

10/21/2010

# Best Practices for CONTENTdm and other OAI- PMH compliant repositories

*creating shareable metadata*



## Best Practices for CONTENTdm and other OAI-PMH compliant repositories: creating shareable metadata

---

© 2010 OCLC, Inc. October 2010

Please direct correspondence to:

Geri Bunker Ingram, [ingramg@oclc.org](mailto:ingramg@oclc.org) or Jason Beatrice Lee, [leeja@oclc.org](mailto:leeja@oclc.org)  
OCLC Digital Collection Services

### Background and acknowledgements:

*Throughout the digital repository landscape, it is increasingly accepted that metadata needs not only to serve the local community but also be suitable for harvesting externally. The challenge is to sustain useful local information while providing context and perspective to both the local and the remote user. Because each metadata standard and each collection management toolset may derive its own 'best practice,' it is incumbent upon each community of practice to provide leadership from its constituents' particular points of view.*

*Thus, in August 2009, OCLC Digital Collection Services (DCS) convened the CONTENTdm Metadata Working Group (MWG) to create a 'best practices' guideline for our community. Discussions followed presentations given at regional and national CONTENTdm Users Groups, and collaborative work was undertaken using the tools familiar to the collective—CONTENTdm, WorldCat Digital Collection Gateway, (Gateway) and various social networking environments. The discussion focused on members' research and publications, and on their efforts to develop, optimize and standardize CONTENTdm metadata element sets such that materials are discoverable easily both in the local CONTENTdm environment as well as across repositories into which their metadata might be harvested according to the standard OAI protocols.*

*OCLC DCS allocated CONTENTdm servers and trained the MWG members to use the Gateway to map qualified Dublin Core metadata and test them against WorldCat.org displays and WorldCat MARC fields. In the course of the work, the MWG untied several knotty issues and made suggestions resulting in significant improvements to the Gateway. In July, 2010, the Gateway was opened to any OAI-PMH compliant repository.*

*OCLC Digital Collection Services would like to thank the participants in the CONTENTdm Metadata Working Group<sup>i</sup>, and their colleagues, for their invaluable contribution to this guide.*

## Contents

---

Creating sharable metadata.....	1
Challenges	
Recommendations	
Opportunities	
Recommended Core Elements for CONTENTdm Digital Collections.....	3
Recommended 'as Appropriate' Elements.....	15
For Further Reading.....	19
Endnotes.....	19
Appendix A: Subjects.....	21
Appendix B: Dates.....	21
Appendix C: Sample schemas	
Photographic collections.....	22
Archival collections.....	24
Appendix D: CONTENTdm compound objects.....	25
Appendix E: Considerations for consortia.....	26
Appendix F: Frequently Asked Questions.....	27
Appendix G: Digital Collection Gateway enhancements.....	33

## **Creating Shareable Metadata:**

### **Challenges**

Essentially there are four types of problems that we see when metadata are viewed outside the context of the collection home. These were generally described in a 2006 article<sup>ii</sup> published by *First Monday*.

Typical problems include:

- Lack of consistency within a single collection.  
-Example: The use of both the Dublin Core <date> and <coverage> elements to record some variant of the resource creation date.
- Too much information.  
-Example: Inclusion of technical information such as date digitized and type of scanner used.
- Lack of key contextual information.  
-Example: Exclusion of a collection name that is essential to make sense of the record.
- Lack of conformance to technical standards.  
-Example: Metadata encoded in XML with character encoding problems.

### **Recommendations**

Likewise, Shreeves, 2006, recommends several general practices which CONTENTdm collection administrators would do well to consider. They include:

- We encourage institutions to think carefully about how they might generate *multiple views* of resources using the metadata already created rather than simply sharing a single record describing everything about a resource.
- An institution should understand what an aggregator needs included in the metadata (learning standards? audience level?) to support its service and, when possible, work to meet those needs.
- Metadata aggregators can more effectively normalize records from metadata providers if all records within a defined set are consistent both semantically and syntactically.

- When multiple values are needed, the metadata element should be repeated.

And from MJ Han et al at the University of Illinois<sup>iii</sup> come these further recommendations. Since their research focused on sharing CONTENTdm collection metadata with OAI harvesters, these are especially relevant to our community:

- Keep a balance between specificity and generality in defining local fields.
- Decide at the outset which locally defined fields are intended only for the local environment and which should be made available to aggregators.
- Be cognizant of how values will be created in the local environment.
- Maximize use of Qualified Dublin Core elements for labeling in the local environment.
- Consider taking field names and definitions, if possible, directly from other metadata standards such as EAD, VRA Core, and CDWA when creating locally developed application profiles.
- Share the logic of mapping decisions with aggregators.

### **Opportunities**

*In the current metadata aggregation landscape, it is safe to assume that users search and browse for resources at an aggregator's site then follow a link back to the home institution for access to the resource itself and any additional metadata. Therefore, when creating metadata for the purposes of inclusion in these aggregations, one can afford to be selective about the data elements included, with the understanding that a user will find his way to the local records for full contextual information. (Shreeves, 2006)*

On July 20, 2009, the OCLC *Digital Collection Gateway* became available to all CONTENTdm 5.1 users in the form of CONTENTdm *WorldCat Sync*. This integrated function enables a CONTENTdm collection administrator to map qualified and simple Dublin Core elements from digital items held in the CONTENTdm collection, to MARC fields, creating and modifying WorldCat records that are synchronized on a schedule set by the collection administrator. The Gateway thus represents a timely opportunity to provide specific Dublin Core metadata schemas for use in CONTENTdm and intended for OAI-PMH harvesting, and underscores a rather urgent need to provide advice to our community.

Below are some notes on creating and configuring metadata for discovery of digital items in WorldCat.org:

- For all fields that you want to display in WorldCat, configure the metadata fields in CONTENTdm so that those fields are mapped to an appropriate Dublin Core element. You

can use any Simple Dublin Core and Qualified Dublin Core elements. We recommend using Qualified Dublin Core elements for the best mapping results.

- Date fields should use consistent date formatting.
- Metadata fields set to hidden in CONTENTdm are not available for use with the Digital Collection Gateway.
- If you opt to make a field “Non-Searchable” in CONTENTdm and map that field into the Digital Collection Gateway, the field will be searchable in WorldCat.org.

## Recommended *Core* Metadata Elements for CONTENTdm Digital Collections

*“An element is a descriptive category of information about the resource.... All of the elements used to describe a resource together make up a **record**.”- NCSU Libraries Core 1.0 Metadata Element Set Best Practices*

The following is a set of guidelines for collections created in CONTENTdm. These guidelines promote the simplification of local information to enable better end-user discovery in an aggregated environment. As with any Best Practices Guide, it is recommended that catalogers follow basic rules of consistency with grammar and syntax (content standard) set forth in resources such as AACR2, DACS, CCO, etc., as well as incorporate the use of controlled vocabularies such as LCSH, AAT, MeSH, and authority lists such as LCNAF and ULAN or 'locally-grown' thesauri as appropriate to the subject matter of a resource. For each digital collection, a collection-level record should be created along with item-level records. Metadata elements should contain labels most useful to the local environment, but should be mapped to standard Dublin Core elements.

