the December 3, 2015 Ohio DPLA Symposium, do you believe that your metadata is ready for sharing/harvesting into DPLA?

38% of respondents say that their metadata is ready for harvesting into DPLA. For remaining institutions, training may be needed for metadata remediation. Some institutions may be closer to the proposed metadata requirements than they think, and may just need a review of sample existing records. A sample harvest of metadata from a smaller collection may also be an option for metadata review.

Q8: What areas do you believe need revision or modification to your metadata in order for it to be harvestable/sharable?

Almost 60% of respondents indicated that they will need to work on copyright statements. This is an area that will likely require additional training; while Community Engagement Centers cannot offer legal advice regarding copyright, representatives can suggest resources for training and guidance in this area. The DPLA publication on rights states, <http://rightsstatements.org/> project suggested statements may also be recommended.

Q9: Based on what you heard at the Symposium, do you have the staff to prepare your metadata for harvesting?

Almost 66% of respondents replied that they have the staff to prepare the metadata for harvesting, though comments indicated that more information about metadata requirements is needed. A small percentage (3%) do not have the staff to prepare metadata, and approximately 30% of institutions either answered "I don't know" or answered via comment. These comments reflect concerns on the level of work potentially involved in metadata preparation for Ohio DPLA. Further analysis is needed to determine the composition of this 30%.

Q 10: DPLA requests that a thumbnail be provided along with the metadata. Could you provide a thumbnail associated with metadata records that would be harvested for the DPLA Project?

Several of the respondents say that if CONTENTdm is able to assist with the thumbnails that they should be able to provide them. This should be researched/confirmed since the majority of respondents (50%) indicated that they use CONTENTdm.

Primary Challenges and Proposed Solutions

Discussion around metadata challenges took place in group meetings, steering committee meetings and at the DPLA Symposium in December 2015. From those discussions and the Metadata Working Group's research on previously-established hubs, some common challenges emerged. These are listed in order of potential impact, with those issues representing the biggest challenges to the project listed first.

The Metadata Working Group has also identified potential solutions to each of these challenges. Improved communication will be paramount in addressing all of these concerns. In addition to the specific solutions provided below, the Metadata Working Group also recommends the creation of a Frequently Addressed Questions (FAQ) document to address commonly-held concerns - those listed below and those which might arise through the course of the pilot.

Challenge 1: Ensuring the quality and consistency of metadata

The DPLA reviews metadata before mapping, checking for issues that need to be addressed before harvest and ingest. The challenge will be to provide DPLA what they need and omit what they do not. This will require the cooperation of the technical DPLA Hub, the Community Engagement Centers, and the contributing partners.

The DPLA Ohio Hub will need to have mechanisms in place to evaluate metadata coming from partner OAI-PMH feeds and ensure that all partner metadata contains the required fields and that those fields have appropriate values. Further, all metadata coming from a single hub must be applied in the same way. Contributing institutions will need to have the ability to correct metadata that does not comply with the prescribed parameters.

Proposed solutions:

- The DPLA Ohio project must provide metadata best practices and guidelines to participating institutions and they should be presented as clearly as possible. The metadata working group has included best practices and guidelines in this document and the DPLA MAP figures prominently in these recommendations.
- Because of the complexity of metadata issues, compounded by the number of partners in a statewide hub, other DPLA hubs have full- or part-time dedicated metadata specialists to assist (potential) partner institutions. A dedicated individual in a metadata specialist position will be critical throughout the pilot, and therefore should be accounted for in any proposed budget for the pilot phase of DPLA Ohio project. If the pilot is successful, it will be important to review the value of this position and how it might need to change in the post-pilot DPLA Ohio program.
- The proposed Community Engagement Centers should facilitate professional development related to metadata remediation for institutions. A related challenge will be

the varying levels of metadata expertise in libraries in Ohio; that challenge is listed below.

- If the DPLA Ohio pilot project is able to extend harvesting beyond the initial group of participating institutions, step-by-step instructions and other strategies may be necessary for remaining institutions, for both metadata editing/preparation and creation.

