

**Colorado State University Libraries
Water Resources Archive Data Dictionary
Version 2.0**

by

Nancy Chaffin Hunter

and

Patricia J. Rettig

**Colorado State University Libraries
Fort Collins, Colorado
May 29, 2012**

Introduction

This data dictionary was initially developed to govern the metadata for the National Endowment for the Humanities (NEH) grant to digitize materials in the Water Resources Archive of Colorado State University Libraries Archives and Special Collections (WRA) for the Western Waters Digital Library (WWDL) [2007]. After considering the scope of the Water Resources Archive, and the potential for additional digitization projects related to water, this data dictionary was broadened to include all materials from the Water Resources Archive.

The WWDL has expressed interest in harvesting metadata for any water-related digitization project at CSU, even if outside the NEH/WWDL grant. The WWDL Metadata Committee issued Field Recommendations (see Appendix A) for participants in the NEH/WWDL. Those recommendations are advisory to this data dictionary.

This data dictionary was originally constructed to conform to the CSU Core Data Dictionary version 1.1 (July 2008). The new version (2.0) includes updates conforming to the CSU Core Data Dictionary version 2.0 (May 30, 2012) as well as additional instructions and examples following over three years of implementation.

Major changes include a change in series/subseries numbering (from "I.A" to "1.1", etc.); etc. (add detail later...)

Table of Contents

Data Dictionary Structure	5
Contributing Institution	6
Contributor	7
Coverage	8
Creator	10
Date	12
Description	14
Format	15
Identifier	16
Language	19
MetadataSchema	20
Publisher	21
Relation	22
Rights	24
Source	25
Subject	27
Technical Metadata	29
Title	30
Type	32
Appendix A. WWDL Field Recommendations	33

Notes:

1. The WRA digital items will be stored in the DigiTool software currently being used for the CSUL-IR. This data dictionary assumes the functionality and limitations of DigiTool. Enhancements to DigiTool, or the use of a different digital asset management system, may require modifications to the metadata and this data dictionary.
2. The WRA will contain some digital objects that comprise the contents of a folder within a collection. Metadata will be at either the folder level or the item level, per the decision of the project manager.
3. Details on Obligation and Repeatability are included in the discussion of each element and its refinements.
4. The WRA digital collection will include complex objects, and the structure and description will use METS or DigiTool's complex object when the objects are ingested. The WRA complex objects will conform to the WRA DD core at the METS object level. Metadata at the item level will be determined by the project manager.
5. If the element may be used in page/item-level metadata for METS/complex objects, it is so noted in the Element Attribute.

6. For the purposes of final decision making regarding the Water Resources Archive metadata, the Head Archivist of the Water Resources Archive is the project manager. At the time that this data dictionary was written, the Head Archivist was Patricia J. Rettig and the Metadata Librarian was Nancy J. Hunter.
7. By using element names based on Dublin Core (DC), and by mapping to DC, the metadata is prepared for OAI harvesting, including the Multi-Site Harvester of the Western Waters Digital Library, or as MODS XML files.
8. When there is a MARC record in Sage or other system for the original resource, the MARC record should be used as the basis for the metadata for the digital object (via copy/paste).

Data Dictionary Structure

The Water Resources Archive Data Dictionary provides the following attributes for the metadata elements:

Element Attribute	Description
Element Name	The unique name that identifies the element.
Standard Referenced	The metadata standards consulted: <ul style="list-style-type: none"> ○ <i>CSU Core v. 2.0</i> ○ <i>Field Recommendations</i> by the WWDL Metadata Committee (2007 revision)
Obligation	Specified within the discussion of each element: <ul style="list-style-type: none"> ○ Mandatory ○ Mandatory if applicable
Repeatable	Specified within the discussion of each element: <ul style="list-style-type: none"> ○ Repeatable ○ Non-repeatable
Definition	A statement that represents the concept and essential nature of the term.
Comment	Additional information about the term or its application as applied in the WRA context.
Refinements	Valid qualifiers for the element.
Schema	Valid schema to be used in the element.
Audience	Lists the intended audience for the element: <ul style="list-style-type: none"> ○ System ○ Manager (curator, repository manager) ○ Staff User ○ End User
Complex Object	Indicates whether the element may be used in the page/item-level metadata for a complex object. Instructions on how the element is to be used for the page/item-level metadata are in the Input Guidelines.
Simple DC Mapping	The DC element to which the WRA DD element maps for metadata sharing via OAI and Multi-Site harvesting. In some instances, qualified DC will be used.
DC-XML Mapping	The DC-XML tag to which this element maps for ingest. See CSU Core.
MARC Mapping	The MARC field to which this element maps for ingest. See CSU Core.
Full View/Label	Whether this element is displayed in the Full View and the label assigned to this element (whether Full View or Object View). See CSU Core.
Input Guidelines	Provides guidance about entering and encoding values for the element and its refinements.
Examples	Instances of how the element is used.

Contributing Institution

Element Name	Contributing Institution
Standard Referenced	WWDL Field Recommendations
Obligation	Mandatory
Repeatable	Yes
Definition	<i>Contributing Institution</i> is intended to aid in the management and preservation of metadata records in a shared environment by identifying the provenance of records and digital objects.
Comment	Used to identify CSU Libraries as the source of metadata contributed to the WWDL.
Refinements	None
Schema	None
Audience	Manager, End User
ComplexObject	No
Simple DC Mapping	None

Input Guidelines:

1. Colorado State University. Libraries (authorized form of the name) will be entered into the ***Contributing Institution*** element.

Examples:

- None given; refer to Input Guidelines above.

Contributor

Element Name	Contributor
Standard Referenced	CSU Core; WWDL Field Recommendations
Obligation	Mandatory if applicable
Repeatable	Yes
Definition	An entity responsible for making contributions to the resource.
Comment	The person(s) or organization(s) who made significant intellectual contributions to the resource but whose contribution is <i>secondary</i> to any person(s) or organization(s) already specified in a <i>Creator</i> element. Examples: editor, transcriber, illustrator, etc.
Refinements	None
Schema	LCAF http://authorities.loc.gov/ [for authorized form of a contributor's name], AACR2 [for instructions on establishing a contributor's name for those names not included in the LCAF]
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Contributor

Input Guidelines:

1. Enter each contributor in a separate *Contributor* element.
2. The form of name should be used consistently in all occurrences, across all resources in the CSUL-IR. The LCAF will be used to determine the established form of a name of a person or entity. If no authority record exists, the Metadata Librarian will consult Sage, CSUL-IR, or other locally developed lists for previously used form of name. Care should be taken to make the name unique and consistent in the database, using AACR2 guidelines for resolving conflicts.
3. Enter personal names as last, first middle, dates as appropriate. Enter corporate body names as established or according to AACR2.

Examples:

- **Contributor:** Hoover, Herbert, 1874-1964. (Folder in the Delph E. Carpenter Papers contains correspondence to/from Herbert Hoover.)

