

FAQ for MODS Guidelines Implementers

FAQ for MODS Guidelines Implementers

This FAQ is designed to help participants in the DLF Aquifer initiative implement the project's [MODS Guidelines](#). It is arranged into several sections:

[General questions](#)

[Tools](#)

[Specific comments on MODS top-level elements](#)

General questions

Q: Where can I get help implementing the Aquifer MODS Guidelines?

A: Any member of the [Aquifer Metadata Working Group](#) would be glad to talk with you.

Q: How can I share my metadata and/or collections with DLF Aquifer?

A: The best place to start is to consult the DLF Aquifer [Collections Submission Documents](#).

Q: There are some features of the Aquifer MODS Guidelines that don't fit well with my local needs. Are you suggesting I use these guidelines locally?

A: No. The needs of a shared environment are very different from a local environment. The Aquifer MODS Guidelines are intended to provide guidance on what metadata you share with the Aquifer project should look like; we don't intend, and in fact realize it is impossible, to specify how your local system should function.

Q: These guidelines seem very library-centered. Can I use them to describe archival materials or museum objects?

A: While MODS itself isn't explicitly for materials held in libraries rather than archives or museums, it is a *bibliographic* metadata element set, and therefore has a bit of that inherent library flavor to it. The *Guidelines* explicitly permit usage of content standards from other communities, such as Describing Archives: A Content Standard, and Cataloging Cultural Objects. It is the hope of the Aquifer Metadata Working Group that MODS records can be produced according the spirit of these guidelines for materials held in any type of cultural heritage institution.

Q: It sure will be a lot of work to make my metadata meet these guidelines. Do I really need to?

A: The Guidelines were written intentionally to set a high standard for Aquifer metadata, so that the Aquifer project can provide better services on that metadata. We figured if anybody could produce higher-quality metadata, it would be DLF institutions. But we also realize that producing metadata for sharing is difficult. The Aquifer Metadata Working Group has therefore produced a "[Levels of Adoption](#)" document to help you prioritize work on your shared metadata based on what Aquifer can accomplish with that work. Also, stay tuned to the [Public Metadata Documents](#) page for tools we'll be releasing to help you assess and improve your metadata.

Q: Why are properties of both an analog original and a digitized copy of a resource mixed together in a single record in these guidelines? Why didn't you follow the lead of FRBR or the Dublin Core 1:1 principle and separate the two?

A: This topic is an extremely complex one, and one that generated a great deal of discussion within the Aquifer Metadata Working Group and those from whom we solicited feedback. The issue is more complicated even than separating metadata about multiple versions; it also involves separating metadata about the content from that of the carrier, a distinction which is not always clear even to experts. An earlier draft of the Guidelines attempted to separate metadata about content from that of its various carriers, but in the end we determined that requiring this separation for all Aquifer participants was too high of a barrier for participation, that it provided very little benefit for the sorts of discovery feasible for a metadata aggregation, and that these distinctions were not very clear for one-of-a-kind materials (rather than published materials). We encourage implementers to implement clear conceptual models underlying their local metadata repositories, and to perform careful mapping of that metadata into MODS records for sharing with Aquifer.

Principles of shareable metadata

Q: Why can't I just share the versions of records I store locally?

A: Your local metadata is likely to be optimized for your delivery system (or suffer from some of its quirks), and leave out information implied by your institutional context. Shared records therefore need to be created that are intended to be used *outside* of that context. For more information, see Shreeves, Sarah L., Jenn Riley, and Liz Milewicz. "[Moving towards shareable metadata](#)." *First Monday* 11, no. 8 (7 August 2006).

Institution as publisher

Q: Since my institution is responsible for digitizing and making a resource available, I want to give us credit for that by making us the publisher of the digital version. Where can I do that in Aquifer MODS?

A: In MODS, the <publisher> element is intended to convey information about the publisher or originator of the original resource that you digitized, which is ultimately the most useful publication-related information for users of your digital resource. If you want to, you can certainly toot your institution's horn about providing the digital facsimile, and the MODS <note> element is the appropriate place to do so.

Mapping issues

Tools

MARC21 to MODS stylesheet for Aquifer

Q: Are there particular MARC tags or subfields that we may want to change the mapping for, if our local cataloging practices differ from what the stylesheet expects?

A: We can't predict all places where you might need to change the mapping, but some we've encountered in our testing that stand out are:

*752 Hierarchical Place Name - we are mapping to <subject><hierarchicalGeographic>, but if your records have a placename that's not a subject (e.g. place of publication) in that tag, the subject mapping would not be appropriate.

*533 Reproduction, subfield f Series - we are mapping to <relatedItem type="series">. If your local practice is to always repeat the series in an 830 tag, that mapping will be redundant and could be deleted.