Challenge 2: CC0 License for Metadata

Contribution of metadata under a CC0 license - a public domain dedication - may be a challenge for institutions for varied reasons: not understanding what CC0 means in general; not understanding if they “own” the metadata and can therefore “give it away;” being uncomfortable with the idea of others mixing, reusing, mashing up their metadata, whether for commercial gain or otherwise.

Proposed solutions:

- Clear explanation that the CC0 license is for the metadata for the digitized or born-digital item and is separate from the copyright status of the item itself.
- Explaining that, while the technical standards are difficult enough, if DPLA couldn't use metadata under CC0, the entanglements arising from varying rights and attribution requirements would significantly hinder the DPLA's work and progress.
- The language of the license (<http://creativecommons.org/publicdomain/zero/1.0/>) is fairly straightforward; but should we try to rephrase it even clearer so there are no misconceptions?
- Examples of the many institutions that are comfortable with CC0 and perhaps any explanations they offer?

Challenge 3: Required Rights Field

DPLA requires that the metadata includes a rights statement for each item harvested. Ideally, at a minimum, rights statements would be clear and accurate. In the future, they may also need to comply with some standard set forth by DPLA.

Proposed solutions:

- Educational materials (such as the FAQ) should make clear that the Rights field covers the rights for that particular item and that the CC0 license still only refers to the metadata shared with the DPLA.
- To support the quality and accuracy of this field, educational documentation about rights statements should be shared with participating institutions, such as the rightsstatements.org white paper.

Challenge 4: Variety of metadata expertise across libraries in Ohio

The wide variety of libraries participating is one advantage of a single state hub to aggregate metadata for harvest by the DPLA. However, with that variety in library type also comes a wide variance in metadata expertise. There are several challenges associated with this.

The DPLA Ohio is committed to keeping barriers for participation in the hub low. In order to do this, the varying levels of metadata expertise will need to be considered.

The challenges identified here will likely be magnified for those institutions beyond the group of identified as the first participants. While a group of initial participants have been identified, the eventual goal is provide an avenue for participation for all institutions in Ohio. In evaluating the survey results, it appears that issues with metadata schemas, cataloging/encoding standards, and metadata harvesting strategies are quite present for those institutions outside the primary group. In discussion at the December 3rd symposium, lack of time, lack of staffing/expertise, and lack of technical infrastructure were identified by potential institutions as possible roadblocks to DPLA Ohio Service Hub participation.

Institutions/individuals that need more assistance than what our metadata best practices documentation will provide. Pilot participants should be served by the documents and support proposed by this and other Working Groups. However, beyond the pilot, we may need to adjust our approach to provide adequate support.

Proposed solutions:

- As mentioned before, best practices and metadata guidelines need to be clear yet thorough.
- The proposed Community Engagement Centers should facilitate professional development related to metadata remediation for institutions.

Challenge 5: Familiarity/capability related to OAI-PMH

In a follow-up survey of those institutions preliminarily identified to participate in the pilot, 29 provided information on their ability to enable OAI-PMH and 9 of them indicated that they either could not enable OAI-PMH or did not know if they could.

Proposed solutions:

- Ascertain if the lack of use of OAI-PMH is an intentional decision or due to a lack of awareness or ability at the contributing institution.
- Clarify to this primary group what would be needed on their part to take part in/apply this protocol.
- Again, engage the Working Groups and the Community Engagement Centers to help provide professional development on this topic if needed.

Challenge 6: Thumbnail/preview images

Many digital asset management systems can generate and expose along with other metadata URLs to thumbnail images. What could/should we do to support institutions that do not have systems with this capability? A follow up survey of potential initial participants indicates that 65% percent of institutions can provide thumbnails, with additional comments linking this capability directly with availability of thumbnails via CONTENTdm.

Challenge 7: Limited Scope of DPLA Collections

The DPLA defines a specific scope of items that are appropriate for ingest into the DPLA. This scope represents a only a subset of the items available in digital collections in Ohio. For example, only records that point to a freely accessible digital resource/item are within DPLA's scope.

