

# Interoperability, OAI-PMH, and Shareable Metadata

## METRO Workshop December 8, 2006

Sarah L. Shreeves ([sshreeve@uiuc.edu](mailto:sshreeve@uiuc.edu)), Coordinator, IDEALS  
University of Illinois at Urbana-Champaign

### Agenda

1. Introduction & Preview of Agenda
  - Purpose & motivation for this workshop
  - Objectives
  - Introductions
2. Why Share?
  - Benefits of sharing
  - Search interoperability frameworks
  - Examples of federated search and aggregated search
3. The Plumbing
  - What is OAI?
  - History of the protocol
  - Basic Concepts & Definitions
  - OAI verbs
  - Examples
4. OAI is easy; Metadata is hard
  - What is shareable metadata
  - How do you create shareable metadata?
5. Implementation options for OAI data providers
  - Implementation options and requirements
  - Organizational issues
  - “Commercial” turnkey solutions
  - Resources for Data Providers
6. Communication and Documentation
  - Talking to your aggregator(s) and vice versa
  - Documenting metadata practices

## **Course Description and Objectives:**

### **Description:**

The Open Archives Initiative Protocol for Metadata Harvesting is now in use by numerous communities to expose and allow aggregation of metadata describing a wide range of collections. Because of protocols like OAI, libraries, archives, and museums need to understand the impact – both positive and negative – that their metadata authoring practices can have within aggregations and networks, as well as what they can do to increase the usefulness of their metadata for service providers and users. This workshop will describe the OAI protocol (what it is, why it exists, and how it works), discuss options for exposing metadata via the OAI protocol, and describe how to make metadata more shareable for use within the OAI protocol or other sharing mechanisms.

### **Objectives:**

The participants in the workshop should come away with:

- An understanding of the importance of interoperability protocols like OAI-PMH;
- A basic understanding of how the OAI protocol works;
- The knowledge necessary to decide whether to become an OAI data provider and what options are available to do so;
- An understanding of the need for interoperable or shareable metadata;
- An understanding of the key components of shareable metadata; and
- The ability to think critically about the shareability of their own metadata.

### **Presenter:**

Sarah L. Shreeves is the coordinator for the institutional repository, IDEALS, at the University of Illinois at Urbana-Champaign (UIUC). She has worked extensively with the OAI protocol since 2001 and is leading the effort to establish best practices for OAI data provider implementations and for shareable metadata. Sarah was the past chair of a Digital Library Federation working group that is establishing guidelines on the use of the Metadata Object Description Standard (MODS) within interoperability protocols like OAI. She received a B.A. in Medieval Studies from Bryn Mawr College, an M.A. in Children's Literature from Simmons College, and an M.S. in Library and Information Science from UIUC.

Sarah can be reached at:

Sarah Shreeves ([sshreeve@uiuc.edu](mailto:sshreeve@uiuc.edu))  
Coordinator, IDEALS  
Grainger Engineering Library and Information Center  
University of Illinois at Urbana-Champaign  
1301 W. Springfield Avenue  
Urbana, IL 61801  
(217)244-3877 or (217)333-4648

## **What Is OAI-PMH?**

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a tool that allows metadata about resources to be aggregated from distributed and heterogeneous collections. Originally developed to facilitate interoperability among e-print archives, OAI-PMH is now in use by numerous communities to expose and allow aggregation of metadata describing a wide range of collections.

## **What Are The Benefits?**

OAI offers a way to unlock the hidden web of digital resources. OAI-PMH strikes a balance between institutional control and centralization. Metadata providers maintain ultimate control over their metadata and their content, while benefiting simultaneously from access to the aggregated metadata.

### **Access**

OAI offers cost-effective item-level access to digital resources through a single discovery mechanism. In most cases currently, item level resources of digital (or print) collections can only be found through the hosting institution's web site. A database of metadata harvested from multiple digital collections can bring us closer to specialized aggregations.