Coverage

Element Name	Coverage
Standard Referenced	CSU Core; WWDL Field Recommendations
Obligation	Mandatory if applicable
Repeatable	Yes
Definition	The spatial or temporal topic of the original resource, the spatial applicability of the original resource, or the jurisdiction under which the original resource is relevant.
Comment	<ul style="list-style-type: none"> o Temporal period may be a named period, date, or date ranges. o Use only when the contents of the object are clearly identified as to place or date.
Refinements	<p>REQUIRED</p> <p>Coverage.Spatial -- Spatial characteristics of the intellectual content of the resource.</p> <p>Coverage.Temporal -- Temporal characteristics of the intellectual content of the resource.</p>
Schema	LCSH for place names and time periods TGN (Getty Thesaurus of Geographic Names) ISO
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Coverage

Input Guidelines:

1. Use as a refined element. Use separate elements for each place or time period.
2. **Coverage.Spatial** will be named place. A jurisdiction may be used. Use LCSH for named places; if no LCSH exists, use TGN. If the place does not exist in either thesaurus, the Metadata Librarian will create the name using the LCSCM guidelines for creating place names.
3. **Coverage.Spatial** may include watersheds, river basins, bodies of water, cities, counties, or states, at the discretion of the project manager. Consult relevant finding aids before LCSH.
4. Temporal characteristics include those aspects of time that relate to the intellectual content of a resource and not its lifecycle.
5. To represent a range of dates in **Coverage.Temporal**, separate the ranges with space hyphen space (-). Enter the dates as YYYY - YYYY or YYYY/MM - YYYY/MM or YYYY/MM/DD - YYYY/MM/DD. To represent two or more specific dates, rather than a range, in **Coverage.Temporal**, separate the dates with space semicolon space (;). Enter the dates as YYYY ; YYYY. The project manager may choose not to represent gaps.
6. Circa dates or approximate dates may be assigned by the project manager for some items. Represent these dates as YYYY? or YYYY? - YYYY? as assigned.
7. If the date of the image is listed as “undated” in the finding aid, do not enter a value in **Coverage.Temporal** unless otherwise directed by the project manager.
8. **Subject** may also contain either geographical or spatial coverage if warranted by LCSH, i.e., the time period is established in LCSH and/or the subject term may be subdivided

geographically. If that is the case, repeat the temporal or geographic element(s) of the LCSH in the appropriate refinement of the *Coverage* element.

Examples:

- *Coverage.Spatial*: Larimer County (Colo.) [For a map of Richards Lake in Larimer County, Colo.]
- *Coverage.Temporal*: 1912 - 1922 [For a folder of correspondence beginning in 1912 and ending in 1922]
- *Coverage.Temporal*: 1925 [For a map dated 1925]
- *Coverage.Temporal*: [blank] [For an image listed in the finding aid as undated]
- *Coverage.Temporal*: 1920? [For an image listed in the finding aid as c. 1920]
- *Coverage.Temporal*: 1920 - 1974 ; 1980 [For a folder with a range and a gap]

Creator

Element Name	Creator
Standard Referenced	CSU Core
Obligation	Mandatory if applicable
Repeatable	No
Definition	The <i>Creator</i> element contains the name(s) of the person(s) or entity(ies) primarily responsible for the intellectual or artistic content of the digital resource.
Comment	The creator of original objects cannot always be known. When the creator is known, this element is mandatory.
Refinements	None
Schema	LCAF http://authorities.loc.gov/ [for authorized form of a creator's name], AACR2 [for instructions on establishing a creator's name for those names not included in the LCAF]
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Creator

Input Guidelines:

1. An archival folder containing multiple items will usually not have a creator of the folder. When creating metadata at the folder level, do not record the creator(s) of the content in the *Creator* element. At the discretion of the project manager, those individuals or entities may be entered in the *Contributor* element.
2. If multiple people or entities are equally responsible for the content of the resource, each will be listed in the *Creator* element, separated by space semicolon space (;).
3. If there are mixed responsibilities, the *Creator* element should be used for the primary responsibility and the *Contributor* element should be used for the secondary responsibility.
4. The form of name should be used consistently in all occurrences, across all projects. Use the Library of Congress Authority File (LCAF) to determine the established form of a name of a person or entity. If no authority record exists, consult Sage, the CSUL-IR, or other locally developed lists for previously used form of name. Care should be taken to make the name unique and consistent in the database, using AACR2 guidelines for resolving conflicts.
5. Enter personal names as last, first middle, dates as appropriate. Enter corporate body names as established or according to AACR2.

Examples:

- *Creator:* Collins, Caspar Wever, 1844-1865 [The artist responsible for the Camp Collins (C.T.) map, as established in the LCAF.]
- *Creator:* Colorado State University. Libraries [The entity responsible for creating the CSU Libraries Policies and Procedures Manual, as established in the LCAF.]

- **Creator:** Mahlman, Jerry David [Author of several monographs on meteorology cataloged in Sage, used as the author. Author's name is variably listed on the title pages as J.D. Mahlman and Jerry D. Mahlman. Name not established in the LCAF.]

Date

Element Name	Date
Standard Referenced	CSU Core ; WWDL Field Recommendations
Obligation	Mandatory when known
Repeatable	No [except within refinements as detailed below]
Definition	The date that the digital resource was created or that the original resource was created.
Comment	Some of the described digital resources will be digital surrogates of physical objects, such as a scanned image of a photograph, a digital photograph of a poster, or a digitized audio file from an audio cassette recording. Other digital resources will be 'born digital', such as an electronic newsletter in Portable Document Format (PDF). In the first case, there is a <i>Date.Original</i> and a <i>Date.Digital</i> . For digitized resources, the <i>Date.Original</i> will be the date the original analog resource was created, and <i>Date.Digital</i> will be the day the digital surrogate of the original analog resource was created. For 'born digital' resources, if the original digital object is used, <i>Date.Original</i> and <i>Date.Digital</i> will be the same; if the original digital object is reformatted and the reformatted version is used, <i>Date.Original</i> will be the date the original digital object is created and <i>Date.Digital</i> will be the date the reformatted version is created.
Refinements	REQUIRED (Use only as a qualified element) <i>Date.Digital</i> : The date that the digital resource was created. <i>Date.Original</i> : The date that the original resource was created.
Schema	ISO 8601 as defined by the World Wide Web Consortium profile at http://www.w3.org/TR/Note-datetime.html and as explained in Input Guidelines
Audience	System, Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	<i>Date.Digital</i> to Date.Created (maps to a qualified DC element due to DigiTool requirements) <i>Date.Original</i> to Source

Input Guidelines:

1. *Date* must always be used with a refinement. At present, only the refinements *.Digital* and *.Original* are part of the CSU Core.
2. Per the ISO 8601 standard, *Date.Digital* should be formatted as: YYYY-MM-DDThh:mm:ss.sTZD, where YYYY is the four-digit year, MM is the two-digit month, DD is the two-digit day of the month, T notes the beginning of time notation, hh is the two-digit hour based on a 24-hour clock, mm is the two-digit minute, ss is the two-digit second, .s is the single digit 1/10th of a second, and TZD is the time zone expressed as relates to Coordinated Universal Time (UTC).
3. *Date.Digital* will be extracted from the digital resource during the ingest process for digital resources in DigiTool.

4. When the date of the original artifact is known, the date in ***Date.Original*** will conform to the ISO 8601 to the extent that the date is known, i.e., YYYY-MM-DD, or YYYY-MM, or YYYY, where YYYY is the four-digit year, MM is the two-digit month, and DD is the two-digit day of the month. Further information about the hour, minute, second of creation is not necessary. Represent a range of dates as YYYY - YYYY.
5. To represent two or more specific dates, rather than a range, separate the dates with space semicolon space (;). Enter the dates as YYYY ; YYYY. The project manager may choose not to represent gaps.
6. When the date of the original artifact is unknown, a value of “undated” (without the quotation marks) will be entered in ***Date.Original***.
7. Circa dates may be assigned by the project manager for ***Date.Original***, either at the item level or in the finding aid. Represent circa date ranges as 1890? - 1903? and single circa dates as 1902?
8. For complex objects, if the date is the same for each item, put it only at the folder level. If items have different dates, put them at the item level in the title field, following a comma space (,).