\*A note about repeating fields: A number of works have been published offering best practices for configuring OAI-harvestable metadata. Although these works recommend repeating fields versus multiple values, in some cases multiple values (separated by a semicolon) are preferred for accuracy depending upon the level of complexity in configuring a collection using your digital collections management software and the OAI harvesting tool. For example, semicolon-separated values can be easily accommodated in CONTENTdm as well as display accurately when synced to WorldCat.org via the Digital Collection Gateway. When in doubt, test your data sets against your chosen OAI harvester.

<b>CORE ELEMENTS</b>	
<i>Element Name</i>	<b>Title</b>
<i>Definition</i>	The name of a resource; a caption
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	Title
<i>Repeatable?</i>	Yes
<i>Best Practice</i>	
<ul style="list-style-type: none"> <li>• Prefer non-numeric description of resource, excluding material-<i>type</i> information if possible.</li> <li>• Prefer dct: alternative (for translated titles, etc.) to be used rather than multiple values for dcterms.title.</li> <li>• Prefer non-use of explanatory or qualifying symbols (e.g. brackets to indicate cataloger-supplied title).</li> <li>• WorldCat.org display mapping: &lt;dcterms.title&gt; maps to wc.Title (MARC 245) and wc.Other Titles (MARC 246).</li> <li>• Secondary titles (dcterms.alternative) and repeating elements (dcterms.Title2) should be mapped to wc.Other Titles.</li> </ul> <p><i>“Make the title descriptive yet brief. Use generic titles to bring together different images of the same subject, if possible (e.g., use Mayor Benjamin Bosse on all photos of him, so they display together by title).” – Metadata Guidelines, Evansville Photos Collection, Evansville Vanderburgh Public Library.</i></p>	



<i>Element Name</i>	Creator
<i>Definition</i>	Entity primarily responsible for creating the intellectual content of a resource
<i>Controlled Vocabulary</i>	LCNAF; ULAN
<i>DC Element Map</i>	Creator
<i>Repeatable?</i>	Not preferred*
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Creators include individual and corporate authors, artists, etc.</li> <li>• Named entities may be repeated in Subject-Name field if deemed appropriate.</li> <li>• Prefer non-use of 'junk value' (e.g. "Unknown,") however, it is appropriate to qualify named entities with "[role]."</li> <li>• "Prefer use of Name (personal or corporate) Authority Source to be used consistently throughout description of a resource and from one resource to another." - <i>Metadata Implementation Guidelines for North Carolina Digital State Documents</i></li> <li>• * Repeatability: Digital Collection Gateway handles repeating dcterms.Creator fields by mapping a main creator ("Creator01") to MARC 100 and shunts additional creators ("Creator02") to MARC 720.</li> <li>• WorldCat.org display mapping: dcterms.Creator maps to wc.Author (MARC 100) and wc.All Authors/Contributors (MARC 720) A third WorldCat.org element (wc.Named Person) is <i>additionally</i> recommended for populating if deemed appropriate.</li> </ul> <p><i>"Do not use honorifics, titles, or nicknames unless it is necessary to disambiguate (e.g., the first name of the person is unknown). Otherwise, these alternate forms of names (such as "Buddy" Jones; Reverend Murrell; Dr. Reed) may be used in the Description field but not as the authoritative version...." – Huntington Digital Library Guidelines, The Huntington Library</i></p>	

<i>Element Name</i>	Description
<i>Definition</i>	Brief account of the content of a resource (e.g., summary or abstract)
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	Description*
<i>Repeatable?</i>	Yes
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Summaries, abstracts, and contextual information can all be used to describe a resource.</li> <li>• Some digital collections management practitioners prefer the local practice of mapping separate Table of Contents, Abstract, and similar local elements, to dcterms.Description.</li> <li>• *Prefer collection-based cataloger decision on enabling full-text searching for this field. <ul style="list-style-type: none"> <li>○ <i>If data type <b>Full text search</b>, prefer <b>no</b> mapping to WorldCat.org. <ul style="list-style-type: none"> <li>▪ Instead use dcterms.Description.Abstract IF simple dcterms.Description has been enabled for full text searching.)</li> </ul> </i></li> <li>○ <i>If data type <b>text</b>, prefer mapping dcterms.Description to wc.Summary (MARC 520 [8 ]) or wc.Abstract (MARC 520 [3 ]).</i></li> </ul> </li> <li>• A third WorldCat.org element (wc.Contents MARC 500) is <i>additionally</i> recommended for populating if deemed appropriate</li> </ul> <p><i>“Also include any other information a searcher might need to find an image through a keyword search or to understand the context of the image: Is there a view of the Mississippi River? Was a photograph taken from the future site of a university library? Does a building no longer exist? What location was a photograph taken from? Is it an aerial view” – WAICU Metadata Guide</i></p>	

<i>Element Name</i>	Contributor
<i>Definition</i>	Additional writer, illustrator, editor, finding aid author, etc.
<i>Controlled Vocabulary</i>	LCNAF; ULAN
<i>DC Element Map</i>	Contributor
<i>Repeatable?</i>	Yes
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Prefer use of Name (personal or corporate) Authority Source to be used consistently throughout description of a resource and from one resource to another.</li> <li>• Contributors are named so because they are judged NOT to have <i>equal</i> responsibility for the creation of a work.</li> <li>• Named entities may be repeated in Subject-Name field if deemed appropriate.</li> <li>• Prefer non-use of 'junk value' (e.g. "Unknown,") however, it is appropriate to qualify named entities with "[role]."</li> <li>• WorldCat.org display mapping: dcterms.Contributor maps to wc.All Authors/Contributors. (MARC 720) A third WorldCat.org element (wc.Named Person) is recommended for populating if deemed appropriate.</li> </ul> <p><i>"Persons or organizations who made significant intellectual contributions to the resource, but whose contribution is usually secondary to the person or organization specified in the Creator element. Examples include co-author, editor, transcriber, translator, illustrator, etc." – Metadata Implementation Guidelines for North Carolina Digital State Documents</i></p>	

<i>Element Name</i>	<b>Publisher</b>
<i>Definition</i>	Person or Corporate/ Organizational entity responsible for producing a resource or a digital copy of a resource
<i>Controlled Vocabulary</i>	LCNAF
<i>DC Element Map</i>	<b>Publisher</b>
<i>Repeatable?</i>	<b>Yes</b>
<i>Best Practice</i> <ul style="list-style-type: none"><li>• Prefer use of Name (personal or corporate) Authority Source to be used consistently throughout description of a resource and from one resource to another.</li><li>• Prefer non-use of 'junk value' (e.g. "Unknown").</li><li>• Prefer "digitized by" or other text prefix to qualify value.</li><li>• WorldCat.org display mapping: dterms.Publisher maps to wc.Publisher (MARC 260 \$b).</li></ul> <p><i>"The entity responsible for making the Resource available in its present form, such as a corporate publisher, a university department, or a cultural institution." – University of Wisconsin Digital Library Data Dictionary</i></p>	