### **Awareness**

The more ways that your collection can be accessed, the higher awareness that end users, colleagues, and funding agencies will have of your collection. Making your digital collection accessible through the OAI Protocol will highlight the amount and variety of digital resources available.

### **Experience**

The OAI Protocol is fast becoming an essential tool in the development of digital libraries. The National Science Digital Library (NSDL) and the European Open Archives Forum, as well as Institute of Museum and Library Services, has chosen OAI-PMH as a strategic tool for uniting digital collections. Institutions with implementations of and experience with OAI data provider services are well positioned to be part of such initiatives.

## How OAI-PMH Basically Works (in a single wordy page)

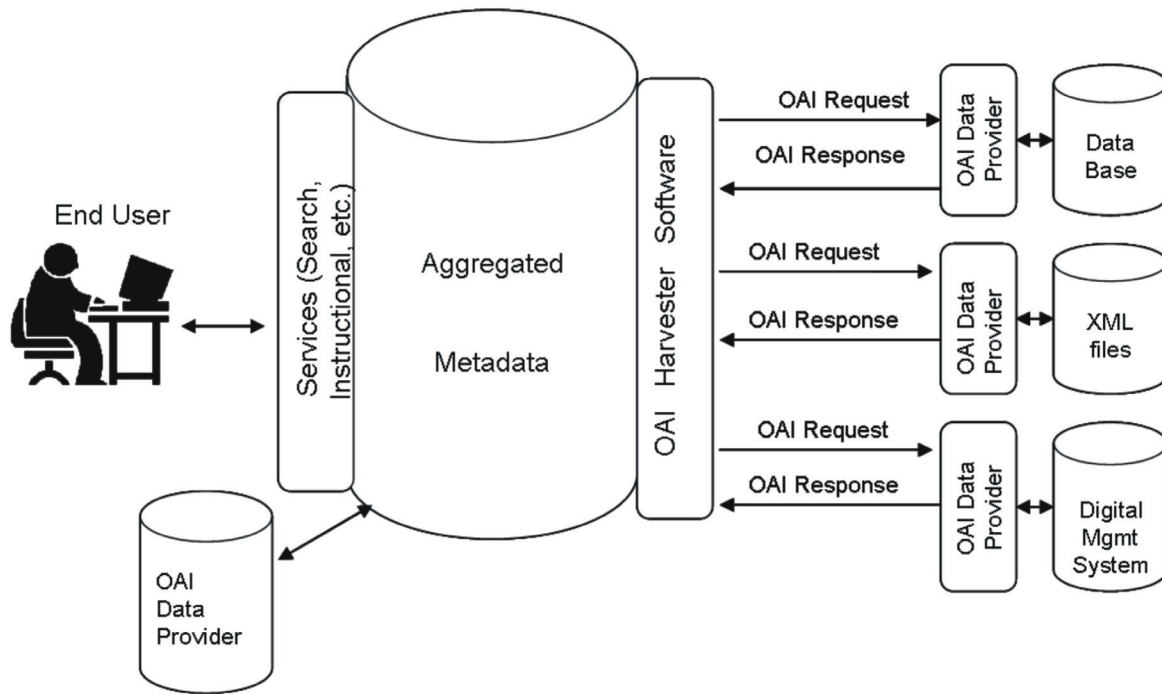
The OAI PMH builds on other standards, notably HTTP, XML for encoding and exchanging information, and the Dublin Core schema for metadata semantics. Service providers may request information from data providers using a standard set of six OAI-PMH verbs. OAI-PMH requests and responses are transmitted via HTTP (which means that you can send OAI requests and receive responses on your web browser). Almost all of the OAI-PMH verbs require and/or allow the use of certain parameters that further define the exact nature and details of the request. Data providers process OAI-PMH requests received and reply with appropriate OAI-PMH responses, which are always in the form of valid XML (including the metadata records themselves) which conform to top-level XML schemas defined by the OAI protocol. An XML schema is a way to define the structure and semantics of XML documents. The service provider can learn who the metadata provider is ('Identify' request), what metadata formats it supports ('ListMetadataFormats' request), and how it has divided its metadata ('ListSets' request). The service provider can also request the metadata itself ('GetRecord', 'ListIdentifiers', 'ListRecords' requests).