Examples:

- For a JPEG 2000 (JP2) access image scanned from a handwritten letter:
 - ***Date.Digital***: 2007-07-03 10T14:09:38-6:00 [Digital creation date extracted from the JP2 when ingested into DigiTool.]
 - ***Date.Original***: 1910-07-14 [Date the letter was written.]
- For a JP2 access image scanned from a sheet map:
 - ***Date.Digital***: 2006-12-10T09:04:06-6:00 [Digital creation date extracted from the JP2 when ingested into DigiTool.]
 - ***Date.Original***: 1893 [When only the year in which the map was created is known.]
- For a JP2 access image scanned from a photograph:
 - ***Date.Digital***: 2006-12-10T09:04:06-6:00 [Digital creation date, extracted from the JP2 when ingested into DigiTool.]
 - ***Date.Original***: 1902? [An approximation of the year the original analog photograph was taken.]
- ***Date.Original***: 1920 - 1974 ; 1980 [For a folder with a range and a gap]

Description

Element Name	Description
Standard Referenced	CSU Core ; WWDL Field Recommendations
Obligation	Mandatory
Repeatable	No
Definition	A general description of the intellectual or artistic content of the digital resource.
Comment	The <i>Description</i> element will be a free-text account generally describing the contents of the digital object. Descriptive comments about the original object that cannot be observed in the digital resource should be entered in the <i>Source</i> element.
Refinements	.Abstract and TableOfContents (Optional, Not Repeatable)
Schema	None
Audience	Manager, Staff User, End User
Complex Object	Yes
Simple DC Mapping	Description

Input Guidelines:

1. Enter brief descriptive text, remarks, and comments about the digital resource. This information can be taken from the resource itself, taken from the finding aid, or provided by the project manager. End the field with a period only if a full sentence.
2. Enter here specialized information not included in other elements; for example, description, technique, and distinguishing features if observable in the digital resource and inscriptions. Features not observable in the digital resource should be included in the *Source* element.
3. **For Complex Objects:** If providing *Description* at the page/item-level for complex objects, describe only the intellectual or artistic content of the digital object itself, not its parent.

Examples:

- **Description:** Articulation of litigation's potential impacts on upper basin; constitutionality of Colorado River Compact [Free text description of the contents of a folder of correspondence.]
- **Description:** Section 25 T12N R70W. Scale 1 in. = 300 ft. Claim made for irrigation, source is Haygood Cannon Creek. Surveyed and filed with Larimer County and State Engineer in 1913. [Taken from the finding aid for a digitized map.]

Format

Element Name	Format
Standard Referenced	CSU Core ; WWDL Field Recommendations
Obligation	Mandatory
Repeatable	No (without refinements)
Definition	The digital manifestation of the resource.
Comment	Use <i>Format</i> to record the Internet Media Type (MIME Media Type) of the digital resource as registered with Internet Assigned Numbers Authority (IANA). Use the refinement <i>Format.Extent</i> to record the size. If the digital resource requires software or hardware external to the resource itself, record that information in a <i>Relation.Requires</i> element.
Refinements	<i>.Extent</i> The size of the digital resource expressed in bytes or playing time in hours, minutes, and seconds. (Mandatory, Not Repeatable).
Schema	MIME Media Type at http://www.iana.org/assignments/media-types/
Audience	System, Manager, Staff User, End User
Complex Object	Yes
Simple DC Mapping	Format (<i>Format</i> and <i>Format.Extent</i>)

Input Guidelines:

1. The *Format* element is provided as part of the extracted technical metadata in DigiTool.
2. The extraction process in DigiTool will record the file size in bytes. NOTE: DigiTool labels this element "File Size" and does not use the caption 'bytes'.
3. For audio and video formats, manually list duration (playing time) in a separate *Format.Extent* field. The playing time should be a numeric value with a caption that is used to qualify the numeric value. Use hour(s), minute(s), and second(s) as captions.

Examples:

For a PDF document, the following would be required:

- *Format:* Application/pdf [MIME Type]
- *Format.Extent:* 45329 [File size in bytes]

For an image ingested as JP2, the following would be required:

- *Format:* Image/jp2 [MIME Type]
- *Format.Extent:* 45329 [File size in bytes]

For a master digital audio file in the aiff format (the master digital resource for the Germans From Russia audio interviews) the following would be required:

- *Format:* audio/aiff [this is not a registered MIME type, so this is a locally created MIME type]
- *Format.Extent:* 4 MB [File size extracted by DigiTool and displayed as megabytes; actual file size is 3989600]
- *Format.Extent:* 45 minutes, 1 second [Playing time, manually entered.]

- ***Relation.Requires:*** Windows Media Player (Microsoft) or QuickTime Player (Apple)
[Software required to play the file.]

Identifier

Element Name	Identifier
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	Yes
Definition	The <i>Identifier</i> is a character string (alphabetic and/or numeric) that is <u>unique within the repository</u> and references the digital resource.
Comment	A character string or record number that clearly and uniquely identifies a digital resource. The <i>Identifier</i> element ensures that individual digital resources can be accessed, managed, stored, recalled, and used reliably.
Refinements	None
Schema	Handle.net ; Input Guideline #4
Audience	System, Manager, Staff User, End User
Complex Object	Yes
Simple DC Mapping	Identifier

Input Guidelines:

1. Use separate *Identifier* elements to enter multiple identifiers.
2. DigiTool will generate a unique *Identifier* when the digital object is ingested into the system that is akin to an accession number. The term used by DigiTool is PID.
3. A handle will be generated by the handle server for digital resources in DigiTool.
4. The following input guidelines apply to manually created and entered *Identifiers* that support logical collections in DigiTool:
 - a. The First *Identifier* will be used to identify the collection from which the resource is taken. It will consist of a structured string of characters comprised of 3 components; each component of the *Identifier* will be a fixed length:
 - i. First component: Fixed length of 4 characters. Value: the 4 character Collection Code for the specific collection in the Water Resources Archive from which the digital object was derived.
 - ii. Second component: Fixed length of 4 characters. Includes notation as to Series and Subseries as arranged in the finding aid at the project manager's discretion. To build the second component of the *Identifier*:
 1. The 1st character is S and stands for Series.
 2. The 2nd and 3rd characters are a number that stands for the Series number in the finding aid. Formerly, finding aids used Roman numerals. Now, use an Arabic numeral. Precede a single-digit Series number with a leading zero (0).
 3. The 4th character is the numeric designation assigned to the Subseries: 1, 2, 3, etc. For Series lacking Subseries arrangement use zero (0) for the 4th character.
 4. Use "aaaa" as a placeholder if series/subseries is not necessary.
 - iii. Third component: Fixed length of 6 characters. The 3rd component of the *Identifier* will be a 6-digit accession number. The beginning digit will not be a zero.

- b. The Second Identifier will be used to identify those resources funded by the NEH/WWDL grant. It will consist of a structured string of characters comprised of 3 components; each component of the *Identifier* will be a fixed length:
 - i. First component: Fixed length of 4 characters. Value: NWDL. This identifies the image as part of an NEH/WWDL project.
 - ii. Second component: Fixed length of 4 characters, consisting of the year of the grant.
 - iii. The 3rd component of the *Identifier* will be a 6-digit accession number. The beginning digit will not be a zero.
5. Two additional identifiers are required:
 - a. An *Identifier* element that is the exact file name of the digital object being ingested, including the extension
 - b. An *Identifier* element that is the exact file name of the digital master, including the extension.
6. For complex objects, identifiers are assigned only at the parent (folder) level.