Proposed solutions:

- The best practices and guidelines must be clear that records/metadata pointing to EADs, theses, dissertations, and other items should be excluded from the feed to DPLA. Also, full-text transcriptions and records that point to one part or an individual page of an object should be excluded.
- DPLA Ohio should consider putting forth a collection development policy to further guide institutions in what collections to submit for harvest to the DPLA. In addition to the specifics put forth by DPLA central, this policy would provide criteria that institutions should consider when deciding whether or not to submit a collection for harvest to the DPLA Ohio project. ~~Along the same lines, remind institutions that harvesting can be directed either collection by collection, or for a portion of a collection—they do not need to do a data-dump of all their allowable digital materials into DPLA. This may require more~~

~~of a qualitative evaluation of the content than the previous challenge, which relates more to those materials that don't qualify—basically, is something WORTH adding to DPLA?~~

Commented [1]: This was just me brainstorming--we should definitely wordsmith this a little more so it fits with the tone of the report

Commented [2]: Collection Development Policy.

Commented [3]: Lily, on the call we decided to frame this as DPLA Ohio should provide a collection development policy to help institutions with this. Does this bullet still capture what you meant?

Challenge 8: Future Enhancements and/or Revisions to the DPLA's Metadata Application Profile (MAP)

As with any project that is dependent on an external entity, DPLA Ohio faces a risk that DPLA will change the Metadata Application Profile on which these recommendations are based.

Proposed solutions:

- To mitigate this risk, a standing Metadata Working Group should be established for DPLA Ohio and charged with monitoring metadata issues, changes to the DPLA MAP and best practices related to metadata.

Additional Challenges

The Metadata Working Group has identified other possible challenges to this project that may need to be addressed.

- **Supporting Organizations that desire to go beyond the minimum DPLA-required metadata**
- **Clear Roadmap for Participation** - Given many factors - including the challenges above and the survey responses - all potential partner institutions in Ohio will not be participating in the pilot. It is essential that institutions that do not participate right away have a relatively clear awareness of the process for becoming a participating institution. This will be a long term concern that can possibly be addressed in conjunction with the Education and Advocacy Working Group.

Proposed Metadata Education & Advocacy Activities

The Proposed Challenges section of this document makes evident that communication of metadata requirements and best practices will be as crucial as establishing them. To ensure quality metadata and participation from a variety of cultural heritage institutions, we will need to

make the goals and directives of metadata management, and the DPLA Ohio project as a whole, clear, concise, and actionable.

To ensure clarity and consistency of data, the Metadata Working Group has created a proposed Metadata Application Profile, including best practices, to provide:

- Examples of required and recommended data entries along with element definitions
- Links to required/recommended vocabularies and related DPLA documentation

Documents for workflow should make clear how data should be collected, examined, migrated, and reused, including:

- FAQ/Glossary to explain key concepts and terms including, but not limited to:
 - The difference between required and recommended metadata
 - CC0 License for Metadata (how the data could be used)
 - OAI-PMH protocol (what it means and why it helps collate and manage data from multiple sources)
- Simple point A to point B flow of data from institution to Service Hub to DPLA (both in text and visuals)
- Recommended tools/processes for migrating and cleaning data
- Organizational chart of the collective (including organizing committees and any paid staff) specifying project responsibilities, with contact information

Ideally, DPLA Ohio will want to present this documentation and any training on its implementation in as many formats as possible for the widest possible reach. In the current proposed Governance structure, primary responsibility for these materials and training will rest with the State Library-based Project Manager, and will be coordinated with Community Engagement Centers and DPLA Community Representatives.

- Documentation should be available on the DPLA Ohio site, and as a printable PDF
- Live instructional webinars and/or streaming AV presentations available on the DPLA Ohio site covering metadata requirements and best practices, and data migration workflows
- In-person, on-site training using above the tools

Proposed Metadata Application Profile

Introduction

This MAP details a proposed set of metadata elements, and includes suggestions for best practices, MARC mapping where appropriate, and recommended controlled vocabulary or syntax for fields.

The MAP has Required and Required When Available fields:

Required: Title, Rights

Required When Available: Collection, Language, Type

Remaining fields are Strongly Recommended, Recommended, or Optional.

For aggregation and inclusion in the DPLA, metadata for the Ohio DPLA Project must be contributed with a [Creative Commons CC0 license](#).

Described Resource Elements

The following fields apply to the attributes of the described resource, not to the digital representation of the object in the case where that representation is a digital surrogate.