The OAI protocol is itself metadata neutral. That is, any (and multiple) metadata schema can be used as long as it has a supporting XML schema to validate against. These schemas are named within the XML document. However, for interoperability purposes, there is a requirement that data providers expose metadata in simple Dublin Core. The Dublin Core (DC) Metadata Element Set (or simple DC) is made up of fifteen elements maintained by the Dublin Core Metadata Initiative (DCMI) (<http://www.dublincore.org/>). Its primary use is for simple resource description and discovery. The DCMI also maintains a set of qualifications to simple DC; when used, the metadata schema is often called qualified DC. Some OAI data providers expose metadata in both simple Dublin Core and another, perhaps more appropriate, schema such as MARC or MODS. It should be noted that this does not necessarily require multiple databases to store the item in each metadata schema. Because the metadata is harvested in XML, it is possible to transform MARC XML records, for instance, into simple Dublin Core XML records through Extensible Stylesheet Language Transformations (XSLT). XSLTs define how to transform one XML document into another. While developing XSLTs is not easy per se, the difficult part of this process is the intellectual mapping from one metadata schema to another.

Each item exposed through OAI is required to have a unique and persistent OAI identifier. An item is the information in a data provider's repository from which an OAI record can be built. An item can be exposed in multiple metadata schemas. An OAI record is the item with its unique OAI identifier, a timestamp, and in a specific metadata schema. This distinction can be difficult to wrap your mind around. But imagine that you have a database of items in the MODS format that you provide access to in both simple Dublin Core and MODS via the OAI protocol. An OAI service provider will ask for a specific item in simple Dublin Core using its OAI identifier but it can also ask for the exact same item in MODS using the same OAI identifier. The information that the service provider receives back is the OAI record.

The timestamp in an OAI record should correspond to the date that the metadata was created or last modified in order to support the concept of incremental harvests. A service provider may choose to harvest all metadata records within a repository, or it may choose to harvest records that are new or changed since a specific date.

## How OAI-PMH Basically Works (in a Picture)



## OAI-PMH Verbs

### 1. Identify

- *Purpose:* Return general information about the archive and its policies (e.g., datestamp granularity)
- *Parameters:* None
- *Example:* <http://aerialphotos.granger.uiuc.edu/oai.asp?verb=Identify>

### 2. ListSets

- *Purpose:* Provide a listing of sets in which records may be organized (may be hierarchical, overlapping, or flat)
- *Parameters:* None
- *Example:* <http://aerialphotos.granger.uiuc.edu/oai.asp?verb=ListSets>

### 3. ListMetadataFormats

- *Purpose:* List metadata formats supported by the archive as well as their schema locations and namespaces

- *Parameters:* identifier – for a specific record (Optional)
- *Example:* <http://aerialphotos.granger.uiuc.edu/oai.asp?verb=ListMetadataFormats>

#### 4. ListIdentifiers

- *Purpose:* List headers for all items corresponding to the specified parameters
- *Parameters:* from – start date (Optional)  
until – end date (Optional)  
set – set to harvest from (Optional)  
metadataPrefix – metadata format to list identifiers for (Required)  
resumptionToken – flow control mechanism (Exclusive)
- *Example:*  
[http://aerialphotos.granger.uiuc.edu/oai.asp?verb=ListIdentifiers&metadataPrefix=oai\\_dc](http://aerialphotos.granger.uiuc.edu/oai.asp?verb=ListIdentifiers&metadataPrefix=oai_dc)

#### 5. GetRecord

- *Purpose:* Returns the metadata for a single item in the form of an OAI record
- *Parameters:* identifier – unique id for item (Required)  
metadataPrefix – metadata format for the record (Required)
- *Example:* [http://aerialphotos.granger.uiuc.edu/oai.asp?verb=GetRecord&identifier=oai:aerialphotos.granger.uiuc.edu:AP-1A-1-1940&metadataPrefix=oai\\_dc](http://aerialphotos.granger.uiuc.edu/oai.asp?verb=GetRecord&identifier=oai:aerialphotos.granger.uiuc.edu:AP-1A-1-1940&metadataPrefix=oai_dc)