Examples:

For a single PDF created from a folder of correspondence:

- WDEC00301.pdf
 - The file name for the PDF document being loaded into DigiTool.
- WDECS01A300001 (legacy identifier using Roman numerals and alphabetic designations for subseries)
 - WDEC stands for the collection: Papers of Delph E. Carpenter and Family, Water Resources Archive, Colorado State University
 - S01 stands for Series I: DEC Correspondence
 - A stands for Subseries A: Filed correspondence
 - 300001 is the 6-digit accession number.
- NWDL2007300001
 - NWDL stands for the NEH/WWDL grant
 - 2007 stands for the beginning year of the grant
 - 300001 is the 6-digit accession number.
- <http://hdl.handle.net/10217/1856>
 - [A handle that will be a shortcut to the actual URL:
http://digitool.library.colostate.edu:1801/webclient/StreamGate?folder_id=0&dvs=1221661347632~512]

Language

Element Name	Language
Standard Referenced	CSU Core
Obligation	Mandatory if applicable
Repeatable	Yes
Definition	The <i>Language</i> element contains the language(s) of the intellectual or artistic content of the resource.
Comment	Use of the <i>Language</i> element implies the language(s) in which the text is written. Images do not usually have a language unless there is significant text in a caption or in the image itself.
Refinements	None
Schema	ISO 639-2 http://www.loc.gov/standards/iso639-2/englangn.html English language names for languages
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Language (ISO 639-2 only)

Input Guidelines:

1. If a *Language* element is appropriate, create one *Language* element with the ISO-639-2 value and a separate *Language* element with the conventional full name of the language in English.
2. If the digital resource contains more than one language, use one *Language* element for the ISO-639-2 values for all relevant languages, separating each language value by space semicolon space (;), and a second *Language* element to list the conventional full names of the languages in English, separated by space semicolon space (;).
3. If special explanation is necessary to identify how language relates to the digital resource, add text to the *Description* element to describe the situation.

Examples:

- For a textual document in English:
 - Language: English
 - Language (ISO): eng
- For a textual document in English and Spanish
 - Language: English ; Spanish
 - Language (ISO): eng ; spa

MetadataSchema

Element Name	Metadata Schema
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	Yes
Definition	The <i>MetadataSchema</i> identifies the data dictionary (title and version or date) used to create the metadata.
Comment	The element <i>MetadataSchema</i> is recorded for quality assurance, to assist in evaluating the metadata in the future, and to aid in any future migration of the metadata.
Refinements	None
Schema	Water Resources Archive Data Dictionary v. 2.0.
Audience	Manager, Staff User
Complex Object	No
Simple DC Mapping	None

Input Guidelines:

1. Use the exact title following to record the *MetadataSchema* for this collection: Water Resources Archive Data Dictionary v. 2.0.
2. If changes are made to the data dictionary during the course of the project, the changes will be recorded in a separate data dictionary, with different version and date, and metadata built using that version will use the new version in *MetadataSchema*.
3. Each version will be maintained in the Institutional Repository. Superseded versions will be clearly labeled superseded, and their metadata will include a *Relation.IsReplacedBy* element.

Examples:

- None given; refer to Input Guidelines above

Publisher

Element Name	Publisher
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	Yes
Definition	The entity responsible for making the digital resource available.
Comment	None
Refinements	None
Schema	AACR2 form of the name for Colorado State University Libraries
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Publisher

Input Guidelines:

1. Colorado State University. Libraries (authorized form of the name) will be entered into the ***Publisher*** element.
2. Optionally, record other entities in a repeated ***Publisher*** element if they are responsible for making the digital resource available, including publishing the original resource in paper, as instructed by the Head Archivist.

Examples:

- Colorado State University. Libraries [Institution making the digital resource available via its Digital Repository.]
- Colorado State University. Dept. of Atmospheric Sciences [Original publisher of the print paper that was digitized.]

Relation

Element Name	Relation
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	Yes (within refinements)
Definition	This element contains a reference to a related resource, i.e., information necessary to describe, find, or link to a related resource.
Comment	<i>Relation.IsPartOf</i> and <i>Relation.IsFormatOf</i> are the only refinements defined for this data dictionary.
Refinements	REQUIRED (Use only as a qualified element) <i>Relation.IsPartOf</i> : identifies the archival collection, series [when applicable], and subseries [when applicable] in which the original object resides, and includes the URL link to the collection finding aid. If the original object is part of a publication series, use this element as well. <i>Relation.IsFormatOf</i> : identifies the original entry as recorded in the finding aid.
Schema	URL, See Input Guidelines for other exact language
Audience	End User, Staff User, Manager
Complex Object	No
Simple DC Mapping	Relation (all Refinements)

Input Guidelines:

1. Refinement is required for the *Relation* element. At present, only *.IsPartOf* and *.IsFormatOf* are defined by this data dictionary.
2. *Relation.IsPartOf* defines the digital object as part of a collection within the Water Resources Archive. Include Series and Subseries as part of the collection name when they exist.
 - a. Transcribe the Series and Subseries titles exactly, including punctuation, capitalization, and spelling. Exclude any dates associated with the Series or Subseries. Follow the citation for the collection with the URL to the finding aid.
 - b. This field should be formulated as follows:
 - i. **For items contained in a collection, with series and subseries:**
 1. *Finding aid title. Series number: Series title; Subseries number: Subseries title. Finding aid: URL to finding aid*
 - ii. **For items lacking subseries:**
 1. *Finding aid title. Series number: Series title. Finding aid: URL to finding aid*
 - iii. **For items lacking series:**
 1. *Finding aid title. Finding aid: URL to finding aid*
3. *Relation.IsFormatOf*: Record here the entry as listed in the finding aid. Include dates but exclude page counts when present. Transcribe the title exactly as it is given in the finding aid, including punctuation, spelling, capitalization, and any dates. Give all bibliographic information listed as part of the finding aid title including author's(s') name as it(they)

- appear(s). Do not record box or folder numbers for the item. When a folder falls under a category, list the category first, followed by a colon space (:), followed by the exact finding aid entry.
4. For items to be harvested by the Western Waters Digital Library, “wwdl” must be entered in an unqualified *Relation* field.
 5. If the digital object is associated with a publication series, use *Relation.IsPartOf*. Use space semicolon space between the series title and number. If there are two or more publication series, use separate fields..
 6. Note that *Source* is used specifically to describe the physical characteristics or conditions of the original. Do not repeat that information using the *Relation* element.

Examples:

Relation.IsPartOf

- Papers of Delph E. Carpenter and Family. Series I: DEC correspondence; Subseries A: Filed correspondence. Finding aid: <http://lib.colostate.edu/archives/findingaids/water/wdec.html> [The source collection for the folder from Archives and Special Collections.]
- Open file report ; no. 72 [Publication series title and number.]

Relation.IsFormatOf

- "Arizona vs. California" B. 1930-1931 [Title as recorded on the folder, including date.]
- Legislation: Code, W.E. Notes on ground-water problems in Colorado. undated [The category followed by the title of a published report as it appears in the finding aid, including date.]
- Contained in: [Maps]. 1897, 1949, undated [Name of folder that contains the original map titled “Map of the San Luis Valley,” which was digitized and assigned metadata as an individual object.]