Required		
DPLA Label	DPLA Property	Usage
Title	dcterms:title	Primary name given to the described resource
Rights	dc:rights	Information about rights held in and over the described resource. Typically, rights information includes a statement about various property rights associated with the described resource, including intellectual property rights

Required When Available		
DPLA Label	DPLA Property	Usage
Collection	dcterms:isPartOf	Collection or aggregation of which described resource is a part
Language	dcterms:language	Language(s) of described resource. Lexvo
Type	dcterms:type	Nature or genre of described resource. DCMIType

Strongly Recommended		
DPLA Label	DPLA Property	Usage
Date	dc:date	Date value as supplied by data provider
Place	dcterms:spatial	Spatial characteristics of described resource, such as a country, city, region, address or other geographical term. Captures aboutness
Subject	dcterms:subject	Topic of described resource

Recommended

DPLA Label	DPLA Property	Usage
Creator	dcterms:creator	Entity primarily responsible for making described resource
Format	dc:format	Physical medium or dimensions of described resource
Publisher	dcterms:publisher	Entity responsible for making the described resource available, typically the publisher of a text

Optional

DPLA Label	DPLA Property	Usage
Alternate Title	dcterms:alternative	Any alternative title of the described resource including abbreviations and translations
Contributor	dcterms:contributor	Entity responsible for making contributions to described resource
Description	dcterms:description	Includes but is not limited to: an abstract, a table of contents, or a free-text account of described resource
Extent	dcterms:extent	Size or duration of described resource
Format	dc:format	Physical medium or dimensions of described resource
Genre	edm:hasType	Captures categories of described resource in a given field. Does not capture aboutness. AAT
Identifier	dcterms:identifier	ID of described resource within a given context
Relation	dc:relation	Related resource
Rights Holder	dcterms:rightsholder	A person or organization owning or managing rights over the resource
Temporal Coverage	dcterms:temporal	Temporal characteristics of the described resource. Captures aboutness

Derived Elements

Attributes of the digital representation of the web resource and aggregation of attributes that apply to the described resource as a whole.

TODO: Break these into Required and Optional?

DPLA Label	DPLA Property	Usage
File Format	dc:format	Web resource format. Internet Media Types
Collection Title	dcterms:title	Name of the collection or aggregation. OAI <SetName>

Commented [4]: TODO - I think we missed this one because its in blue. :) Might be something to address in the next draft.

Collection Description	dcterms:description	Free-text account of aggregation, for example an abstract or content scope note. OAI <SetDescription>
Data Provider	edm:dataProvider	The organization or entity that supplies data to DPLA through a Provider.
Is Shown At	edm:isShownAt	Unambiguous URL reference to digital object in its full information context
Preview	edm:Preview	The URL of a thumbnail, extract or other type of resource representing the digital object for the purposes of providing a preview
Standardized Rights Statement	edm:rights	The value given here should be the rights statement that applies to the digital representation as given (for example) in edm:isShownAt when these resources are not provided with their own edm:rights. This property requires a controlled set of values. URI.

Element Details

Elements are in order alphabetically; see Appendix A for explanations of controlled vocabulary/syntax acronyms. See table below for explanation of properties.

Explanation of

Label	The unique name used in the Metadata Application Profile.
Status	Status labels for elements can include Required, Strongly Recommended, Recommended, and Optional.
Description	Brief description of the element.
Refines	The Dublin Core metadata field of which the described term is a sub-property.
Repeatable	“Yes” means that a field may be used multiple times in an item record; “No” means that a field can only be used once in an item record.

Controlled Vocab/Syntax	Recommended vocabulary for element, if available. (For data quality and consistency)
DC 1.1 Mapping	The Dublin Core element to which the metadata field name maps.
MARC Mapping	The MARC field to which the Dublin Core metadata element is crosswalked.
Notes	Applicable notes for the element.
Example	Examples of best practices for the described field.

DPLA Property	dcterms:alternative
Label	Alternate Title
Status	Optional
Description	Any alternative title of the described resource including abbreviations and translations
Refines	dc:title ; dcterms:title
Repeatable	Yes
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	dc:title
Qualified DC Mapping	dc.title.alternative
MARC Mapping	246, subfield a and b; also, may use 210, 222, 240, 242, 243, and 247
Notes	The distinction between titles and alternative titles is application-specific; some titles are better known by their alternative title.
Example	The book Columbus 400 has five known alternative titles including: Business 400 of Columbus; Men of the Ohio Capital; et al.

