Geospatial Multistate Archive and Preservation Partnership (GeoMAPP)

Archival Metadata Elements for the Preservation of Geospatial Datasets

September 21, 2011

version: 1.0 (final)

Table of Contents

Geospatial Multistate Archive and Preservation Partnership (GeoMAPP)
Archival Metadata Elements for the Preservation of Geospatial Datasets

1. Paper Goal

Offer a suggested archival metadata element set to facilitate and promote the long-term preservation of geospatial datasets that is consistent and compatible with archival metadata identified to preserve digital assets.

2. Introduction

To promote the preservation, discovery and use of geospatial data, GIS producers will create metadata to accompany their geospatial datasets that includes:

- **Who**: creator of data
- **What:** title and description of data
- **Where:** geographical extent of data
- **Why:** reason data was created
- **When:** when the data was created
- **How:** how the data was created

The Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM)[1] defines a standard metadata model that facilitates and promotes consistency in geospatial metadata creation, which benefits the description of datasets for long term preservation.

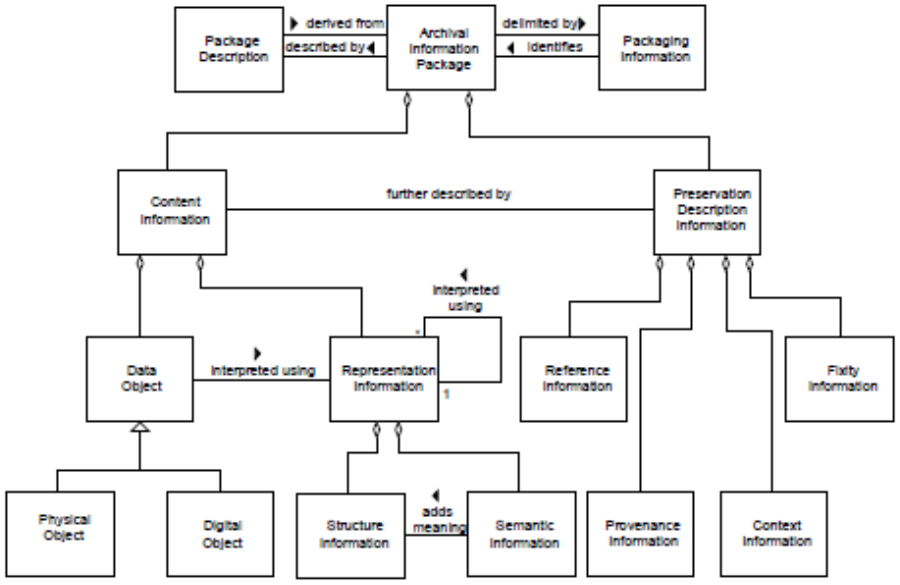
Supplying the FGDC Metadata as part of the geospatial dataset creation also benefits the long-term preservation of the dataset, as it includes key descriptive and source information that will be useful to future users of the dataset, especially when considering the access may occur long into the future. As datasets are transferred to an archives, the archives staff may then supplement the original object's metadata with additional metadata that will facilitate the management and long-term preservation of the dataset. These metadata elements may include

- **Technical Details:** hardware, software environment, or application dependencies
- **Administrative Details:** donor information, dataset ownership information, use and access rights or restrictions
- **Preservation Details:** information about the processing the archives performs to prepare the dataset for long-term preservation, and ongoing activities to ensure long-term access

The Open Archival Information System (OAIS) framework[2] offers a metadata framework to organize metadata intended to support the long-term preservation and archiving of digital objects. In addition to capturing the basic descriptive metadata, it also includes important technical metadata to facilitate long-

[1] Federal Geographic Data Committee Geospatial Metadata Standards: http://www.fgdc.gov/metadata/geospatial-metadata-standards
[2] Consultative Committee for Space Data Systems. "Reference Model for an Open Archival Information System (OAIS); CCSDS 650.0-B-1 Blue Book." (January 2002). http://public.ccsds.org/publications/archive/650x0b1.PDF

term access, and administrative metadata that records the ongoing curation of the digital object through the remainder of its life within the archives.[3]