<i>Element Name</i>	<b>Subject:</b>
<i>Definition</i>	Terms, keywords or phrases, describing the content of a resource
<i>Controlled Vocabulary</i>	LCSH, LCNAF, AAT, TGN
<i>DC Element Map</i>	<b>Subject</b>
<i>Repeatable?</i>	Yes
<p><b>Best Practice</b></p> <ul style="list-style-type: none"> <li>• Terms include topics, events and geographic, personal and corporate named entities</li> <li>• Prefer use of standard controlled vocabularies and name authority sources.</li> <li>• WorldCat.org display mapping: prefer map dcterms.Subject to MARC 650 if controlled, to MARC 653 if uncontrolled</li> </ul> <p><i>“Use subject terms that describe what an object is as well as what it is about. Example 1: Mural painting and decoration; Derry (Northern Ireland); Ireland—History—Easter Rising, 1916.” – Guidelines for Metadata Application in the Claremont Colleges Digital Library</i></p>	

<i>Element Name</i>	<b>Identifier</b>
<i>Definition</i>	Unique numeric or alphanumeric character string used to locate or label a resource
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	<b>Identifier</b>
<i>Repeatable?</i>	Yes *
<p><b>Best Practice</b></p> <p><b>Locate:</b> Gateway selects the first dcterms.Identifier that contains a URL and makes it the default value for the resolution URL</p> <ul style="list-style-type: none"> <li>• Gateway will place the URL in MARC 856\$u and will include a text string reading <i>item resolution URL</i>. <ul style="list-style-type: none"> <li>○ If your resolution URL is in a field other than the first dcterms.Identifier field, you will map it separately. <ul style="list-style-type: none"> <li>▪ Use the <i>Edit metadata map</i> function.</li> <li>▪ Choose the <b>WorldCat Item View</b>.</li> <li>▪ <b>Click</b> on the yellow box in the <b>Find a copy online</b> section, and map the URL</li> </ul> </li> </ul> </li> <li>• *Repeatability: It will take all other URLs in repeating dcterms.Identifier fields, and place them in repeating 856 fields but with no \$3 text. <ul style="list-style-type: none"> <li>➤ Re: Thumbnail display images: <ul style="list-style-type: none"> <li>▪ CONTENTdm supplies the Reference URL to dcterms.Identifier. This not only provides the resolution URL but also automatically generates the thumbnail for WorldCat.org.</li> <li>▪ OTHER OAI-compliant repositories: To display your thumbnail image in WorldCat.org, with Gateway Ver.2.3, select the yellow box labeled <i>Click to map thumbnail URL field</i> under the rectangle anchoring the position for a thumbnail. Then associate one of your source metadata fields with the thumbnail URL.</li> </ul> </li> </ul> </li> </ul> <p><b>Label:</b> Examples include accession number, ISBN, photo negative job/roll/frame number, call number, URI, etc.</p> <ul style="list-style-type: none"> <li>• Digital Collection Gateway automatically populates a value for a non-URL dcterms.Identifier (MARC 024).</li> </ul> <p><i>“If contributing a digital resource to a collaborative digital collection, consider prefixing the character string with an institutional code to keep your resources distinguishable from those owned by other institutions.” –Mountain West Digital Library Metadata Group</i></p>	

<i>Element Name</i>	<b>Language</b>
<i>Definition</i>	Depicted language(s) via text, audio, and/or video, of a resource
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	<b>Language</b>
<i>Repeatable?</i>	Yes
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Prefer standard authorized format [MARC code] of depicted language.</li> <li>• Multiple values are often used when a resource contains more than one language.</li> <li>• WorldCat.org display mapping: dcterms.Language maps to wc.Language (MARC 546)</li> </ul> <p><i>“Separate terms by semi-colon (;) and a space. For example, for French and English: fre; eng” – Metadata Supplement for Fashion Plate Collection, Claremont Colleges Digital Library</i></p>	

<i>Element Name</i>	<b>Rights</b>
<i>Definition</i>	Copyright & intellectual property permissions concerning legal use, access and reproduction of a resource
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	<b>Rights</b>
<i>Repeatable?</i>	<b>Yes</b>
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Prefer free text statement of rights to a 'lonely' url.</li> <li>• Rights statements should provide references or contact information. Additional clarification can be indicated via linking to an institutional policy statement or other web resource.</li> <li>• World.org mapping: dcterms.Rights maps to MARC 540. The following WorldCat.org element is recommended for populating only if strictly deemed necessary: wc.Responsibility</li> </ul> <p><i>“These statements should be given in the form: Rights status. Reproduction/use restrictions. Further information.” – Core 1.0 Metadata Element Set Best Practices, NCSU Libraries</i></p>	



<i>Element Name</i>	<b>Type</b>
<i>Definition</i>	The characteristic that identifies a resource by genre
<i>Controlled Vocabulary</i>	DCMI
<i>DC Element Map</i>	<b>Type</b>
<i>Repeatable?</i>	Yes
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Moving images, three-dimensional objects and sound recordings are all examples of Resource Types.</li> <li>• Prefer DCMI Type Vocabulary for controlled list of authorized terms: <a href="http://dublincore.org/documents/dcmi-type-vocabulary/">http://dublincore.org/documents/dcmi-type-vocabulary/</a></li> <li>• WorldCat.org display mapping: dcterms.Type maps to wc.Genre/Form (MARC 655)</li> </ul> <p><i>“This element should be populated from the DCMI type vocabulary, a controlled listing of genre types. It may be automatically populated, based on characteristics of the repository.” – NCSU Libraries Core 1.0 Metadata Element Set Best Practices</i></p>	

<i>Element Name</i>	<b>Format</b>
<i>Definition</i>	The media form of the resource
<i>Controlled Vocabulary</i>	MIME, AAT
<i>DC Element Map</i>	<b>Format</b>
<i>Repeatable?</i>	<b>Yes</b>
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Prefer use of MIME type (Internet Media Type) or two-part (type/subtype) identifier in a single string: <a href="http://www.iana.org/assignments/media-types/">http://www.iana.org/assignments/media-types/</a></li> <li>• WorldCat.org display mapping: dcterms.Format maps to wc.Notes (MARC 500) Note that dcterms.Format.Extent maps to MARC 300 and dcterms.Format.Medium maps to MARC 340.</li> </ul> <p><i>“New media types and applications are always emerging. If the resource format being described is not yet part of the MIME type list, select a broad category of object format for the first part of the MIME type, then use the file name suffix for the second half.” – University of Louisville CONTENTdm Cookbook</i></p>	