Rights

Element Name	Rights
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	Yes
Definition	The <i>Rights</i> element contains a statement of copyright permission or special conditions for use of the digital resource.
Comment	The value of the <i>Rights</i> element is defined by the Coordinator of Archives and Special Collections and will consist of a URL link to the rights statement.
Refinements	None
Schema	URL
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Rights

Input Guidelines:

1. The ***Rights*** element will minimally be a hyperlink to the rights statement relevant to that resource.
2. Specific rights information may be added to the ***Rights*** element as free text when conditions warrant. Use the same field and separate with a space semicolon space (;).

Examples:

- <http://teaching.colostate.edu/guides/copyright/> [Link to the Colorado State University rights page]

Source

Element Name	Source
Standard Referenced	CSU Core
Obligation	Mandatory if Applicable
Repeatable	Yes
Definition	The analog resource from which the described digital resource is derived.
Comment	The Source element describes <u>physical characteristics</u> of the original analog resource from which a digital resource is derived, either in whole or in part. The <u>citation</u> of the original is entered in Relation.IsFormatOf . Some digital resources are “born digital” and derive from no pre-existing analog resource; in these cases, the Source element is not used.
Schema	None
Audience	Manager, Staff User, End User
Complex Object	Yes
Simple DC Mapping	Source

Input Guidelines:

1. Enter source information in order of importance. Use separate **Source** elements to enter multiple sources or *clearly separate each entry* by a semicolon and a space within an element. Usually there will be only one source from which the present digital resource has been derived.
2. Whenever possible, include a unique standard identifier such as an ISBN or ISSN. If no standard identifier exists, use a local call number, control number, accession number, or barcode. Identify the institution associated with such locally derived numbers; in most cases this will be “Colorado State University. Libraries.”
3. Clarify the nature of the relationship between the two resources by using an initial phrase such as “Excerpted from:”, “Original:”, etc. as determined by the Head Archivist.
4. When listing the number of items in the original, abbreviate as follows:
 - a. page(s) = p.
 - b. volume(s) = v.

Do not abbreviate other items counted, such as leaf (leaves), slides, etc. Leave a space between the number and the word or abbreviation. (See example 2 below.)
Alternatively, copy the conventions used in the finding aid.
5. Information about the condition of the physical object may be included here, or may be included in a separate project-defined element.
6. When giving physical dimensions, use the unit of measurement most appropriate for the resource.
7. Include notes about any materials that are part of the original object which were not scanned as part of the digital item. These include duplicate pages, irregular items, oversize items, and copyrighted items. (A target should be included in the digital object indicating the absence of these types of materials.)

8. If any pages/items that are part of a digital item have any confidential type markings, add a note: "Papers marked confidential are open for public access."
9. When both digital files and physical files are received from a donor but do not exactly correspond, note differences as necessary.
10. **For Complex Objects:** The physical format and dimensions for page/item-level metadata may be provided when there are differences between the multiple pages. Example: A book is digitized; all the textual pages are 23 cm. tall, but there are fold-out maps measuring 23 cm. tall by 30 cm. wide.

Examples:

1. For a digitized map:
 - **Source:** Original: sheet map; 24 in. x 36 in. [Physical description of the original analog resource that was digitized.]
 - **Relation.IsFormatOf:** Case #2031. Haygood Cannon Ditch [Number and title of the map in the Larimer County District Court Map Collection. Add dates if present in the finding aid]
 - **Relation.IsPartOf:** Larimer County District Court Map Collection. Finding aid: <http://lib.colostate.edu/archives/findingaids/water/wlcc.html> [The source collection for the image from Archives and Special Collections. Together with the **Relation.IsFormatOf**, the original map can be easily located when necessary.]
2. For a folder of digitized correspondence:
 - **Source:** Original: Folder of correspondence. 27 p. [Physical description of the original folder and its contents.]
 - **Relation.IsFormatOf:** "Arizona vs. California" B. 1930-1931 [Folder title of the file of correspondence in the Papers of Delph E. Carpenter and Family.]
 - **Relation.IsPartOf:** Papers of Delph E. Carpenter and Family. Series I: DEC correspondence; Subseries A: Filed correspondence. Finding aid: <http://lib.colostate.edu/archives/findingaids/water/wdec.html> [The source collection for the image from Archives and Special Collections. Used with **Relation.IsFormatOf**, the original folder can be easily located.]
3. For a folder of digitized slides:
 - **Source:** Original: 24 35 mm slides [Physical description of the original resource that was digitized.]
 - **Relation.IsFormatOf:** CNPP & ID [Central Nebraska Power & Irrigation District]. 1949-1956, undated [Folder title of the group of slides in the Papers of Maurice L. Albertson.]
 - **Relation.IsPartOf:** Papers of Maurice L. Albertson. Series 10: Audio/visual materials; Subseries 1: Slides. Finding aid: <http://lib.colostate.edu/archives/findingaids/water/wmla.html> [The source collection for the image from Archives and Special Collections. Used with **Relation.IsFormatOf**, the original folder can be easily located.]
4. For a folder of 32 slides that came with a CD of 50 scans of slides, some the same, some with no corresponding slides:
 - **Source:** Original: 32 35 mm slides and CD of scanned slides

Subject

Element Name	Subject
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	Yes
Definition	A term used to describe the content of the digital resource.
Comment	The values in this element refer to what the content of the resource is <u>about</u> or what it <u>is</u> , expressed by topical terms, headings, key terms, phrases, names, or other terms for significant people, places, and events, etc. The application of <i>Subject</i> is to be determined at the project level.
Refinements	None
Schema	Multiple controlled vocabularies: examples: Library of Congress Subject Headings (LCSH), Art and Architecture Thesaurus (AAT), Library of Congress Authority File (LCAF)
Audience	Staff User, End User
Complex Object	No
Simple DC Mapping	Subject

Input Guidelines:

1. Library of Congress Subject Headings and, when appropriate, LCAF personal or corporate names will be assigned in the *Subject* field.
2. Subjects and names will be taken from the finding aid and applied appropriately, when there are subjects or names assigned to the finding aid.
3. If LCSH terms have subdivisions (topical, geographic, temporal, or genre), the subdivisions will be separated by <space>--<space> rather than – (double hyphen instead of long dash) to accommodate DigiTool requirements.
4. If multiple LCSH/LCAF subjects are assigned to the resource, each will be entered into its own *Subject* field, i.e., only one subject/name will be entered into a single field.
5. Record genre terms in a *Subject* field, using the Art and Architecture Thesaurus. When multiple genre terms apply, record each in its own *Subject* field, i.e., only one genre term will be entered into a single field.
6. Names of rivers and watersheds should be listed in both *Subject* and *Coverage.Spatial* fields.
7. Do not use a period at the end of the LCSH, LCAF, or AAT term unless the term itself ends in a period.

Examples:

- *Subject:* Water -- Law and legislation [LCSH]
- *Subject:* Water resources development [LCSH]
- *Subject:* South Platte River Watershed (Colo.-Neb.) [LCSH]
- *Subject:* Arkansas River Watershed -- Maps [LCSH]
- *Subject:* Aerial photographs [AAT]

Technical Metadata

Standard Referenced	Z39.87, PREMIS
Obligation	Mandatory; individual elements are not specified in this Data Dictionary
Comment	Technical metadata is important for the long term management of digital resources. Virtually all technical metadata is embedded within the resource itself; however, the design model adopted by Z39.87 presumes that technical metadata will be stored outside the digital resource, and this Data Dictionary applies that model to all digital formats, not just still images. (To date, Z39.87 is the only nationally approved and published technical metadata standard and is restricted to digital still images.)
Audience	System, Manager
Complex Object	Yes
Simple DC Mapping	None

Input Guidelines:

1. Use DigiTool to extract technical metadata during the ingest process.
2. Only *Type*, *Format*, and *Format.Extent* are mapped to simple Dublin Core for OAI harvesting.