The OAIS model defines a taxonomy of archival information object classes that includes:

- **Content Information** – identifies the object that is being preserved. May consist of a file identifier, file size. Note: Other detailed information about the Digital Object can be found in the Representation Information and Preservation Description Information sections.
- **Representation Information** – information that describes the representation formats of the digital object or sequence of bits. This provides necessary information needed to render the digital object. May include links to format specifications.
    - **Structure Information** – imparts meaning about how the information is organized.
    - **Semantic Information** – provides additional meaning on the special meanings associated with all of the elements of the Structural Information, operations that may be performed on each data type, and their inter-relationships.
- **Preservation Description Information** – allows the understanding of the Content Information over an indefinite period of time.
    - **Reference** – mechanism(s) that identifies and describes mechanisms used to assign identifiers to the Content Information.
    - **Context** – documents the relationships between the Content Information and its environment.
    - **Provenance** – documents the history of the Content Information. Provides the source or origin, any changes, and who has had custody of it since it was originated.
    - **Fixity** – provides the Data Integrity checks or Validation/Verification keys used to ensure that the particular Content Information object has not been altered in an undocumented manner. (authentication mechanisms within the repository)
- **Packaging Information** – information that describes how the Content Information is packaged within a particular environment (e.g. METS).
- **Descriptive Information** – information that helps users locate and access the Content Information object.

[3] ibid. p. 4-37.

While the OAIS model provides general categories of information that should be represented in the archival metadata, it does not define a specific metadata model. The table that follows provides a suggested list of archival metadata elements for the preservation geospatial datasets. The table was largely informed by three sources:

1. **CEDARS project preservation metadata elements**[4]: the CEDARS project aimed "to address the strategic, methodological and practical issues" to "provide guidance for libraries in the best practice of digital preservation."[5] The project published an extensive preservation metadata model to support the documentation of the preservation of digital objects. The model was designed to be flexible enough to accommodate several digital object types, and was therefore allowed the reasonable addition of geospatial-specific metadata. The table that follows is largely an extension of the CEDARS metadata table, augmented to include geospatial-related preservation metadata.

2. **National Geospatial Digital Archive report: *An Investigation into Metadata for Long-lived Geospatial Data Formats* by Nancy Hoebelheinrich**: This report produced from the NGDA project provides a recommendation on several categories of metadata elements "necessary for archiving complex geospatial data as well as what if any, auxiliary data sources are needed for correctly understanding the data."[6] The table that follows includes the geospatial metadata recommended in the Hoebelheinrich paper. This paper is also useful in that it includes crosswalks to FGDC, the Center for International Earth Science Information Network (CIESIN) Geospatial Electronic Records (GER) geospatial metadata model, and to PREMIS.

3. **CIESIN GER Data Model:** [7] This offers an even more extensive data model originating with the purpose of managing geospatial electronic records.

The benefit of combining the CEDARS preservation metadata with the NGDA metadata provides a metadata model based on the OAIS framework that will support the preservation of digital content, in general, and geospatial digital resources, in particular. This approach offers archives with a preservation data model that can be utilized for not just their geospatial datasets, but can also support the preservation metadata needs for other digital artifacts such as images, documents, audio, etc.

3. Archival Metadata Elements for the Preservation of Geospatial Datasets – Data Model Diagrams

The diagrams and table that follows presents an OAIS-based metadata model to support the preservation of digital assets including geospatial digital resources. The table is organized based on the OAIS model. It includes the metadata element and a brief description. The table also provides columns to denote whether a data element is required (*req*), and whether the data element is repeatable (*rep*). Each archives will need to identify required fields based on its own policies, technical capabilities, and labor resources. The table also includes a column for additional comments, and a column to record the source

[4] Stone, Andy, and Day, Michael. "Cedars Preservation Metadata Elements, Cedars Project Document AIW02." (1999). http://www.ukoln.ac.uk/metadata/cedars/papers/aiw02/