<i>Element Name</i>	<b>Date</b>
<i>Definition</i>	Publication date of a resource, or date a resource is issued
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	<b>Date</b>
<i>Repeatable?</i>	<b>Yes</b>
<i>Best Practice</i>	
<ul style="list-style-type: none"> <li>• Prefer ISO 8601 W3C Date/Time Format standard.</li> <li>• Prefer non-use of 'junk value' (e.g. "Unknown").</li> <li>• Follow a consistent standard method of inputting date ranges or uncertain dates (*some 'communities of practice' reference both the <i>Date-Digital</i> and the <i>Date-Original</i>. See Appendix B: Dates) Digital Collection Gateway maps the first available dcterms.Date (Date01) only</li> <li>• WorldCat.org display mapping: dcterms.Date maps to wc.Publisher (MARC 260\$c).</li> </ul> <p><i>"Similarly, if you will describe both physical and digital manifestation properties in your local system using unique field names, consider whether you intend to follow the Dublin Core one-to-one principle, in which case only metadata about one manifestation will be mapped and made available to aggregators." – Metadata for Special Collections in CONTENTdm: How to improve interoperability of Unique Fields through OAI-PMH</i></p>	

## Recommended 'as Appropriate' Elements

CONTENTdm practitioners may additionally choose to map the *Filename* element to <dc:identifier> depending on the collection, but should not be mapped to WorldCat. One of the discussion threads within the Metadata Working Group resulted in the suggestion of having either two records (one being a record based upon the source material, the other being an admin record about the digital surrogate) tied to a single digital object or the other solution of simply adding SOME metadata fields at the end of each record with SOME of the admin data about the digital surrogate, with nothing mapped and hidden, BUT SEARCHABLE IN CONTENTdm, and only for admin use.

<b>Recommended ‘as Appropriate’ Elements</b>	
<i>Element Name</i>	Source
<i>Definition</i>	Original object which the digital surrogate represents
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	Source
<i>Repeatable?</i>	Yes
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Prefer use of free text description incl., Collection Name, Accession Number, Physical Dimensions for graphic materials and Repository information</li> <li>• Prefer “Original Format” or other text prefix to qualify value.</li> <li>• WorldCat.org display mapping: ( MARC 786 [08])</li> </ul> <p><i>“Enter information about the original item before digitization as follows: genre of item: collection name, name of box, number of bin. Ex: 35 mm color slide: Larry Oglesby Collection, Morro Bay FT, bin #8” – Data Dictionary for Larry Oglesby Collection, LOC—Claremont Colleges Digital Library</i></p>	

<i>Element Name</i>	<b>Relation</b>
<i>Definition</i>	Named digital collection where a resource resides
<i>Controlled Vocabulary</i>	None
<i>DC Element Map</i>	<b>Relation.IsPartOf</b>
<i>Repeatable?</i>	Yes *
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• May contain URL of the digital collection(s) homepage.</li> <li>• * Some 'communities of practice' reference both the <i>Physical Collection</i> and the <i>Digital Collection</i></li> <li>• WorldCat.org display mapping: dcterms.Relation maps to MARC 787. maps to wc.Series and wc.SeriesTitle.</li> </ul> <p><i>"The described resource is a physical or logical part of the referenced resource." – University of Wisconsin Digital Library Data Dictionary</i></p>	

<i>Element Name</i>	<b>Location</b>
<i>Definition</i>	Spatial characteristics which describe the content of a resource
<i>Controlled Vocabulary</i>	TGN, GNIS, LCNAF
<i>DC Element Map</i>	Coverage.Spatial
<i>Repeatable?</i>	Yes
<p><i>Best Practice</i></p> <ul style="list-style-type: none"> <li>• Prefer use of standard controlled vocabularies and name authority sources.</li> <li>• Some 'communities of practice' reference geographic information system coordinates, such as those made available by Google Earth®</li> <li>• WorldCat.org display mapping: dcterms.coverage.spatial maps to MARC 522.</li> </ul> <p><i>“Currently recommended by the “Collaborative Digitization Project Dublin Core Metadata Best Practices” guide for use only ‘in describing maps, globes, and cartographic resources or when place or time period cannot be adequately expressed using the Subject element.’ Coverage spatial refers to the extent or scope of the content of the resource (e.g. place shown on a map or in a photograph, or geographic locations that are the topic of a manuscript), not the place of publication or digitization.” - Metadata Best Practices Guide, Western Michigan University Libraries</i></p>	

<i>Element Name</i>	<b>Time Period</b>
<i>Definition</i>	Era (ex: Colonial Period) which describes the content of a resource
<i>Controlled Vocabulary</i>	AAT
<i>DC Element Map</i>	Coverage.Temporal
<i>Repeatable?</i>	Yes
<p><b>Best Practice</b></p> <ul style="list-style-type: none"> <li>• Prefer ISO 8601 W3C Date/Time Format standard for dates.</li> <li>• Prefer use of standard controlled vocabularies.</li> <li>• WorldCat.org display mapping: dcterms.Coverage.Temporal maps to MARC 648</li> </ul> <p><i>“Usually a date or range of dates, but can be a named time period (e.g., Renaissance). Temporal coverage ‘refers to the time period covered by the intellectual content of the resource (CDP Dublin Core Metadata Best Practices (CDPDCMBP)),’ not the date of publication or digitization. It can refer to the time period shown in an image, the topic of a written manuscript, the time period covered in a series of diary entries, or, for art objects or artifacts, the date or time period of creation of the piece.” - Metadata Best Practices Guide, Western Michigan University Libraries</i></p>	

## For Further Reading

Digital Library Federation. 2007. Best Practices for Shareable Metadata.

<http://webservices.itcs.umich.edu/mediawiki/oai/p/index.php/ShareableMetadataPublic>

<sup>i</sup> Members of the original CONTENTdm Metadata Working Group, Aug-Dec 2009