#### 6. ListRecords

- *Purpose:* Retrieves metadata records for multiple items
- *Parameters:* from – start date (Optional)  
until – end date (Optional)  
set – set to harvest from (Optional)  
resumptionToken – flow control mechanism (Exclusive)  
metadataPrefix – metadata format (Required)
- *Example:*  
[http://aerialphotos.granger.uiuc.edu/oai.asp?verb=ListRecords&metadataPrefix=oai\\_dc](http://aerialphotos.granger.uiuc.edu/oai.asp?verb=ListRecords&metadataPrefix=oai_dc)

## **Glossary – Open Archives Initiative Protocol for Metadata Harvesting**

### ***Aggregator***

An OAI aggregator is both a Service Provider and a Data Provider. It is a service that gathers metadata records from multiple Data Providers and then makes those records available for gathering by others using the OAI-PMH. (from *OAI for Beginners*).

### ***Archive***

The term "archive" in the name *Open Archives Initiative* reflects the origins of the OAI – in the E-Prints community where the term archive is generally accepted as a synonym for repository of scholarly papers. Members of the archiving profession have justifiably noted the strict definition of an “archive” within their domain; with connotations of preservation of long-term value, statutory authorization and institutional policy. The OAI uses the term “archive” in a broader sense: as a repository for stored information. Language and terms are never unambiguous and uncontroversial and the OAI respectfully requests the indulgence of the professional archiving community with this broader use of “archive.” (from FAQ on OAI Web site)

### ***Data Provider***

A Data Provider maintains one or more repositories (web servers) that support the OAI-PMH as a means of exposing metadata. (from FAQ on OAI Web site)

### ***Datestamp***

In OAI-PMH, every metadata record disseminated by a repository includes a datestamp value indicating the date (and optionally time down to the second) when that record was created or last modified. Datestamps enable selective harvesting of metadata by last modified date and time.

### ***Dublin Core (DC)***

Dublin Core (DC) is a metadata schema maintained by the Dublin Core Metadata Initiative (DCMI). The Dublin Core Metadata Element Set (DCMES) contains fifteen elements for simple resource description and discovery. Designed to be flexible and simple, DC’s elements are all optional and repeatable. DC has been extended with further optional elements, element qualifiers and vocabulary terms. See <http://www.dublincore.org/>.

### ***Document Type Definition (DTD)***

A DTD is a formal specification of the structure of a class of XML documents. DTDs (or XML Schemas which serve essentially the same purpose) are used to validate the correctness and completeness of XML document instances. By definition all OAI-PMH responses must be valid XML document instances.

### ***E-print***

An e-print is an author self-archived document. In the sense that the term is ordinarily used, the content of an e-print is the result of scientific or other scholarly research.

***Flow control***

The management of the flow of data between Data Provider and Service Provider in order to assure that neither end of the transaction suffers overload. Flow control is generally realized in the OAI-PMH through the use of *resumption tokens*.

***Harvester (see Service Provider)***

***Harvesting***

In the OAI context, harvesting refers specifically to the gathering together of metadata from a number of distributed repositories into a combined data store.

***Identifier***

In OAI-PMH an identifier is a unique key for an metadata item in a repository. Each item of metadata in a repository must have a locally unique, persistent identifier. If the repository is registered with the Open Archives Website (<http://www.openarchives.org/community/>), and if the repository conforms to the recommended guidelines for generating OAI Identifiers, then the OAI Identifiers of that repository will be globally unique among all registered OAI providers.