Title

Element Name	Title
Standard Referenced	CSU Core
Obligation	Mandatory
Repeatable	No
Definition	A name given to the digital resource.
Comment	Typically, a <i>Title</i> will be a name by which the digital resource is formally known or the name given to the digital resource by the creator or publisher. <i>Title</i> may also be an identifying phrase or name of the digital resource supplied by the metadata creator, project manager, or archivist. Neither <i>Title</i> nor <i>Title.Alternative</i> needs to be unique.
Refinements	<i>Title.Alternative</i> (Mandatory if Applicable, Repeatable) Any form of the title used as a substitute or alternative to the formal title of the digital resource.
Schema	None
Audience	Manager, Staff User, End User
Complex Object	Yes
Simple DC Mapping	Title (<i>Title</i> and <i>Title.Alternative</i>)

Input Guidelines:

1. Transcribe the title if one is available from the source or the finding aid. For more guidance in constructing titles, the project manager should consult established cataloging rules such as *Describing Archives: A Content Standard (DACCS)* (preferred) or *Anglo-American Cataloging Rules (AACR2)*. It is recommended that only **one** set of rules be used for water projects. Retain punctuation from the finding aid, including quotes and square brackets.
2. To aid in title sorting of results, remove any initial article and append to the end of the title, following a comma, for transcribed titles. Avoid using initial articles in supplied titles if at all possible.
3. If the title is taken from the finding aid, capitalize according to the finding aid. If the title is supplied, capitalize only the first letter of the title and proper nouns contained within the title. If the title is taken from the resource itself, capitalize per the original.
4. In general, use the punctuation provided with the title or, if the title is created by the project manager, use English language punctuation standards. The project manager may designate a standard style manual such as *Chicago Manual of Style*, *University Style Manual* (Colorado State University), etc.
5. For published materials, if the finding aid lists the author's name first, remove it and all bibliographic information listed so that only the title of the published work appears in the *Title* field.
6. The *Title* element is repeatable only with the refinement *Alternative*. The *Title* element should be used only once and all other *Title* elements should be expressed as *Title.Alternative*.
7. *Title.Alternative* can be used for other titles useful for access, i.e., caption title, former title, spine title, collection title, series title, artist's title, object name, translation of title, and other variations of the title.

8. File names, accession numbers, call numbers, or other identification schemes should be entered in the *Identifier* element.
9. When one of a set of folders is turned into a digital object, include the folder numbering information at the end of the title, following a period space (.).
10. **For Complex Objects:** At the item level, transcribe the title exactly and include the date after the title, followed by a comma space (,).

Examples:

- **Title:** “Arizona vs. California” L [Title from folder name in the Papers of Delph E. Carpenter and Family.]
- **Title:** Case #2031: Haygood Cannon Ditch [Title from Larimer County District Court Map Collection finding aid, constructed by treating the Case # as the title and the title of the map itself as a subtitle. Do not treat such constructed titles as a supplied title.]
- **Title:** Humphreys Dam [Colorado]. Folder 1 of 2 [Title of folder from Papers of Robert E. Glover followed by folder numbering information.]
- **For Complex Objects:**
 - **Title:** Glen Canyon Project [Folder-level title for folder of slides]
 - **Title:** Lake Powell, December 1965 [Item-level title of first slide]

Type

Element Name	Type
Standard Referenced	CSU Core; WWDL Field Recommendations
Obligation	Mandatory
Repeatable	No
Definition	The nature or genre of the content of the resource.
Comment	The <i>Type</i> element lists terms describing general categories, functions, genres, or aggregation levels for content. To describe the physical or digital manifestation of the resource, use the <i>Format</i> element.
Refinements	None
Schema	DCMI Type Vocabulary http://dublincore.org/documents/dcmi-type-vocabulary/
Audience	Manager, Staff User, End User
Complex Object	No
Simple DC Mapping	Type

Input Guidelines:

1. Images of written language are assigned as Text in *Type*.
2. Some digital resources may require more than one *Type*, such as a scanned page containing both a picture and text.
3. Digital representations of three-dimensional objects should be assigned “Image”, “Still Image”, etc. as appropriate, rather than “Physical Object”. The use of “Physical Object” is confined to databases describing the objects, not to their digital surrogates.
4. To conform to WWDL Metadata Committee Field Recommendations, Still Image and Moving Image must also be represented in their camelback form: StillImage, MovingImage.
5. Record all DCMI Types and all WWDL Types in the same *Type* field. If there is more than one type, separate each type by space semicolon space (;).

Examples:

- Still Image ; Image ; StillImage ; Map ; Text [For a map that will be contributed to the Western Waters Digital Library, which requires that Image be used in conjunction with Still Image, that Map be assigned to cartographic resources, and that StillImage be represented in camelback form. The map also contains English language text that adds to the intellectual content of the resource.]

Appendix A
**WESTERN WATERS DIGITAL LIBRARY
METADATA COMMITTEE
FIELD RECOMMENDATIONS**

**WESTERN WATERS DIGITAL LIBRARY
METADATA COMMITTEE
FIELD RECOMMENDATIONS**

OVERVIEW: The WWDL Metadata Committee was established as prescribed in the IMLS grant proposal to explore and address the issues that challenge uniform application of metadata standards in a collaborative project. Initially, the Committee focused on descriptive metadata issues that affect searching in the centralized site.

The recommendations of this committee were reviewed by a second Metadata Committee formed for the NEH grant proposal in September/October 2007. This review was deemed necessary because of the increased importance of collecting technical metadata, the release of a revised *Western States Dublin Core Metadata Best Practices (WSDCMBP)* version 2.1.1 in September 2006, and the time that has elapsed between receipts of the two grants.

The current Metadata Committee has approved the original recommendations for nine elements based on the *WSDCMBP*, as well as requirements for metadata for compound objects. The recommendations are linked below; those with an asterisk are the elements that have been identified as required for searching the aggregated site. Those elements with an additional asterisk have been identified as required for capture of technical metadata.

- Title*
- Creator*
- Subject*
- Description*
- Date Original*
- Date Digital**
- Type*
- Format**
- Identifier*
- Compound Objects (special instructions)

The current Metadata Committee recommends that the following elements also be required:

- Rights Management
- Digitization Specifications**
- Contributing Institution
- Contributor (when applicable)
- Source (when applicable)

The additional four elements defined in the *WSDCMBP* are **highly** recommended but not required:

Publisher
Language
Relation
Coverage

The committee also recommends use of the following element for the geographic area of coverage of the resource:

Place Name (when applicable)

This document should be used in conjunction with the *WSDCMBP*.

r:10/26/07 (deb)

Element/Field **TITLE**

Mandatory

Repeatable

DC Definition: The name given to the resource

DC Comment: Typically, a *Title* will be a name by which the resource is formally known.

WS Comment: The name given to the resource by the creator or publisher; may also be an identifying phrase or name of the object supplied by the contributing institution.

Instructions:

Apply the *WSDCMBP* instructions for transcribing or creating a title for the digital resource.