DPLA Property	dcterms:isPartOf
Label	Collection
Status	Required when available
Description	Collection or aggregation of which described resource is a part
Refines	dc:relation ; dcterms:relation
Repeatable	Yes
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	dc:relation
Qualified DC Mapping	dc.relation.isPartOf
MARC Mapping	

Notes	The term 'collection' can be applied to any aggregation of physical or digital items. It is typically used to refer to collections of physical items, collections of digital surrogates of physical items, collections of 'born-digital' items and catalogues of such collections.
Example	Ohio Postcard Collection

DPLA Property	dcterms:contributor
Label	Contributor
Status	Optional
Description	Entity responsible for making contributions to described resource
Refines	dc:contributor
Repeatable	Yes
Controlled	LCNAF
Vocab/Syntax	
DC 1.1 Mapping	dc:contributor
Qualified DC Mapping	
MARC Mapping	700, 710, 711, 720 (Added Entry – Personal name, Corporate name, Conference name, Uncontrolled Name).
Notes	Examples of a Contributor include a person, an organization, or a service.
Example	Editor: Charles W. Seward; Photographer: Herb Topy

DPLA Property	dcterms:creator
Label	Creator
Status	Strongly recommended
Description	Entity primarily responsible for making described resource
Refines	dc:creator ; dcterms:contributor
Repeatable	Yes
Controlled	LCNAF
Vocab/Syntax	
DC 1.1 Mapping	dc:creator
Qualified DC Mapping	
MARC Mapping	100 1# (Main Entry- - Personal Name), or 110 2# (Main Entry- - Corporate Name), 111 1# (Main Entry- - Conference Name) or 700/710/711
Notes	Examples of a Creator include a person, an organization, or a service. Use Library of Congress Authorities to guide formatting of names. http://authorities.loc.gov
Example	President Edmund B. Paxton; Columbus Writers Club

DPLA Property	dc:date
Label	Date
Status	Strongly recommended
Description	Date value as supplied by data provider
Refines	
Repeatable	Yes
Controlled Vocab/Syntax	ISO 8601 (W3CDTF)
DC 1.1 Mapping	dc:date
Qualified DC Mapping	
MARC Mapping	260 ## subfield c (Date of publication, distribution, etc.)
Notes	Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme. Date refinements are generally useful in situations where more than one date is needed, and the difference between the dates may be important to users.
Example	1940-02-20; 1940-02; 1940; February 20, 1940

DPLA Property	dcterms:description
Label	Description
Status	Optional
Description	Includes but is not limited to: an abstract, a table of contents, or a free-text account of described resource
Refines	dc:description
Repeatable	Yes
Controlled Vocab/Syntax	
DC 1.1 Mapping	dc:description
Qualified DC Mapping	
MARC Mapping	520, 545, 300, 500, 505
Notes	Since the Description field is a potentially rich source of indexable terms, care should be taken to provide this element when possible. Best practice recommendation for this element is to use full sentences, as description is often used to present information to users to assist in their selection of appropriate resources from a set of search results.
Example	Bill Moss posed with Tina Turner and Eddie Castleberry, the program director of WVKO.

DPLA Property	dcterms:extent
Label	Extent

Status	Optional
Description	Size or duration of described resource
Refines	dc:format
Repeatable	Yes
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	
Qualified DC Mapping	dc:format.extent
MARC Mapping	300 subfield a
Notes	Because the refinement Extent is used in a variety of situations, it generally consists of both a numeric value and a caption that is needed to interpret the numeric value. Best practice is to separate the numeric value and the caption with a space, whether the caption appears before or after the value.
Example	21 minutes; 899 kb; 250 pages; 35 mm

DPLA Property	dc:format
Label	Format
Status	Strongly recommended
Description	Physical medium or dimensions of described resource
Refines	
Repeatable	Yes
Controlled	IMT
Vocab/Syntax	
DC 1.1 Mapping	
Qualified DC Mapping	
MARC Mapping	340; 856 subfield q
Notes	The file format, physical medium, or dimensions of the resource. Recommended best practice is to use a controlled vocabulary.
Example	image/gif; jpeg; drawing; photograph