***Interoperability***

Interoperability is the ability of systems, services and organizations to work together seamlessly toward common or diverse goals. In the technical arena it is supported by open standards for communication between systems and for description of resources and collections, among others. Interoperability is considered here primarily in the context of resource discovery and access. (from *OAI for Beginners*)

***Item***

In OAI-PMH, item is the label for the totality of metadata contained in a repository about a single resource. Each item of metadata in a repository has a unique identifier and may be disseminated from the repository in one or more metadata formats. A metadata item disseminated in a particular metadata format is an OAI-PMH metadata record.

***Metadata***

Structured information about resources (including both digital and non-digital resources). Metadata can be used to help support a wide range of operations on those resources. In the context of services based on metadata harvested via OAI-PMH, the most common operation is discovery and retrieval of resources. (from *OAI for Beginners*)

***Open Archives Initiative (OAI)***

OAI is an initiative to develop and promote interoperability standards that aim to facilitate the efficient dissemination of content.

***OAI-PMH (OAI Protocol for Metadata Harvesting)***

OAI-PMH is a lightweight harvesting protocol for sharing metadata between services.



***Protocol***

A protocol is a set of rules defining communication between systems. FTP (File Transfer Protocol) and HTTP (Hypertext Transport Protocol) are examples of other protocols used for communication between systems across the Internet. (from *OAI for Beginners*)

***Record***

In OAI-PMH a record is a metadata item disseminated in a specific metadata format. An OAI metadata record is uniquely identified by the associated item's OAI Identifier, the metadata format of the record (metadataPrefix), and the record's datestamp value.

***Repository***

In OAI-PMH a repository is a network accessible server that is able to process OAI-PMH requests correctly. It accepts OAI Requests and generates valid OAI Responses in reply.

***Resource***

In OAI-PMH a resource is an object the metadata is "about". The nature of resources is not defined in the OAI-PMH. Resources may be digital or non-digital.

***Resumption Token***

When responding to certain OAI-PMH requests, data providers may optionally send an incomplete response. For instance when asked to list all records in a repository containing 1,000 items, a data provider may wish to send only 100 at a time. A ResumptionToken is included in any partial response and may be used by the harvester when requesting the next part of response.

***Service Provider***

A Service Provider issues OAI-PMH requests to data providers and uses the metadata returned as a basis for building value-added services. (from FAQ on OAI Web site)

***Set***

In the OAI-PMH a Set is an optional construct for grouping items in a repository. Sets allow selective harvesting of metadata.

***Value-added service***

A service that is based on harvested metadata, and adds value for its users by means which may include normalization and enriching of the harvested metadata for example. Types of services which may be offered include search services, citation linking, overlay journals, and peer-review services, among others. (from *OAI for Beginners*)

***XML (Extensible Markup Language)***

XML is a language for creating other languages. It defines a means of describing data. XML can be validated against a DTD or schema setting out the elements of the language created. XML mappings exist for a number of metadata record formats.

***XML namespace***

An XML namespace is a collection of names, identified by a URI reference [RFC2396], which are used in XML documents as element types and attribute names. XML namespaces differ from the "namespaces" conventionally used in computing disciplines in that the XML version has internal structure and is not, mathematically speaking, a set. (Definition quoted from W3C—Namespaces in XML at <http://www.w3.org/TR/REC-xml-names/>)

### ***XML schemas***

XML Schemas express shared vocabularies and allow machines to carry out rules made by people. They provide a means for defining the structure, content and semantics of XML documents. (Definition quoted from W3C Architecture Domain—XML schema at <http://www.w3.org/XML/Schema>)

## **Open Archives Initiative Resources:**

### Open Archives Initiative

- <http://www.openarchives.org>
- <http://www.openarchives.org/documents/FAQ.html>
- <http://www.openarchives.org/OAI/openarchivesprotocol.htm>
- <http://www.openarchives.org/OAI/2.0/guidelines-static-repository.htm>
- <http://www.openarchives.org/community/>