Sheila Bair	Western Michigan University	<a href="mailto:bair@wmich.edu">bair@wmich.edu</a>
Dachun Bao	National Defense University	<a href="mailto:baod@ndu.edu">baod@ndu.edu</a>
Amalia (Molly) Beisler	University of Nevada Reno	<a href="mailto:abeisler@unr.edu">abeisler@unr.edu</a>
Megan Bernal	Depaul University	<a href="mailto:MBERNAL2@depaul.edu">MBERNAL2@depaul.edu</a>
Laura Capell	University of Southern Mississippi	<a href="mailto:laura.capell@usm.edu">laura.capell@usm.edu</a>
MingYu Chen	University of Houston	<a href="mailto:mchen15@uh.edu">mchen15@uh.edu</a>
Mei Ling Chow	Montclair University	<a href="mailto:chowm@mail.montclair.edu">chowm@mail.montclair.edu</a>
Kevin Clair	Penn State University	<a href="mailto:kmc35@psulias.psu.edu">kmc35@psulias.psu.edu</a>
Lee Dotson	University of Central Florida	<a href="mailto:ddotson@mail.ucf.edu">ddotson@mail.ucf.edu</a>
Mario Einaudi	The Huntington Library	<a href="mailto:meinaudi@huntington.org">meinaudi@huntington.org</a>
Allegra Gonzalez	Claremont Colleges Digital Library	<a href="mailto:Allegra_Gonzalez@cuc.claremont.edu">Allegra_Gonzalez@cuc.claremont.edu</a>
Deborah Green	University of Idaho	<a href="mailto:dgreen@uidaho.edu">dgreen@uidaho.edu</a>
Myung-Ja (MJ) Han	University of Illinois U-C	<a href="mailto:mhan3@illinois.edu">mhan3@illinois.edu</a>
Rachel Howard	University of Louisville	<a href="mailto:rachel.howard@louisville.edu">rachel.howard@louisville.edu</a>
Amanda A Hurford	Ball State University	<a href="mailto:aahurford@bsu.edu">aahurford@bsu.edu</a>
Andrea Kappler	Evansville Vanderburgh Public Library	<a href="mailto:andreak@evpl.org">andreak@evpl.org</a>
Deborah Keller	US Army	<a href="mailto:deborah.eb.keller@us.army.mil">deborah.eb.keller@us.army.mil</a>
Kate Kluttz	North Carolina State Library	<a href="mailto:kate.kluttz@ncdcr.gov">kate.kluttz@ncdcr.gov</a>
Lyn MacCorkle	University of Miami	<a href="mailto:LMaccork@miami.edu">LMaccork@miami.edu</a>
Sandra McIntyre	Mountain West Digital Library	<a href="mailto:sandra.mcintyre@UTAH.EDU">sandra.mcintyre@UTAH.EDU</a>
Gail McMillan	Virginia Tech	<a href="mailto:gailmac@vt.edu">gailmac@vt.edu</a>
Ann Olszewski	Cleveland Public Library	<a href="mailto:ann.olszewski@cpl.org">ann.olszewski@cpl.org</a>
Jennifer Palmentiero	SE NY Library Resources Council	<a href="mailto:jennifer@senylrc.org">jennifer@senylrc.org</a>
Kitty Pittman	Oklahoma State Library	<a href="mailto:kpittman@oltn.odl.state.ok.us">kpittman@oltn.odl.state.ok.us</a>
Gayle Porter	Chicago State University	<a href="mailto:gporter@csu.edu">gporter@csu.edu</a>
Gayle Spears	Atlanta University Center	<a href="mailto:gspears@aucr.edu">gspears@aucr.edu</a>
Jill Strass	St. Olaf University	<a href="mailto:strass@stolaf.edu">strass@stolaf.edu</a>
Glee M Willis	University of Nevada Reno	<a href="mailto:willis@unr.edu">willis@unr.edu</a>
Ling Wang	University of Illinois Chicago	<a href="mailto:lwang@uic.edu">lwang@uic.edu</a>
Noelia Ramos	Map Library of Catalonia	<a href="mailto:noelia.ramos@icc.cat">noelia.ramos@icc.cat</a>
Shilpa Rele	University of Miami	<a href="mailto:s.rele@miami.edu">s.rele@miami.edu</a>
Cheryl Walters	Utah State University	<a href="mailto:cheryl.walters@usu.edu">cheryl.walters@usu.edu</a>
Trashinda Wright	Atlanta University Center	<a href="mailto:twright@aucr.edu">twright@aucr.edu</a>
ZeeZee Zamin	Louisiana State University/LOUIS	<a href="mailto:zehra@lsu.edu">zehra@lsu.edu</a>

<sup>ii</sup> Moving towards shareable metadata by Sarah L. Shreeves, Jenn Riley, and Liz Milewicz  
**First Monday**, volume 11, number 8 (August 2006),  
URL: [http://firstmonday.org/issues/issue11\\_8/shreeves/index.html](http://firstmonday.org/issues/issue11_8/shreeves/index.html)



<sup>iii</sup>Han, Myung-Ja, Cho, Christine, Cole, Timothy W. and Jackson, Amy S. (2009) 'Metadata for Special Collections in CONTENTdm: How to Improve Interoperability of Unique Fields Through OAI-PMH', *Journal of Library Metadata*, 9: 3, 213 – 238. URL: <http://dx.doi.org/10.1080/19386380903405124>

### *Appendix A: Moving Towards Marketing with Metadata*

We have long recognized the need for effective marketing to increase discovery and delivery of digital collections. Enhancing descriptive metadata can move us in the right direction. Websites such as Flickr have adopted Web 2.0 social metadata standards such as tagging, in order to improve searchability for digital image material, and can leverage existing metadata to augment the user experience. There exists opportunity to further optimize descriptive metadata in otherwise well-aggregated digital collections. For example, there are many archival collections of historical material related to topics such as gold mining, railroad production, and other industries. The metadata used to describe these types of images can be quite literal and catalogers sometimes ‘miss the point’-- failing to apply such key, albeit at times colloquial, descriptors as “boomtowns,” “Gold Rush,” or “Wild West.”

While many controlled vocabularies are limited in their ability to incorporate this type of higher-level description, catalogers are encouraged to develop their own local controlled vocabularies based upon a convergence of subject terms (nouns, adjectives and verbs describing main topics) technical and style-based terms (unique image attributes such as image orientation, lens perspectives, and photographic techniques) and concept terms (ideas portrayed in an image). In WorldCat.org, the ability to create/name lists of items and apply social tags to items allows a high level of flexibility in accessing and managing content. Thus, the further integration of digital content into WorldCat.org represents a unique opportunity for the special collections community to begin experimenting with these types of terminologies-focused workflow tasks to increase discovery.

### *Appendix B Dates*

<b>Date type</b>	<b>DATE example</b>
Known year-month-day	2001-10-19
Known year-month	2001-10
Known year	2001
One year or another	1892 or 1893

---

Circa year-month	circa 1843-02
Decade certain	1970s
Before a time period	before 1867
After a time period	after 1867

*-Guidelines for Metadata Application in the Claremont Colleges Digital Library*

#### About Dates in CONTENTdm:

1. CONTENTdm supports the “date” data type and is consistent with the ISO standard yyyy-mm-dd, yyyy-mm and yyyy. You must use the date data type in order to provide searchable dates in CONTENTdm. However, many CONTENTdm users also provide a date field using the text data type. The fields shown in the latter five examples above would need to be configured as “text”.
2. To enter a range of years, use the following guidelines:
  - a. CONTENTdm Project Client- Use the yyyy-yyyy standard. Upon saving your metadata, the CONTENTdm Project Client will break out every date in the range.
  - b. CONTENTdm Web Add- Type every single year in the date range separated by semicolon-space.