*856 mapping based on indicator values - we are mapping 856 with 2nd indicator 1 (Version of resource) to <location><url> for the resource. Many records that use the "single record" approach to describe both physical and digital manifestations in one record use this indicator for the URL. But sometimes that indicator is used for a URL to only a small portion of a resource (e.g. table of contents). If that has been your practice, you may want to change the mapping to <relatedItem type="constituent"> or some other mapping, or delete mapping of 856 with 2nd indicator 1 in the stylesheet.

Q: How does the stylesheet determine which 856 field should get a usage="primaryDisplay" attribute when mapped to MODS <location><url>? The guidelines say there should be one and only one with that attribute.

A: We have a protocol based on the 856 second indicator. If there's an 856 with second indicator 0 (Resource), it gets the "primaryDisplay". If no indicator 0 is present but an 856 has indicator 1 (Version of Resource), it gets "primaryDisplay". If neither 0 nor 1 indicators are present, an 856 with indicator "blank" (no information) gets the attribute. If there are multiple 856's with the indicator that's highest in protocol, the first one appearing in the record gets the attribute. 856 with indicator 2 (Related resource) is not mapped for Aquifer MODS - as an aggregator, we're focusing on the resources described. If this protocol would result in inappropriate assignment of the attribute for your records, you may want to change the mapping.

Specific comments on MODS top-level elements

titleInfo

Q: What's your general approach to titleInfo?

A: Because titles are such an important metadata element, these guidelines require the use of at least one <titleInfo><title> element. Choice and format of titles should be based on a content standard such as AACR2 or CCO.

Q: What if there is no title on an item?

A: You must supply one, but do not enter it in square brackets or use other punctuation to indicate that fact. Instead, use the displayLabel attribute to indicate that the title is supplied. Please consider using a content standard such as AACR2 to give you guidance about how to create consistent supplied titles.

Q: Does the entire title string go in the <title> element?

A: Not always. There are more specific subelements such as <subTitle>, <partName>, and <partNumber> which should be used if applicable.

name

Q: What's your general approach to name?

A: Because the name of the author or creator of the intellectual content of a resource is necessary for a minimal citation, the guidelines recommend the use of at least one <name> element for this purpose, if applicable. Use of the name element is one of the minimum requirements for participation in Aquifer when it is applicable and available.

Q: What if the author or creator of a resource is unknown or anonymous?

A: Do not use the name element in this case.

Q: How should names appear in this element?

A: All names should appear in the <namePart> sub-element. Either the entire name can appear in one <namePart> element or each part of the name can be wrapped in separate <namePart> elements. These guideline prefer the latter. Name constituents should appear in the same same order as their authorized form, or, if no authority is used, in last name, comma, first name form.

typeOfResource

Q: What's your general approach to typeOfResource?

A: We require the use of at least one <typeOfResource> element, in order to give aggregators the option to display to end users high-level information regarding the nature of the original resources being described. This element describes the original item, whether an analog format that has been digitized, or a born-digital file. The element can be repeated, and it can also possess attributes specifying that the record represents a collection, or represents manuscript content.

genre

Q: Why should I include <genre> in my institution's shareable metadata records?

A: The concept of genre is a very useful one in helping users navigate through increasingly complex information spaces. Aggregators of digital content find it an especially useful option in building interfaces that allow users to explore large retrieval sets when they do searches, or as a facet in providing a browsing capability. Genre terms applied to digital resources are most useful when they are selected from a standardized thesaurus or controlled vocabulary (e.g., [Library of Congress Subject Headings](#), [Art and Architecture Thesaurus](#), [Thesaurus of Graphic Materials II: Genre and Physical Characteristic Terms](#)). Even broadly relevant genre terms that appear in most thesauri — such as "Books," "Photographs," or "Sound recordings" — are more useful to aggregators and users than no genre terms at all. When a group of digital resources has similar genre characteristics, but has metadata containing no genre terms, it may be possible to supply an appropriate term or two as part of the process of extracting shareable metadata records from your local content management system.

originInfo

Q: What's your general approach to originInfo?

A: <originInfo> data can apply to many different versions of a resource; the DLF Aquifer MODS Guidelines are flexible in that they allow <originInfo> data to be present describing any version of a resource. They strongly recommend to be prudent in what <originInfo> data to include, however, by only including data likely to be useful in an aggregated environment. Dates are considered extremely important for end-user access, but refrain from including dates that can be considered technical metadata, for example, the date an analog resource was digitized. The use of the keyDate attribute on one and only one date is useful to tell the aggregator which date to use for indexing and sorting.