### OAI for Beginners: An Open Archives Forum Online Tutorial

- <http://www.oaforum.org/tutorial/>

### University of Illinois OAI-PMH Data Provider Registry

- <http://gita.granger.uiuc.edu/registry/>

### XML Schema Validator

- <http://www.w3.org/2001/03/webdata/xsv>

### XML Tools at W3C

- <http://www.w3.org/XML/#software>

### Examples of Service Providers:

- OAster: <http://oaister.umdl.umich.edu/o/oaister/>
- IMLS Digital Collections and Content: <http://imlsdcc.granger.uiuc.edu/>
- Digital Library Federation: <http://www.hti.umich.edu/i/imls/>
- DLF Aquifer Portal: <http://www.hti.umich.edu/a/aquifer/>
- See list of service providers: <http://gita.granger.uiuc.edu/registry/services/>

## **What is Shareable Metadata?**

Shareable or interoperable metadata is simply metadata which can be fruitfully utilized outside of its local environment. High quality metadata<sup>1</sup> may or may not be truly shareable metadata. That is, metadata may be of high quality within its local context, but may be compromised when taken out of this context for various reasons. There are some particular characteristics that make quality metadata more useful in a shared environment.

### **Use of standards (metadata formats, vocabularies, content guidelines)**

The use of standards enables the better integration of metadata records from one source with records from other sources.

### **Proper context**

In a shared environment, metadata records will become separated from any high-level context applying to all records in a group, and from other records presented together in a local environment. It is therefore essential that each record contain the context necessary for understanding the resource the record describes, without relying on outside information.

### **Content coherence**

Metadata records for a shared environment need to contain enough information such that the record makes sense standing on its own, yet exclude information that only makes sense in a local environment. This can be described as sharing a 'view' of the native metadata.

### **Consistency**

Even high-quality metadata will vary somewhat among metadata creators. All decisions made about application of elements, syntax of metadata values, and usage of controlled vocabularies, should be consistent within an identifiable set of metadata records so those using this metadata can apply any necessary transformation steps without having to process inconsistencies within such a set.

### **Technical conformance**

Metadata should conform to the technical specifications of the aggregated environment (for example, XML for an OAI environment) and should be properly encoded.

### **Documentation**

Metadata creators should publicly document the decisions made about their metadata formats, vocabularies, and content guidelines particularly if the metadata itself does not allow such communication or if local standards are used.

## **Benefits of creating shareable metadata**

Creating shareable metadata requires an investment of time. However, there are many benefits gained from making this investment.

---

<sup>1</sup> For an excellent discussion of what metadata quality is, see: Bruce, T.R. and D.I. Hillmann. (2004) "The continuum of metadata quality: defining, expressing, exploiting," in *Metadata in Practice*, Ed. by Diane Hillmann and Elaine Westbrooks. Chicago: ALA Editions.

The first and perhaps most significant benefit to creating shareable metadata is that it will be interoperable, or meaningful when combined with metadata from other sources. By using metadata schemas and rules for creating metadata values similar to those used by others, your resources can meaningfully appear in search results alongside related resources from other metadata providers.

When creating true shareable metadata, your resources are more likely to be found when pooled together with resources from other providers, rather than not being retrieved by searchers due to inconsistencies or gaps in description. Your resources therefore will receive more exposure, and end-users will have the opportunity to make previously unseen connections between your resources and those from other metadata providers.

Finally, creating shareable metadata increases the number of access points for your resources available to end-users. Aspects of a resource not previously explicitly described are often added when metadata creators think in terms of shareable metadata.

### **Some Best Practices for Shareable Metadata**

*(from Best Practices for Shareable Metadata*

*<http://oai-best.comm.nsd.gov/cgi-bin/wiki.pl?PublicTOC>)*

#### ***Appropriate Representation of Resources***

- Think about uses of your metadata in a federated environment, and tailor your records for those purposes.
- Use metadata formats appropriate to your resources and your intended communities.
- Use metadata elements and construct values for those elements appropriate for a shared environment.
- Include the appropriate context for the resource in a metadata record.

#### ***Granularity of Description***

- Expose metadata records at the smallest level of granularity appropriate for the resources being described.
- Do not expose individual metadata records for digital objects that are only subordinate parts to a single item.