*-Metadata Implementation Guidelines for North Carolina Digital State Document*

### *Appendix C: Metadata Schemas*

The following are examples of CONTENTdm metadata schemas that represent the vetted work of the MWG:

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab		add field
1	Title	Title	Text	No	Yes	No	Yes	No	move to ▼	edit   delete
2	Subject-Names	Subject	Text	No	Yes	No	No	Yes	move to ▼	edit   delete
3	Subject-Places	Subject	Text	No	Yes	No	No	Yes	move to ▼	edit   delete
4	Subject-Topics	Subject	Text	No	Yes	No	No	Yes	move to ▼	edit   delete
5	Description	Description	Text	No	No	No	No	No	move to ▼	edit   delete
6	Creator	Creator	Text	No	Yes	No	No	Yes	move to ▼	edit   delete
7	Publisher	Publisher	Text	No	Yes	No	No	No	move to ▼	edit   delete
8	Contributors	Contributors	Text	No	Yes	No	No	No	move to ▼	edit   delete
9	Date Original	Date	Date	No	Yes	No	No	No	move to ▼	edit   delete
10	Type	Type	Text	No	Yes	No	No	No	move to ▼	edit   delete
11	Format	Format	Text	No	Yes	No	No	No	move to ▼	edit   delete
12	Identifier	Identifier	Text	No	Yes	No	No	No	move to ▼	edit   delete
13	Source	Source	Text	No	Yes	No	No	No	move to ▼	edit   delete
14	Language	Language	Text	No	Yes	No	No	No	move to ▼	edit   delete
15	Rights & Usage	Rights	Text	No	Yes	No	No	No	move to ▼	edit   delete

For photographic collections (above) and archival collections (below)

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab		add field
1	Title	Title	Text	No	Yes	No	Yes	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
2	Title-Alternative	Title-Alternative	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
3	Creator	Creator	Text	No	Yes	No	No	Yes	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
4	Description	Description	Text	Yes	Yes	No	Yes	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
5	Abstract	Description-Abstract	Text	Yes	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
6	Table Of Contents	Description-Table Of Contents	Text	Yes	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
7	Transcription	Description	Full Text Search	Yes	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
8	Subject	Subject	Text	No	Yes	No	No	Yes	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
9	Publisher	Publisher	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
10	Contributors	Contributors	Text	No	Yes	No	No	Yes	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
11	Date.Original	Date	Date	No	Yes	No	Yes	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
12	Date.Digital	None	Date	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
13	Type	Type	Text	No	Yes	No	Yes	Yes	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
14	Format	Format	Text	No	Yes	No	Yes	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
15	Identifier	Identifier	Text	No	Yes	No	Yes	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
16	Source	Source	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
17	Language	Language	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
18	Relation	Relation-Is Part Of	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
19	Location	Coverage-Spatial	Text	No	Yes	No	No	Yes	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
20	Time Period	Coverage-Temporal	Text	No	Yes	No	No	Yes	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
21	Rights & Usage	Rights	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
22	Audience	Audience	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
23	Relation-References	Relation-References	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete
24	Digitization Specifications	None	Text	No	Yes	No	No	No	<input type="button" value="move to"/> <input type="button" value="v"/>	edit   delete

## *Appendix D Compound Objects*

### **Addendum on the treatment of compound objects with respect to OAI harvesting**

#### **Authors:**

Geri Bunker Ingram, MLIS  
OCLC Digital Collection  
Services

Myung-Ja "MJ" Han  
Metadata Librarian  
Assistant Professor of Library  
Administration  
University of Illinois at  
Urbana-Champaign

Sheila Bair, MLIS  
Metadata & Cataloging  
Librarian  
Western Michigan University

#### **Context:**

During the drafting of the Best Practices Guide version 1.7, discussion arose among the Metadata Working Group concerning the special case of sharing metadata from CONTENTdm Compound Objects. Users may employ diverse strategies for sharing metadata, regardless of the material type or formats that are assembled as compound objects, and regardless of the OAI-PMH harvester that will be employed. A request was made to attach a statement to the guide explaining the implications of metadata schema definition and CONTENTdm field configuration when a collection containing Compound Objects is destined to be harvested.

#### **CONTENTdm Definitions:**

**COMPOUND OBJECT** –any two or more CONTENTdm items that are logically and structurally assembled together. Each compound object comprises:

- A metadata record describing the object itself, (known as *object-level metadata*).
- A metadata record (known as *page-level metadata*) for each of the composite pages or items that make up the compound object.

**ITEM**—a single digital file and its affiliated metadata. In cases where there is metadata only—e.g., an image has not yet been scanned, the metadata is known as a “metadata only item”.

**COMPOUND OBJECT CLASSES:**

- Document—a series of related items
- Monograph—a series of items related in hierarchical fashion
- Post card—a series of exactly two items that may be displayed on one screen using the compound object viewer (by default labeled “front” and “back”);
- Picture cube—a series of exactly six items (designed originally for scans of realia)

**DOCUMENT DESCRIPTION (VIEW):** One of several views of the compound object available from the ‘compound object viewer’. The metadata that displays through this view is the *object-level* metadata.

**PAGE DESCRIPTION (VIEW):** One of several views of the compound object available from the ‘compound object viewer’. The metadata that displays through this view is the *page-level* metadata.

## Sharing metadata

With CONTENTdm, one can set a collection to be harvestable generally as long as the harvester is compliant, and one can also set a collection to be harvested by the Digital Collection Gateway specifically. With the former, CONTENTdm collection administrators can decide whether to enable the page-level metadata to be harvested. This is done in CONTENTdm Administration in the Server/Settings/OAI configuration function. With the Gateway, page-level metadata are never harvested, therefore the object-level metadata must be carefully considered. For other OAI harvesters, CONTENTdm collection administrators can decide whether and how fully to allow harvest of page-level metadata. Collection administrators should verify for every collection that the OAI configuration settings are correct for that particular collection. The implications for discovery and delivery vary depending upon the type of object at hand, and how well the Compound object -level (*metadata of the object itself*) is represented. Collection administrators must determine whether the document description (object-level metadata) is

enough for resource discovery/retrieval outside of the context of the native CONTENTdm environment. If a harvester provides direct links back to the object in its repository environment, (as in worldcat.org), and if the object-level metadata is extensive enough to allow discovery of the object, then end-users can link directly to the original collection and re-issue the specific search criteria to retrieve relevant objects with 'hits' highlighted on each page of each compound object across the collections on the server.

**Example--Enhancing discovery of buried information**

One of the CONTENTdm collections at Western Michigan University is a collection of Civil War diaries and letters assembled as compound objects. They employ the Library of Congress' "20 percent rule"<sup>iii</sup> for subject headings at the object level, except in cases of special information of interest to Civil War researchers. For instance, in all the diaries, subject headings at the object level contain the names of battles in which the diarist participated even though the description of the battle may comprise only a small percentage of the total text.

**Special considerations for textual transcripts**

The Document and Monograph classes of compound object in CONTENTdm are used mainly to handle text-rich objects. Searchable text transcripts are handled as metadata within a CONTENTdm schema. I.e., not only can every field of the metadata be made searchable, but above and beyond that, one field in each record may contain a searchable transcript of the text of the item. The *Full text search* field data type can be used for one field in each schema. In the case of a compound object, the object level metadata itself, and each of its item level metadata, may contain up to 128,000 characters in this Full text search field (often re-labeled "Transcript" in practice).CONTENTdm administrators decide whether to make this field harvestable or not, i.e., map the field to one of the DC elements.