The Metadata Committee recommends excluding initial articles when transcribing titles into the

Title field, to aid in the proper ordering of retrieval sets. If local practice requires initial articles,

titles will not display in their proper alphabetical sequence in the MultiSiteServer. Note that this includes titles with foreign initial articles.

For compound objects:

For the document level metadata, follow *WSDCMBP* instructions for transcribing/creating the title.

For page level metadata, *Title* is usually the filename, truncated to leave only the last bit, such as "page001". It is imported, truncated, and inserted in the *Title* element automatically through CONTENTdm's compound object creator.

Element/Field **CREATOR**

Mandatory (if available)

Repeatable

DC Definition: An entity primarily responsible for making the content of the resource

DC Comment: Examples of a *Creator* include a person, an organization, or a service. Typically, the name of a *Creator* should be used to indicate the entity.

WS Comment: A person or entity primarily responsible for creating the intellectual content of the resource. Examples of creators include authors of written documents, artists, photographers, collections of natural specimens or artifacts, organizations that generate archival collections, etc.

The form of the name should be under authority control and entered in inverted order. The last name first followed by a comma and the first name, middle name, etc.

Instructions:

Apply the *WSDCMBP* instructions for transcribing or identifying the creator of a resource.

For compound objects:

Creator should be added only to the document level metadata.

Do not add *Creator* to page level metadata. The *Creator* field should be left blank.

Revised 10/18/2007

Element/Field **SUBJECT**

Mandatory

Repeatable

DC Definition: A topic of the content of the resource

DC Comment: Typically, *Subject* will be expressed as keywords, key phrases, or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

WS Comment: What the content of the resource is about or what it is, expressed by headings, keywords, phrases, names, or other terms for significant people, places and events, etc. A classification code may also be assigned.

WWDL Instructions:

Use only terms or phrases from a controlled vocabulary in the *Subject* element. The preferred controlled vocabulary is *Library of Congress Subject Headings*.

A classification scheme such as Dewey DC or LC Classification may be use **in addition** to the subject terms.

For compound objects:

Subject should be added only to the document level metadata.

Do not add *Subject* to any page level metadata. The *Subject* element should be left blank.

Element/Field **DESCRIPTION**

Mandatory

Repeatable

DC Definition: An account of the content of the resource

DC Comment: **Description** may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

WS Comment: Descriptive comments about the original object that cannot be observed in the digital resource should be entered in the **Source** element.

Instructions:

Apply the *WSDCMBP* instructions for adding a description of the contents of the resource in the **Description** field.

For compound objects:

Typically, the **Description** is applied to both the document level metadata and the page level metadata.

When possible, include the full text of the page in the **Description** field for page level objects. This may be transcribed or scanned and created into text, depending on the circumstances.

Element/Field **DATE ORIGINAL**

Mandatory (if applicable)

Repeatable

DC Definition: none (There is no DC equivalent for this element)

DC Comment: none (There is no DC equivalent for this element)

WS Comment: Creation or modification date for the *original* resource from which the digital object was derived or created.

WWDL Instructions:

This date is not mandatory for ‘born digital’ resources.

This field maps to the DC DATE in local CONTENTdm collections.

For all new metadata records created for the WWDL, contents of the *Date Original* and *Date Digital* fields must conform to the W3C-DTF (Date and Time Formats) at <http://www.w3.org/TR/NOTE-datetime.html>. This is the YYYY-MM-DD, or YYYY-MM, or YYYY format.

For all existing metadata records, the Metadata Committee is recommending that institutions correct any *Date Original* or *Date Digital* fields that do not comply with the W3C-DTF. This includes items with circa dates and dates using other formatting. It appears that the most common error appearing in the *Date* fields is ca. for circa dates. Circa dates can still be expressed within the metadata; however, it should be done utilizing the DCMI Period Syntax (1998?-2002?). For single circa dates, the following syntax should be utilized: 1998?. For more information on how the *Date* fields should be formatted, refer to the *Western States Dublin Core Metadata Best Practices, Version 2.0*. However, since searchability is not currently affected by the presence of non-W3C-DTF formatted data, the Metadata Committee will not make date conversion mandatory at this time. Currently, CONTENTdm does not allow users to perform date range searches or to search for specific dates within a date range when using the W3C-DTF format. Although CONTENTdm will expand dates if the *Date* field is formatted as DATE, the formatting CONTENTdm supports is not compatible with the W3C-DTF (i.e., CONTENTdm requires items using the DATE format to be entered in YYYY, MM-YYYY or DD-MM-YYYY format). When CONTENTdm implements W3C-DTF support, the Metadata Committee will reconsider mandatory conversion of records to conform to W3C-DTF dates.

For compound objects:

Use only for the document level metadata. Do not add *Date Original* to the page level metadata. The *Date Original* field should be left blank.

Element/Field **DATE DIGITAL**

Mandatory

Repeatable

DC Definition: A date of an event in the lifecycle of the resource (Equivalent to DC: Date)

DC Comment: typically, Date [Digital] will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3C-DTF] and follows the YYYY-MM-DD format.

WS Comment: Date of creation or availability of the *digital* resource. The contributing institution may approximate the date a resource was digitized.

WWDL Instructions:

Use ***Date Digital*** for born digital as well as digitized resources.

This field maps to the DC DATE in local CONTENTdm collections.

For all new metadata records created for the WWDL, contents of the ***Date Original*** and ***Date Digital*** fields must conform to the W3C-DTF (Date and Time Formats) at <http://www.w3.org/TR/NOTE-datetime.html>. This is the YYYY-MM-DD, or YYYY-MM, or YYYY format.

For all existing metadata records, the Metadata Committee is recommending that institutions correct any ***Date Original*** or ***Date Digital*** fields that do not comply with the W3C-DTF. This includes items with circa dates and dates using other formatting. It appears that the most common error appearing in the ***Date*** fields is ca. for circa dates. Circa dates can still be expressed within the metadata; however, it should be done utilizing the DCMI Period Syntax (1998?-2002?). For single circa dates, the following syntax should be utilized: 1998?. For more information on how the ***Date*** fields should be formatted, refer to the *Western States Dublin Core Metadata Best Practices, Version 2.0*. However, since searchability is not currently affected by the presence of non-W3C-DTF formatted data, the Metadata Committee will not make date conversion mandatory at this time. Currently, CONTENTdm does not allow users to perform date range searches or to search for specific dates within a date range when using the W3C-DTF format. Although CONTENTdm will expand dates if the ***Date*** field is formatted as DATE, the formatting CONTENTdm supports is not compatible with the W3C-DTF (i.e., CONTENTdm requires items using the DATE format to be entered in YYYY, MM-YYYY or DD-MM-YYYY format). When CONTENTdm implements W3T-DFT support, the Metadata Committee will reconsider mandatory conversion of records to conform to W3C-DTF dates.

For compound objects:

Use only for the document level metadata. Do not add ***Date Digital*** to the page level metadata. The ***Date Digital*** field should be left blank.

Revised 10/18/2007

Element/Field **TYPE**

Mandatory

Repeatable

DC Definition: The nature or genre of the content of the resource.

DC Comment: Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type vocabulary [DCMITYPE]). To describe the physical or digital manifestation of the resource, use the *Format* element.

WS Comment: Best practice for Western States metadata is to use only the DCMI Type vocabulary.

WWDL Instructions:

Apply the *WSDCMBP* instructions for adding *Type* to the metadata.

Use the recommended DCMI Type Vocabulary at

<http://www.dublincore.org/documents/dcmi-type-vocabulary>.