DPLA Property	edm:hasType
Label	Genre
Status	Optional
Description	Captures categories of described resource in a given field. Does not capture aboutness
Refines	edm:is RelatedTo
Repeatable	
Controlled	AAT ; TGM ; LCGFT
Vocab/Syntax	
DC 1.1 Mapping	
Qualified DC Mapping	

MARC Mapping	655 #7 subfield a (Index Term- - Genre/Form) plus subfield 2=local (for DCMI Type); =gmGPC (for Thesaurus for Graphic Materials); =aat (for Art & Architecture Thesaurus); =lcgft (Library of Congress Genre/Form Terms for Library and Archival Materials).
Notes	If the resource is composed of multiple mixed types then multiple or repeated Type elements should be used to describe the main components.
Example	Poems; Letters; Articles; Maps

DPLA Property	dcterms:identifier
Label	Identifier
Status	Optional
Description	ID of described resource within a given context
Refines	dc:identifier
Repeatable	Yes
Controlled	URI
Vocab/Syntax	
DC 1.1 Mapping	dc:identifier
Qualified DC Mapping	
MARC Mapping	856 40 subfield u (Electronic Location and Access/URI) Only use for mapping the main URI that refers to the resource and begins with "http."
Notes	Recommended best practice is to identify the resource by means of a string conforming to a formal identification such as an established file naming structure.
Example	0385424728 [ISBN]; Westerville_00099

DPLA Property	dcterms:language
Label	Language
Status	Required when available
Description	Language(s) of described resource
Refines	dc:language
Repeatable	
Controlled	ISO 639-2 ; Lexvo
Vocab/Syntax	
DC 1.1 Mapping	dc:language
Qualified DC Mapping	
MARC Mapping	041 0# subfield a (language code); 008/35- 37
Notes	Either a coded value or text string can be represented here. If the content is in more than one language, the element may be

	repeated. Recommended best practice is to use a controlled vocabulary.
Example	en; en-US; eng; English

DPLA Property	dcterms:spatial
Label	Place
Status	Strongly recommended
Description	Spatial characteristics of described resource, such as a country, city, region, address or other geographical term. Captures aboutness
Refines	dc:coverage ; dcterms:coverage
Repeatable	Yes
Controlled	LCSH ; TGN ; USGS
Vocab/Syntax	
DC 1.1 Mapping	dc:coverage
Qualified DC Mapping	dc:coverage.spatial
MARC Mapping	651 #0 (for LCSH place names) or #7 with the specific vocabulary source provided in subfield 2 (TGN, for example)
Notes	Spatial topic and spatial applicability may be a named place or a location specified by its geographic coordinates. A jurisdiction may be a named administrative entity or a geographic place to which the resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names [TGN]. Where appropriate, named places can be used in preference to numeric identifiers such as sets of coordinates or date ranges.
Example	Ohio; Columbus, Ohio; Downtown; 96 S Grant Avenue

DPLA Property	dcterms:publisher
Label	Publisher
Status	Strongly recommended
Description	Entity responsible for making the described resource available, typically the publisher of a text
Refines	dc:publisher
Repeatable	Yes
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	dc:publisher
Qualified DC Mapping	
MARC Mapping	260 subfield b (if born digital) or 533 subfield c (if reformatted)
Notes	The intent of specifying this field is to identify the entity that provides access to the resource. If the Creator and Publisher are the same,

	do not repeat the name in the Publisher area. If the nature of the responsibility is ambiguous, the recommended practice is to use Publisher for organizations, and Creator for individuals. In cases of ambiguous responsibility, use Contributor.
Example	The Ohio Sentinel Publishing Company; F.J. Heer Publishing Co.