Q: Why does Aquifer require a date in an <originInfo><date> element as a minimum for participation, "if known"? What does "if known" mean in this context?

A: Dates are important to users of an aggregated collection such as Aquifer in several ways. If a resource contains a date of creation, publication or issuance, we require that the date be present in the metadata, because a date can be essential in allowing a user determine what version of a work is being presented, and it can help them place the work in context temporally.

Dates and date ranges can also be used to narrow, qualify, or sort search results; this can give users researching historical or temporal aspects of a topic a valuable tool. This feature will be only partially helpful, however, if some records in the collection don't contain a date; those records would be "missed" in any search that qualifies by a date range.

We recognize that undated materials may be described in metadata contributed to Aquifer, and that a particular content standard may not allow for supplying approximate, inferred or questionable dates for these resources. We will accept records without dates under those conditions. However, we would prefer a supplied date or date range (years, decade, even century), if you can possibly provide one that reasonably represents the time period for the content, over having no date in the <originInfo> element at all. Best practice would be to give the qualifier attribute value (approximate, inferred or questionable) for such dates.

Q: Does using the W3CDTF date format (YYYY-MM-DD) require that I include a month and day? What if I only know the year?

A: W3CDTF allows leaving off the more granular parts of the date. 1876 and 1876-02 (to indicate February 1876, but no day known) are both legal W3CDTF values.

Q: Can I include the date my institution digitized or claimed copyright on a digital surrogate of an item?

A: While this information is important for you to track locally, it is of minimal utility in a shared environment, and may actually get in the way of date indexing by Aquifer. It is best to omit this information from shared records.

language

Q: When is language required in a record?

A: <language> is required for resources that cannot be properly understood without language. Examples range from written correspondence or an audio recording of a speech. If language is vital to making sense of the resource, include it in the record.

Q: Besides the fact that it's required, why should I include language in the record?

A: Language is a powerful way to narrow search results. It allows a user to exclude materials in languages he/she doesn't understand.

physicalDescription

abstract

Q: Do I really need an abstract for every resource?

A: We recommend including a summary or abstract for most resources. Research has shown that users focus attention on this area of a metadata record, and besides being helpful in evaluating usefulness of a resource, indexing keywords in the abstract makes the resource more "findable". This is particularly helpful when no table of contents is available, and when the title is not very descriptive.

tableOfContents

Q: Why is Table of Contents recommended?

A: A table of contents, whether it is transcribed from a table within in a resource or derived from chapter or section titles, can give users a quick overview of what the resource contains and how it is organized, and keyword indexing of the <tableOfContents> element often makes the content more "findable" to users.

Q: What should or shouldn't be included in <tableOfContents>?

A: Page numbering is usually omitted in a metadata record; volume or section numbering along with the section titles helps show the organization, but if all you have is numbering, the TOC should be omitted. Linking to an online table of contents using the xlink attribute may be helpful for user resource evaluation, but not for keyword indexing. The displayLabel attribute may be helpful for contents that are partial or differ from standard text contents tables.

targetAudience

Q: What's your general approach to targetAudience?

A: Some resources may be intended for audiences that are at a certain intellectual level or share certain intellectual interests. <targetAudience> should contain this information, if it is available.

Q: Should targetAudience be used for MPAA or RIAA audience ratings?

A: No. Use <accessCondition> with "type" attribute value "restrictionOnAccess".

Q: What if my resources don't have a defined target audience, they're for everybody? Do I need to include this element with a value like "general"?

A: We don't recommend including boilerplate values for elements like this. It's best to just omit the targetAudience element if there's no specific audience that the resource is designed for.

note

Q: When should I use a <note> in my records?

A: Use <note> in MODS records only when you have information that cannot be conveyed in another, more specific, MODS element. If it is possible to map information from your local content management system into a more specific MODS element than <note>, you should do so in order to insure that your data content is labeled and presented correctly and, consequently, is as useful as possible to users.

Q: I'm mapping information from MARC records into MODS so that we can make these records harvestable. It would be so much easier for me just to map all the MARC 5XX fields into MODS <note> elements and be done with it.

A: It may seem easier, but aggregators who harvest your shareable metadata records can do much more with them in terms of general presentation, labeling, and indexing if you take the time to map your MARC 5XX notes correctly into MODS. For example, a MARC 506 field conveys information to end users about conditions placed on access to a resource. This information could be generically treated if you map it to MODS <note>, but can be appropriately highlighted if you map it correctly to MODS <accessCondition>. The same is true of information in the MARC 520, or Summary field. It can be much more appropriately handled by aggregators and understood by users if correctly mapped to MODS <abstract> rather than generically treated as a <note>.

subject

Q: How do I declare a subject vocabulary?