*Appendix E: Consortium issues*

**Addendum on considerations for consortia using OAI harvesting tools; adding value from the members' point of view [draft 1.5]**

**Authors:**

Jason B. Lee  
Metadata Coordinator,  
WorldCat Digital Content  
OCLC Digital Collection  
Services

Lyn MacCorkle  
Digital Project Development  
& Repositories Librarian,  
Digital Initiatives & Resources  
University of Miami Libraries

Sandra McIntyre  
Program Director, Mountain  
West Digital Library

Gayle Porter  
Special Formats Catalog  
Librarian, University Library  
Chicago State University

Taylor Surface  
Senior Product Manager  
OCLC Digital Collection  
Services

Cheryl Walters  
Head of Digital Initiatives,  
Utah State University

**Context:**

A consortium is defined as an “agreement, combination, or group (as of companies) formed to undertake an enterprise beyond the resources of any one member.” During the drafting of the Best Practices Guide ver. 1.7, discussion arose among Metadata Working Group members concerning digital production & syndication challenges from a consortial viewpoint. A task group was formed in order to identify these [primarily workflow-oriented] issues in order to set forth an additional suite of recommended guidelines and to propose and communicate some specific resolutions in the WorldCat Digital Gateway environment.

**Considerations for Consortia:**

We have identified several overlapping core considerations for institutional members of a consortium using OAI harvesting tools in order to contribute digital content to a central server (outside of the institution). These core considerations, which may affect workflows at both the institution- and consortium-levels, include but are not limited to, metadata practices, communication strategy, and coordination of tasks.



*Note:* In the CONTENTdm-specific scenarios we reference here, there are two distinctly different issues present:

1. One CONTENTdm license is owned by the Consortium and shared among institutions.
2. One CONTENTdm license is owned as above, PLUS one or more CONTENTdm licenses are owned by member institutions.

### *Appendix F: Frequently Asked Questions regarding the Digital Collection Gateway*

(see also <http://www.oclc.org/gateway/>)

1. Does the Digital Collection Gateway only allow a single registration (username and password) per server, and do all of the libraries in the consortium have to share login information?

Modifying or issuing Gateway license KEYS to accommodate multiple users, as well as multiple repositories, is the recommended workflow for consortia. A Gateway license key may allow up to 50 separate usernames for individual control of collections. The consortia should have some centralized control where all of the metadata is managed. This enables many user logins to the Gateway, facilitated by coordination with the repository system administrator to allow the metadata to be shared by OAI. Currently, any existing CONTENTdm user that is part of a consortium can send an e-mail request to [contentdmsupport@oclc.org](mailto:contentdmsupport@oclc.org) and request that their key be modified to 'allow xx number of users'. Once the change is implemented, **each** library consortia member would be able to create a separate Gateway registration @ <https://worldcat.org/DigitalCollectionGateway/register.jsp> [see Figure A below].

The screenshot shows the 'WorldCat® Digital Collection Gateway' registration page. At the top, there is a navigation menu with 'Home', 'Reports', and 'Manage Account' tabs. Below this is a 'Registration Information' section with a form. The form contains the following fields: 'E-mail Address', 'Confirm E-mail Address', 'User Name', 'Password', 'Confirm Password', 'Phone Number', and 'Key'. A 'Register' button is located at the bottom right of the form.

- Figure A: Digital Collection Gateway online registration page

2. Is there a way that multiple people can manage a repository in Digital Collection Gateway? It appears that when an admin *delegates* a collection to another person, he/she can no longer see or manage it.

In the Digital Collection Gateway interface, only one person can manage a repository at a time, but that means only that one person has control of the editing. Any user can go into the Manage Account tab and assign a collection to themselves or someone else. In other words, if 'Jason L.' is out on vacation for a while, then 'Taylor S.' can assign the "entire repository" collection to himself and manage the metadata map and sync schedule.

3. The set up and configuration for WorldCat Sync tasks is located in the Server tab in the CONTENTdm Web Administration area, which may only be accessible to staff at the institution-level. Therefore, who would need to perform the initial setup to enable each collection to be uploaded to the Digital Collection Gateway?

We recommend that staff write policies and procedures to clearly describe administrative tasks in OAI harvesting, such as initial registration/set-up & log-in information, record sync schedule,

and selection of collections. These procedures need not be lengthy or laborious, but should be communicated and distributed to all institutions within the consortium. Both the consortium staff and institutional staff need to coordinate their workflows to make sure that initial setup has been completed for each institution that wants to have their records added to the Gateway.

4. Would staff from both the consortia as well as the member library need to 'keep track' of which collections have been uploaded to the Gateway?

We recommend that consortia staff develop a reporting structure and make information standard and easily visible across stakeholder groups. Consortia staff should keep an up-to-date account of management of digital records through the OAI harvesting tool, so that members are aware of which records have been uploaded and to prevent duplication of effort. The Gateway now provides a monthly activity summary for an entire repository which details the number of records added, updated, and deleted on a collection by collection basis. Staff from both groups also need to be in agreement as to which collections are 'ready' to be uploaded to the Gateway as metadata is revised or updated in the repository, in preparation for a manual 'push' or automated regularly-set upload. Gateway users also now have the ability to block certain records from their collections from being loaded to WorldCat even if they are "published" in CONTENTdm. Staff from each institution who works with digital collections should understand and follow the consortia policies for managing their records.

5. What happens if digital records from a member library are harvested by the consortia, and then both the consortia and the member library upload those records to WorldCat?

Digital Collection Gateway, OCLC's self-service OAI harvesting tool, recently added an important identifier de-duplication enhancement for digital content uploaded to WorldCat. The Gateway will now verify that no other records in WorldCat contain the same item URL which will reduce the introduction of this type of duplication in WorldCat. Best practice calls for a consortium to identify a digital content syndication coordinator and task him/her with

responsibility to coordinate contribution with an eye to quality and uniqueness, while minimizing duplication of effort among the membership.

6. In the consortial environment, what kind of metadata-specific practices do the partners need to agree upon?

Member libraries contributing digital content to a central server should agree on consistency in metadata-sharing practices by adopting a standard metadata style guide. Additionally, proprietary information such as rights, provenance, donor, etc., should be taken into consideration when determining what metadata is displayed locally, but not mapped for harvesting. For example, some consortia find it important to describe the process, equipment and specifications used to create the digital surrogate, although this information is often only useful within the local context. Mountain West Digital Library provides a non-Dublin Core field for this purpose (*Digitization Specifications*) which they adopted from the BCR/CDP DC Metadata Best Practices guide. Additionally, preservation data relating to archival master files are less useful in the aggregated environment, although a valuable best practice at the local level for migration purposes.

Consortia are also encouraged to develop a ‘common field properties’ schema that can be used flexibly for different types of materials such as theater programs, oral histories, and correspondence. Additionally, agreement and consistency (particularly in *level of granularity*) among the consortium on the intellectual content contained within digital collection records, particularly support the harvesting of shareable metadata related to:

- Subject & Genre information
- Geographic information
- Controlled vocabularies and name authorities
- *Required, Optional, and Recommended*, as well as *Searchable* designators
- Multiple field values vs. Repeating fields
- Display of qualifiers in the OAI environment
- *Original Date* vs. *Digitized* or *Published Date*
- Formatting conventions for Date, Language and other metadata fields

## *Appendix G: Digital Collection Gateway enhancements*

Announced enhancements to the Digital Collection Gateway

February 1, 2010 (CONTENTdm)

1. Record delete support – The Gateway will now mirror your activity in CONTENTdm. If you delete an item from your CONTENTdm repository the Gateway will remove the item from WorldCat to keep both systems fully in sync. Also, if you completely remove a collection from your CONTENTdm server you can ask the Gateway to remove all records associated with that collection from WorldCat.
2. Monthly activity report – The Gateway will now provide a monthly activity summary for your entire repository which details the number of records added, updated, and deleted on a collection by collection basis.
3. Identifier de-duplication – The Gateway will verify that no other records in WorldCat contain the same item URL which will reduce the introduction of this type of duplication in WorldCat.
4. Record blocking – You will now have the ability to block certain records from your collection from being loaded to WorldCat even if they are “published” in CONTENTdm.