#### ***Crosswalking Metadata***

- Map metadata from more robust formats to simpler ones.
- Plan for both mapping values between fields and for transforming data values themselves to meet the expectations of the target metadata format.
- Repeat elements when your target metadata schema allows it.
- Include titles and appropriate context in your mapped metadata.
- Exclude indications of unknown or inapplicable data, and artifacts of descriptive practices not applicable to the target metadata format.
- Stepped crosswalking may be beneficial.

### ***Describing Versions and Reproductions***

- Adhere to the one-to-one principle when practical.
- When it is necessary to provide access to multiple versions of a resource, carefully select a strategy from options used by others in your community of practice.

### ***Linking Issues***

- URIs in metadata records should permanently point to the resource being described.
- URIs in metadata records should include indication of the protocol used.
- Provide a single URL pointing to a resource in an appropriate context.

### ***Providing Supplemental Documentation to Metadata Aggregators***

- Provide documentation on choices made when providing metadata for a shared environment

### ***Expressing Rights for Metadata Records***

- Use mechanisms in metadata sharing protocols for expressing rights over metadata when they exist; add this information to your supplemental documentation when these mechanisms don't exist.
- Only restrict use of metadata for use in a shared environment in specialized circumstances.

### ***Specific Classes of Data Elements***

#### *Titles*

- Provide a title in every record. Supply one, if necessary, according to established standards.
- Express multiple titles in repeated fields.
- Make the distinction between title and sub-title clear through the metadata format used or through standard punctuation.

#### *Names*

- Include all known names expected by your community of practice.
- Format names consistently, according to authority files or standards expected by your community of practice.
- Provide as granular an encoding of a name as possible in the metadata schema used.
- Express multiple names in repeated fields.

## Dates

- Date elements in records should contain values important for discovery of the resource by end-users.
- When providing multiple dates, clearly indicate the relationship of each to the resource, and repeat the relevant date element for each date.
- Include only easily-parsable values in date elements.
- Present dates in a consistent format, according to established machine-readable standards.

## Subjects

- Choose subject values from relevant controlled vocabularies consistently and explicitly.
- Repeat subject information in more specific fields when they are available in the metadata format being used.
- Express multiple subjects in repeated fields.

## Language of Content

- Supply a language element when relevant to the resource.
- Format the value of the language element according to the rules of the metadata format in use.
- Express multiple titles in repeated fields.
- Supply the language of the metadata record only in a metadata element specifically designed for this purpose.

## Geographic Places

- Choose geographic place values from relevant controlled vocabularies consistently and explicitly.

## Identifiers

- Include recognized standard identifiers when available.
- Include a URI or DOI linking to the resource when available.
- Explicitly encode the nature of an identifier provided.
- Express multiple identifiers in repeated fields.

## Rights over Resources

- Include rights information about a resource in the most granular format possible.
- State rights information in plain language intended for the end-user of a resource.

## Bibliographic Citation

- Provide a bibliographic citation either to identify the work described in a metadata record or to reference a related work.
- Format the citation using standards applied in your community of practice.

## Types of Resources

- Present format and type/genre information in all records.
- Choose type values from relevant controlled vocabularies consistently and explicitly.
- Express multiple type terms in repeated fields.

## Choosing a Metadata Standard for Use with OAI

The use of multiple metadata formats (at least one in addition to the required simple Dublin Core) is strongly encouraged, and this list is a necessarily incomplete list of potential metadata formats for use with OAI

### **MODS: Metadata Object Description Standard**

MODS may be a good option for an additional metadata schema to expose via OAI for data providers who:

1. locally engage in descriptive practices heavily influenced by resource description standards in libraries, and,
2. have as a primary audience for resources described via OAI records a community well-versed in library descriptive practices, yet also want robust records in a format accessible to service providers outside the core library community.

The MODS v.3.2 XML Schema is available at  
<http://www.loc.gov/standards/mods/v3/mods-3-2.xsd>.