A: Use the authority="xxx" attribute for the <subject> in question.

Q: What if the subject is a local term?

A: Hopefully you maintain a controlled list to manage it. In this case, use authority="local" as an attribute for the subject term. If you do not control the term in any way, do not include an authority attribute.

Q: You say that <subject> is required. Is it OK to use a genre term (i.e. sculpture)?

A: <subject> is required only if applicable. Certain resources, especially those that are art or realia, may not be said to have any subject at all. In these cases, use the appropriate genre term under the top-level <genre> element instead of here in <subject>.

classification

Q: How is classification used in the guidelines?

A: Aggregators, including Aquifer, typically do not use <classification> for search and browse functionality. Because of this, <classification> is optional in the guidelines. If you choose to include <classification>, however, the guidelines recommend that the values conform to authorities referenced in the Source Codes for Classification maintained by the Library of Congress at <http://www.loc.gov/marc/sourcecode/classification>.

relatedItem

Q: What's your general approach to relatedItem?

A: Aggregators, including the Aquifer initiative, are unlikely to follow large numbers of links to related records or parse deeply recursive MODS records effectively. Instead, they are likely to assume one shared record should result in one record presented in a search result set. With this expectation in mind, the DLF Aquifer MODS Guidelines recommend the use of <relatedItem> in only three cases. The first is "to point to a full metadata record for a related item," where the <relatedItem> element would include an xlink:href attribute pointing to a related record. In this case, the aggregator is likely to display the link to a user but not to retrieve and index the record. The second and third cases involve including MODS elements within <relatedItem> rather than relying on a link: "to provide contextual information useful for full description of the resource" and "to provide additional information about intellectual constituent units of the resource being described." In the former case, the data is likely to describe a resource at a higher hierarchical level, such as a collection or series, and in the latter case, the data is likely to describe analytical components of a resource. In these cases, the intention should be that the aggregator include the data within <relatedItem> in a search index.

identifier

Q: How is <identifier> used in these guidelines?

A: <identifier> can be used to identify either the digital version or the analog original. Only universal identifiers are recommended, so local call numbers should be avoided. A type attribute must be used to identify the type of identifier.

Q: How does <identifier type="URI"> differ from <location><url>?

A: When a URL is used in <identifier type="URI">, it should link to the resource in its repository context. The link in <location><url> is a link to the resource with its contextual material.

location

Q: Why do the guidelines require one <location><url> to have the attribute usage="primary display"?

A: Indicating which url is the primary url for users to access allows service providers to use this url as the link in the brief display. Typically, the title of a resource is converted into a link. If there are multiple urls provided without an indication of which is primary, the service provider may just choose the first listed.

Q: The guidelines indicate that the primary display url should point to the resource in context. Why shouldn't I just provide a link directly to the resource itself?

A: The best practice of the URL pointing to the resource in context is important because many service providers do not display the full metadata record harvested from a data provider. Thus, if the primary link to a resource is to a stand-alone version of the resource (such as a JPG image only), an end-user will have no context except for the metadata on the service provider's site. This does not serve the end-user well, nor does it serve the data provider well as the end-user cannot easily navigate to other parts of the data provider's collection. At a minimum, the URL should point to a page that contains the resource and a navigation bar that allows users to reach the collection homepage.

accessCondition

Q: Why do the guidelines require an <accessCondition> element?

A: Including information about access rights and restrictions is essential to promote the widest possible use of Aquifer resources. The Aquifer Metadata Working Group decided that it is important to state as explicitly as possible the access conditions for each item. We also want to promote the use of standard license agreements such as Creative Commons.

part

extension

Q: What's your general approach to the <extension> element?

A: The guidelines do not recommend the use of the <extension> element except in where there is well-documented and community driven information to place there. The <extension> element is meant to hold information - generally local information - that the other elements in MODS cannot accommodate. Usually this information will not be meaningful or useful to a service provider because it cannot be interpreted. The exception is the use of well documented community-based information for which there is not another appropriate place within the MODS schema (for example, the Asset Action package).

recordInfo

Q: What's your general approach to recordInfo?

A: <recordInfo> is required by these guidelines because even though it contains information that end users will not find useful, it provides aggregators with the information they need to use metadata records appropriately. These guidelines require the use of at least one <languageOfCataloging> sub-element containing two <languageTerm> sub-elements: one containing the "text" value for the "type" attribute, another with the "code" value for the "type" attribute. Information about how a MODS record was generated, and by what rules, is useful to aggregators, and if available, should appear in <recordOrigin>.

Q: Do you require any particular language authority?

A: ISO639-2b <http://www.loc.gov/standards/iso639-2/faq.html>