DPLA Property	dc:relation
Label	Relation
Status	Optional
Description	Related resource
Refines	
Repeatable	Yes
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	dc:relation
Qualified DC Mapping	
MARC Mapping	530
Notes	Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system. Relationships may be expressed reciprocally (if the resources on both ends of the relationship are being described) or in one direction only, even when there is a refinement available to allow reciprocity. If text strings are used instead of identifying numbers, the reference should be appropriately specific. For instance, a formal bibliographic citation might be used to point users to a particular resource.
Example	If an item is related or a part of another resource; for example a picture being scanned from a book than a bibliographic citation would be given to the book. Ex:Columbus Men Worthwhile: Presented in Cartoon, by Billy Ireland, OH 741.5 I652c, (p.219)

DPLA Property	dc:rights
Label	Rights
Status	Required
Description	Information about rights held in and over the described resource. Typically, rights information includes a statement about various property rights associated with the described resource, including intellectual property rights
Refines	
Repeatable	Yes
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	dc:rights

Qualified DC Mapping	
MARC Mapping	506, 540
Notes	The Rights element may be used for either a textual statement or a URL pointing to a rights statement, or a combination, when a brief statement and a more lengthy one are available.
Example	A user of any image in this collection is solely responsible for determining any rights or restrictions associated with the use, obtaining permission from the rights holder when required, and paying fees necessary for a proposed use.; http://www.photohio.org/columbus-citizen-journal-and-columbus-ci

DPLA Property	dcterms:rightsholder
Label	Rights Holder
Status	Optional
Description	A person or organization owning or managing rights over the resource
Refines	
Repeatable	
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	
Qualified DC Mapping	
MARC Mapping	542 subfield d
Notes	Since, for the most part, people and organizations are not typically assigned URIs, a person or organization holding rights over a resource would be named using a text string. People and organizations sometimes have websites, but URLs for these are not generally appropriate for use in this context, since they are not clearly identifying the person or organization, but rather the location of a website about them.
Example	Herb Topy; The Ohio State University

DPLA Property	dcterms:subject
Label	Subject
Status	Strongly recommended
Description	Topic of described resource
Refines	dc:subject
Repeatable	Yes
Controlled	LCSH ; AAT
Vocab/Syntax	
DC 1.1 Mapping	dc:subject

Qualified DC Mapping	
MARC Mapping	650, 600, 651, 610, 653
Notes	Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme. Select subject keywords from the Title or Description information, or from within a text resource. If the subject of the item is a person or an organization, use the same form of the name as you would if the person or organization were a Creator or Contributor.
Example	Dogs; Airplanes; Rhodes, James A. (James Allen), 1909-2001

DPLA Property	dcterms:temporal
Label	Temporal Coverage
Status	Optional
Description	Temporal characteristics of the described resource. Captures aboutness
Refines	dc:coverage ; dcterms:coverage
Repeatable	Yes
Controlled	LCSH
Vocab/Syntax	
DC 1.1 Mapping	dc:coverage
Qualified DC Mapping	dc.coverage.temporal
MARC Mapping	033 subfield a; 533 subfield b
Notes	Recommended best practice is to select a value from a controlled vocabulary. Where appropriate time periods should be used in preference to numeric identifiers such as sets of date ranges.
Example	19th Century; 1850s

DPLA Property	dcterms:title
Label	Title
Status	Required
Description	Primary name given to the described resource
Refines	dc:title
Repeatable	No
Controlled	
Vocab/Syntax	
DC 1.1 Mapping	dc:title
Qualified DC Mapping	
MARC Mapping	245 subfields a and b
Notes	Typically, a Title will be a name by which the resource is formally known. If in doubt about what constitutes the title, use the Alternate Title element and include the variants in second and subsequent Alternate Title iterations.

Example	Handsome Homes of Columbus, Ohio; Historical Collections of Ohio
---------	--

DPLA Property	dcterms:type
Label	Type
Status	Required when available
Description	Nature or genre of described resource
Refines	dc:type
Repeatable	Yes
Controlled	DCMI Type
Vocab/Syntax	
DC 1.1 Mapping	dc:type
Qualified DC Mapping	
MARC Mapping	655 #7 subfield a (Index Term- - Genre/Form) plus subfield 2=local (for DCMI Type)
Notes	Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary [DCMITYPE]. To describe the file format, physical medium, or dimensions of the resource, use the Format element.
Example	Image; Sound; Text

Appendix A: Controlled Vocabularies and Syntax

The table below provides links to the Controlled Vocabularies or Syntax prescribed by the Proposed Metadata Application Profile section.