May 16 (CONTENTdm and other OAI-PMH compliant repositories)

1. **Add “constant data”**. This new feature allows you to add a field of information to every record in your collection. With this you have the control to add context into a shared collection that may not be included locally. You can also include provenance information as you share your collections outside their local realm.

2. **Mappings for Qualified Dublin Core.** When you use CONTENTdm to synchronize your metadata with WorldCat you have the ability to use Qualified Dublin Core for more control. Thanks to the work of the Metadata Working Group the Gateway will now include default mappings for each of the Qualified Dublin Core elements supported by CONTENTdm. (The full set of 55 elements is planned to be included in the next major release of CONTENTdm.) Using the Gateway you can adjust the default mappings to optimize the WorldCat.org display.
3. **Naming “Entire Repository”.** This feature enhances your ability to work with multiple repositories. Some OAI repositories have a single collection (“Set” in OAI terms). This feature helps you keep track of multiple OAI repositories more easily. This advanced feature is being added to support general OAI support in the community.
4. **Enhanced Error/Warning Information.** The WorldCat Sync Report now includes summary information about each warning and error with links to detailed descriptions to guide your next steps.
5. **Duplicate URL detection.** When you add records to WorldCat, the Gateway will pre-search for other records in WorldCat with existing URLs that match your item URL. This new feature reduces duplication of records harvested from a primary site and secondary aggregations. The WorldCat Sync report will include a cross-reference OCLC number of the WorldCat record with the matching item URL.
6. **Updated Tutorial.** The tutorial, *Using the WorldCat Digital Collection Gateway*, is being updated to reflect new features added in the last couple of months.
  - a. **Field Splitting.** The Gateway allows you to specify a field delimiter and “split” the field into multiple, separate source fields. This allows you to work-around certain

display issues in the WorldCat.org display and sets up your data for new features coming soon.

- b. **Record Blocking**. The Gateway allows you to key on the contents of your source metadata fields to “block” records from being added to WorldCat. Use this feature to manage access/visibility of certain items within your collections.
- c. **Prefix/Suffix**. The Gateway allows you to add text to the beginning and/or end of a source metadata field. This allows you to do things like expand a DOI or Handle into a full URL.
- d. **Material type mapping**. You can assign a WorldCat material type icon based on the source metadata field contents. For example, you can use the contents of the dcterms:format field to assign a WorldCat material type icon.
- e. **Collection Cross-referencing**. Each time you upload a collection record to WorldCat the Gateway inserts a special URL in it to link it to all the item records from your collection. Likewise, each item record will include a collection record link in it when a collection record exists.

### **July 18, 2010 / version 2.1**

1. Metadata map editing allows same-name source fields to be mapped independently. If your source metadata has multiple entries in each record with the same metadata tag you can now map each entry separately. This allows multiple occurrences of a tag like dcterms:subject in each record to be moved to different locations in the WorldCat.org display.
2. Qualified Dublin Core and European Semantic Elements now supported. The Gateway now chooses the most detailed Dublin Core related metadata schema available from your repository as the default for each collection. The Gateway will select from European Semantic Element set, Qualified Dublin Core (DC Terms), OCLC DC Terms (available only

on CONTENTdm servers), and Simple Dublin Core (oai\_dc). This gives you more control over mappings to the WorldCat.org display.

3. Collection record creation (non-CONTENTdm repositories). The Gateway will utilize information provided by your OAI repository in the setDescription field to create a Collection Description in WorldCat. A collection description allows users to put your collection in context.
4. A-Z Contributor List. The Gateway will utilize information provided by your OAI repository to add information about your institution and your repository to the A-Z Contributor List. The Gateway will gather this information from the description field when your repository *Identify*'s itself.
5. Language mapping. The Gateway will utilize information in your metadata record's dcterms:language field to construct more precise MARC mappings. The Gateway will use the language information you supply in ISO 693-2,3 standard language tables and fill out more of the language details in the resulting record.
6. Performance improvements. The Gateway now has streamlined metadata map management so record previews, collection navigation, and map changes will all process much more quickly. The Gateway is also more sensitive to times when your repository may be responding slowly and will adjust timeouts based on the interaction with your repository.



7. Reports improved. The Gateway now provides more details in WorldCat Sync Reports. Harvest errors now provide detail about the OAI interaction and error message. Repository XML coding errors are also identified in report details. Harvests that are interrupted for an error now include more details on number of records processed before the interruption.

### **October 17, 2010 / version 2.3**

1. Description Identifier. User-assigned Resolution URL Gateway users can select the source metadata field that contains the URL of the digital item being described. This URL is referred to as the Resolution URL. The Gateway uses dcterms:identifier as the default source field for Resolution URLs.
2. User-assigned Thumbnail URL. Gateway users can select the source metadata field that contains the URL of the thumbnail of the digital item being described. This is referred to as the Thumbnail URL. The WorldCat.org interface will soon be adapted to include Thumbnail URLs in its display. The Gateway does not have a default for Thumbnail URLs for non-CONTENTdm repositories.  
This feature shows up as a yellow box around the Thumbnail section of the WorldCat.org display.
3. WorldCat Persistent Identifiers (CONTENTdm only) "As records are uploaded to WorldCat the digital items described by the records will be assigned a WorldCat Persistent Identifier (WPI). The WPI is a short URL of the form "http://worldcat.org/oclc/<OCLC#>/viewonline" which can be used for citation of digital items. WPI's provide a resolver service for CONTENTdm digital items that is stable even as the digital item is moved from collection to collection or server to server. All that is required of the server administrator is that they sync collections with WorldCat and sync the OCLC#'s into the metadata of their CONTENTdm items.  
This feature shows up as a yellow box in the Find a Copy Online section of the WorldCat.org display."
4. Apply a map to multiple collections & sync w/o preview. "From the Manage Account tab the user can select multiple collections from the collection list and sync them with

WorldCat. This provides a new workflow for users with lots of collections in their repository. Either create a map for one collection, then apply it to all, or apply the default map to all collections. This feature shows up on the Manage Account page in each Repository Section's list of Active Collections, select Map/Sync Selected."

5. Additions to Show Unmapped Fields. "Journal information details are now added to Show Unmapped Fields. The user can assign to Journal Title, Volume, and Issue."

#### Bug Fixes

1. "Centering the collection list on "active" collection" "When a user's repository has a long list of sets the Gateway centers the display on the last-edited collection if the user has left the Collection List page (e.g., the Home tab)."
2. WorldCat Sync Report - Links from error detail to error summary Error reports now have correct links from the error detail section to the error summary section