The Metadata Committee recognizes that local needs may call for other terms to be used in the *Type* field. Other terms can be included, but the DCMI Type term **must also be present**.

Following is a list of common digital objects and the appropriate DCMI Type:

- Photographs = Image
- Video = MovingImage (May add Video to the *Type* field if local practice requires it, or include the term video and the specific format in *Format* field e.g. video/mpeg)
- Audio = Sound (May add Audio to the Type field if local practice requires it, or include the term audio and the specific format in Format field, e.g. audio/mp3)
- Text (including images of text, such as handwritten documents in JPEG) = Text (If the digital object is an image of text, e.g. JPEG image of a handwritten document, use Image/jpeg in the *Format* field to reflect this and for local file management)
- Cartographic Materials = Image; Map [Use of both terms is preferred.] There is no term for cartographic materials in the DCMI Type Vocabulary; however, the Metadata Committee knows that maps and other cartographic resources will constitute an important part of the WWDL, and it is expected that users will want to limit searches to those kinds of materials. Just as Text is the term to use for digital representations of what we would consider text when looking at the source, so should digital versions of what we would recognize to be cartographic materials be identified as such. The Metadata Committee recommends using the term Map within the *Type* element, to help distinguish such resources for retrieval. Image is the DCMI term, and Map will be the project term. At the present there are no GIS datasets included on the site, but the Metadata Committee feels it is likely that those kinds of resources will be included at some point in the future. Note that for these resources, the DCMI type is Dataset. Recommendations for the project term for GIS datasets have not yet been determined.
- Charts/diagrams = Image

The following table may better illustrate the requirements for the *Type* element:

Common/generic/popular term	DCMI-Type (mandatory terminology)	Notes
Photograph	Image	May use narrower term 'StillImage' as well as 'Image'. May use additional term 'photograph' if required by local practice.
Video	MovingImage	May use additional term 'video' if required by local practice. Do not include the more inclusive term 'Image'
Audio	Sound	May use additional term 'audio' if required by local practice.
Text	Text	Note: facsimiles or images of texts are still of the genre text, including jpg and pdf. For such objects do not include the term Image.
Cartographic materials	Image ; Map	Use of both terms is preferred. May use narrower term 'StillImage' as well as 'Image'.
Charts and/or diagrams	Image	May use narrower term 'StillImage' as well as 'Image'.

Note that retrieval of digital objects by end-users is **dependant on the appropriate terminology included in the metadata**. Per our grant parameters, the WWDL metadata is based on the *Western States Dublin Core Metadata Best Practices*, which recommends using a controlled vocabulary such as the DCMI Type Vocabulary. The above table indicates instances when it is preferable to go beyond the DCMI types when they are inadequate to represent the content of our collections, but in order to have digital objects retrieved properly, the DCMI type terminology **must be included** in the *Type* field.

For compound objects:

Use the appropriate DCMI Type term in the document level metadata only. Do not add *Type* to the metadata for the page level metadata. The *Type* field should be left blank.

Element/Field **FORMAT**

Mandatory

Repeatable

DC Definition: The physical or digital manifestation of the resource.

DC Comment: Typically, *Format* may include the media-type or dimensions of the resource.

Format may be used to determine the software, hardware or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

WS Comment: Use the *Format* element to record the Internet Media Type (IMT scheme). Use the Extent refinement to record a resource's file size and/or duration. Use the Medium refinement to describe an item's physical (as opposed to its digital) nature. The *Format* element is reserved for describing the *access* file only (be it image, audio, or video). Technical metadata relating to the digitization process (i.e., scanner model, scanner resolution, color schemes, file size of the *master* file, etc.) should be recorded in the *Digitization Specifications* element.

WWDL Instructions:

Apply the *WSDCMBP* instructions for adding *Format* to the metadata.

Terminology for *Format* should come from the Internet Media Type (MIME type). Size may be included in either a separate field (*Format Extent*) mapped to *Format*, or in the same field as the MIME type, separated per *WSDCMBP* guidelines. The list of MIME types is available at:

<http://www.iana.org/assignments/media-types> or

<http://www.graphcomp.com/info/specs/mime.html>

For compound objects:

Use the appropriate MIME type term in the document level metadata only. Do not add *Format* to the page level metadata for the chapter/page level metadata. The *Format* field should be left blank.

Revised 10/18/2007

Element/Field **IDENTIFIER (Resource Identifier in *WSDCMBP*)**

Mandatory

Repeatable

DC Definition: An unambiguous reference to the resource within a given context

DC Comment: Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).

WS Comment: A character string or record number that clearly and uniquely identifies a digital object or resource. The *Identifier* element ensures that individual digital objects can be accessed, manage, stored, recalled and used reliably. Input ISSN, ISBN, other international standard numbers, [and] local naming conventions that describe the original in *Source*.

WWDL Instructions:

Apply the *WSDCMBP* instructions for adding an identifier to the *Identifier* field.

Participants may use their OCLC code as the beginning portion of the file name, but it is not required. The *Identifier* element is more critical for local databases than the MultiSiteServer.

For compound objects:

Use the full filename, including extension, of the object. This is imported automatically into the *Identifier* field by selecting "filename" from the drop-down box for that field in the Template Creator when importing compound objects into CONTENTdm.

Element/Field **PLACE NAME**

Recommended, if applicable

Repeatable

DC Definition: none.

DC Comment: none.

WS Comment: The geographic area of coverage of a resource.

WWDL Instructions:

This field maps to the DC *Coverage* field.

Give the place name in SACO format, which is used for Library of Congress Subject Headings. For resolution of differing forms of a place name, use a standard gazetteer, such as the *Columbia-Lippincott Gazetteer of the World* or Webster's *Geographical Dictionary*.

For compound objects:

Place name should be added only to the document level metadata. Do not add **Place name** to any page level metadata. The **Place name** element should be left blank.

Element/Field **GEOGRAPHIC COORDINATES**

Recommended, if applicable

Repeatable

DC Definition: none.

DC Comment: none.

WS Comment: The geographic coordinates of the geographic area of coverage of a resource.

WWDL Instructions:

This field maps to the DC *Coverage* field.

When available, enter the geographic coordinates that define the location of the object. Examples are the latitude and longitude or UTM coordinates that define the boundaries of a map, or the UTM coordinates of a resource that can be used to link the resource to a map.

For compound objects:

Geographic Coordinates should be added only to the document level metadata.

Do not add ***Geographic Coordinates*** to any page level metadata. The ***Geographic Coordinates*** element should be left blank.

Compound Objects

Apply only four fields to the metadata for individual pages within compound objects. Those four fields are:

Title: This is usually the filename, truncated to leave only the last bit, such as "page001". Imported, truncated, and inserted in the **Title** field automatically through CONTENTdm's compound object creator.

Description: OCR'd or transcribed text for that page, which must have the same filename as its corresponding image, but be grouped in a separate directory. Imported and placed in the **Description** field automatically through CONTENTdm's compound object creator.

Identifier: The full filename, including extension, of the page image. Imported into the **Identifier** field automatically by selecting "filename" from the drop-down box for that field in the Template Creator.

Source: Title or truncated title of book, manuscript number, or other unique character string to identify the document to which the page belongs. Imported automatically by filling in value for that field in the Template Creator.

The compound object itself (document level) gets a full metadata record. The following fields should NOT be populated for individual pages

Creator
Subject
Date
Type

If these fields are populated, then individual pages will be retrieved in a search of the aggregated index, which would be confusing for the end-user.