Element Label	Controlled Vocabulary/Syntax	Link
Contributor & Creator	LCNAF - Library of Congress Name Authority File	http://authorities.loc.gov/
Date	ISO 8601	http://www.iso.org/iso/home/standards/iso8601.htm
Format	IMT - Internet Media Types	http://www.iana.org/assignments/media-types/media-types.xhtml
Genre	AAT - Art and Architecture Thesaurus	http://www.getty.edu/research/tools/vocabularies/aat/
Genre	LCGFT - Library of Congress Genre/Form Terms	http://id.loc.gov/authorities/genreForms.html ; https://www.loc.gov/aba/publications/FreeLCSH/GENRE.pdf
Genre	TGM - Library of Congress Thesaurus for Graphic Materials	http://www.loc.gov/pictures/collection/tgm/
Identifier	URI	https://tools.ietf.org/html/rfc3986
Language	ISO 639-2	http://www.loc.gov/standards/iso639-2/php/code_list.php
Language	Lexvo	http://www.lexvo.org/
Place	LCSH - Library of Congress Subject Headings	http://authorities.loc.gov/
Place	TGN - Thesaurus of Geographic Names	http://www.getty.edu/research/tools/vocabularies/tgn/
Place	USGS- U.S. Geological Survey	http://www.usgs.gov/science/tab-term.html
Subject	LCSH - Library of Congress	http://authorities.loc.gov/

	Subject Headings	
Subject	AAT - Art and Architecture Thesaurus	http://www.getty.edu/research/tools/vocabularies/aat/
Temporal Coverage	LCSH - Library of Congress Subject Headings	http://authorities.loc.gov/
Type	DCMI Type	http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/

Appendix B: Metadata Best Practices Resources

The Metadata Working Group originally planned to include a Best Practices section in this report, but realized that the project is at too early a stage at this point to make these recommendations. Instead, the Metadata Working Group recommends that a standing Metadata Working Group take on this document once the project is underway. However, in order to help guide them in this work, we include our draft materials in this appendix.

Use DPLA docs like one on Temporal and Geographic metadata as reference

<https://docs.google.com/document/d/1lfiJ8yoZf1fAoR5vmJoHpWQO63eKeL8HDGVupCocfoM/edit>

Beginning DRAFT Text

Intro to DPLA Metadata Model: http://dp.la/info/wp-content/uploads/2015/03/Intro_to_DPLA_metadata_model.pdf

In researching recommendations from the DPLA and other hubs, a commonly cited challenge is consistent quality metadata from contributing institutions. The DPLA Ohio project aims to keep barriers for participation low, so collections are required to have very few fields in order to participate, though other fields may be provided for harvest. In all cases, the quality of a collection's metadata at the local level will be reflected in the resulting record at the DPLA. Participating in the DPLA Ohio project has the potential to increase visibility to an institution's collections - and any metadata problems that exist therein.

Therefore, the Working Group charged with Metadata issues for the pilot should establish and communicate best practices for creating metadata, as well as best practices for editing existing, problematic metadata.

These best practices may include, but need not be limited to:

- Use the notes and examples in the DPLA Ohio Metadata Application Profile to guide metadata creation - each element table contains recommended controlled vocabularies or standards, as well as examples.
- From the DPLA Recommendations for Creating Geographic Metadata in this document: <https://docs.google.com/document/d/1lfiJ8yoZf1fAoR5vmJoHpWQO63eKeL8HDGVupCocfoM/edit> "The best way to ensure a correct match is to use unambiguous data. A GeoNames URI is great if you have the ability to add them to your data. Geospatial coordinates are also useful. If you are unable to supply these, consistently using place names in a hierarchy (Erie, Pennsylvania, or even Erie, Erie County, Pennsylvania, United States) will result in better matches."
 - Latitudes and Longitudes should be expressed in whole numbers and decimals (not degree- minutes)
 - If contained in the same element, latitude and longitude should be given in the following order: "latitude, longitude"

- Place names in a hierarchy should be in the same element. For example two dc.coverage elements, one with Erie and one with Pennsylvania, will result in problematic matching on the DPLA side. Instead, use just one dc.coverage element, with the value "Erie, Pennsylvania"