### **Qualified Dublin Core**

Qualified Dublin Core may be a good option for an additional metadata schema to expose via OAI for data providers who:

1. have a need for more granularity of description than is available in simple Dublin Core but not a fundamentally different approach to resource description, and,
2. use controlled vocabularies that they wish to specify within their metadata records, and,
3. have resources of interest to many different knowledge communities with disparate descriptive metadata practices.

There does not exist a single canonical Qualified Dublin Core XML schema. However, a XML schema for Qualified DC can be created through the importation of the necessary namespaces and schemas.

<http://dublincore.org/schemas/xmls/qdc/2003/04/02/dc.xsd>

<http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcterms.xsd>

<http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcmitype.xsd>

### **MARCXML**

MARCXML may be a good option for an additional metadata schema to expose via OAI for data providers who:

1. locally describe resources in MARC according to AACR2r, and,
2. have as a primary audience for resources described via OAI records the core library community.

The MARCXML XML Schema is available at  
<http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>.

### **CDWA Lite**

CDWA Lite may be a good option for an additional metadata schema to expose via OAI for data providers who:

1. wish to describe works of art and material culture
2. intend their metadata to be used by specialist audiences in the art domain

The CDWA Lite XML Schema version 1.1 is available at  
<http://www.getty.edu/CDWA/CDWALite/CDWALite-xsd-public-v1-1.xsd>.

### **ETD-MS: Electronic Theses, and Dissertations Metadata Standard**

ETD-MS may be a good option for an additional metadata schema to expose via OIA for data providers who:

1. are primarily exposing metadata about electronic theses and dissertations.
2. wish to contribute to aggregations of electronic theses and dissertations such as the Networked Digital Library of Theses and Dissertations

The ETD-MS XML Schema is available at  
<http://www.ndltd.org/standards/metadata/etdms/1.0/etdms.xsd>.



## Metadata Resources

Caplan, Priscilla. *Metadata Fundamentals for All Librarians*. Chicago: ALA Editions, 2003.

A good introduction to a range of metadata issues and formats. Aimed at librarians but quite accessible to other communities as well.

Digital Library Federation & National Science Digital Library. Best Practices for Shareable Metadata. <http://oai-best.comm.nsdlib.org/cgi-bin/wiki.pl?PublicTOC>

A community based effort to define best practices for shareable metadata. Much of the work in the pre-conference is drawn on this document. This is still a work in progress so changes may occur.

Digital Library Federation. (2005) *MODS Implementation Guidelines for Cultural Heritage Materials*. Draft for Public Comment and Review. [http://www.diglib.org/aquifer/DLF\\_MODS\\_ImpGuidelines\\_ver4.pdf](http://www.diglib.org/aquifer/DLF_MODS_ImpGuidelines_ver4.pdf).

This document represents explicit guidelines for how to implement MODS specifically for sharing records.

Getty Research Institute. *Introduction to Metadata: Pathways to Digital Information*. (Edited by Murtha Baca) [http://www.getty.edu/research/conducting\\_research/standards/intrometadata/index.html](http://www.getty.edu/research/conducting_research/standards/intrometadata/index.html)

Another good basic introduction to metadata.

Hillmann, Diane I. and Elaine Westbrook, eds. *Metadata in Practice*. Chicago: ALA Editions, 2004.

Presents a series of case studies from libraries, museums, and other communities working with metadata. Contains some specific case studies focused on OAI and metadata interoperability.

National Science Digital Library. *NSDL Metadata Primer*. <http://metamanagement.comm.nsdlib.org/outline.html>.

The NSDL primer for institutions interested in contributing metadata to the NSDL via the OAI protocol, but also more broadly applicable.

Research Libraries Group. *Descriptive Metadata Guidelines for RLG Cultural Materials*. 2004. [http://www.rlg.org/en/pdfs/RLG\\_desc\\_metadata.pdf](http://www.rlg.org/en/pdfs/RLG_desc_metadata.pdf)

Guidelines for institutions interested in contributing to RLG's Cultural Materials database, but very useful for any institution with cultural heritage materials.