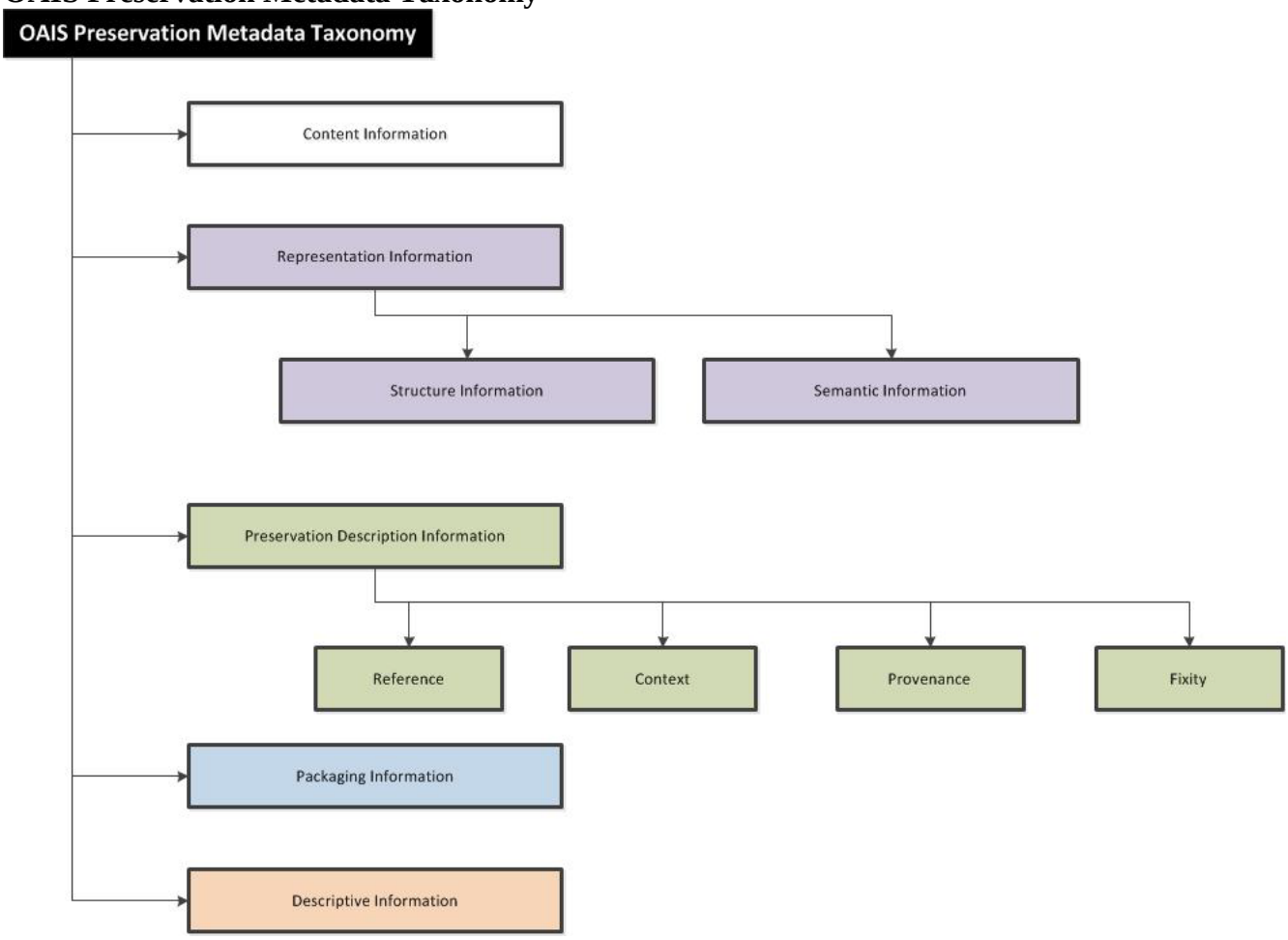
[5] The CEDARS Project CURL Exemplars in Digital ARchiveS. http://www.ukoln.ac.uk/services/elib/projects/cedars/

[6] Hoebelheinrich, Nancy. "An Investigation into Metadata for Long-lived Geospatial Data Formats." (2008). p. 5. http://www.ngda.org/reports/InvestigateGeoDataFinal_v2.pdf

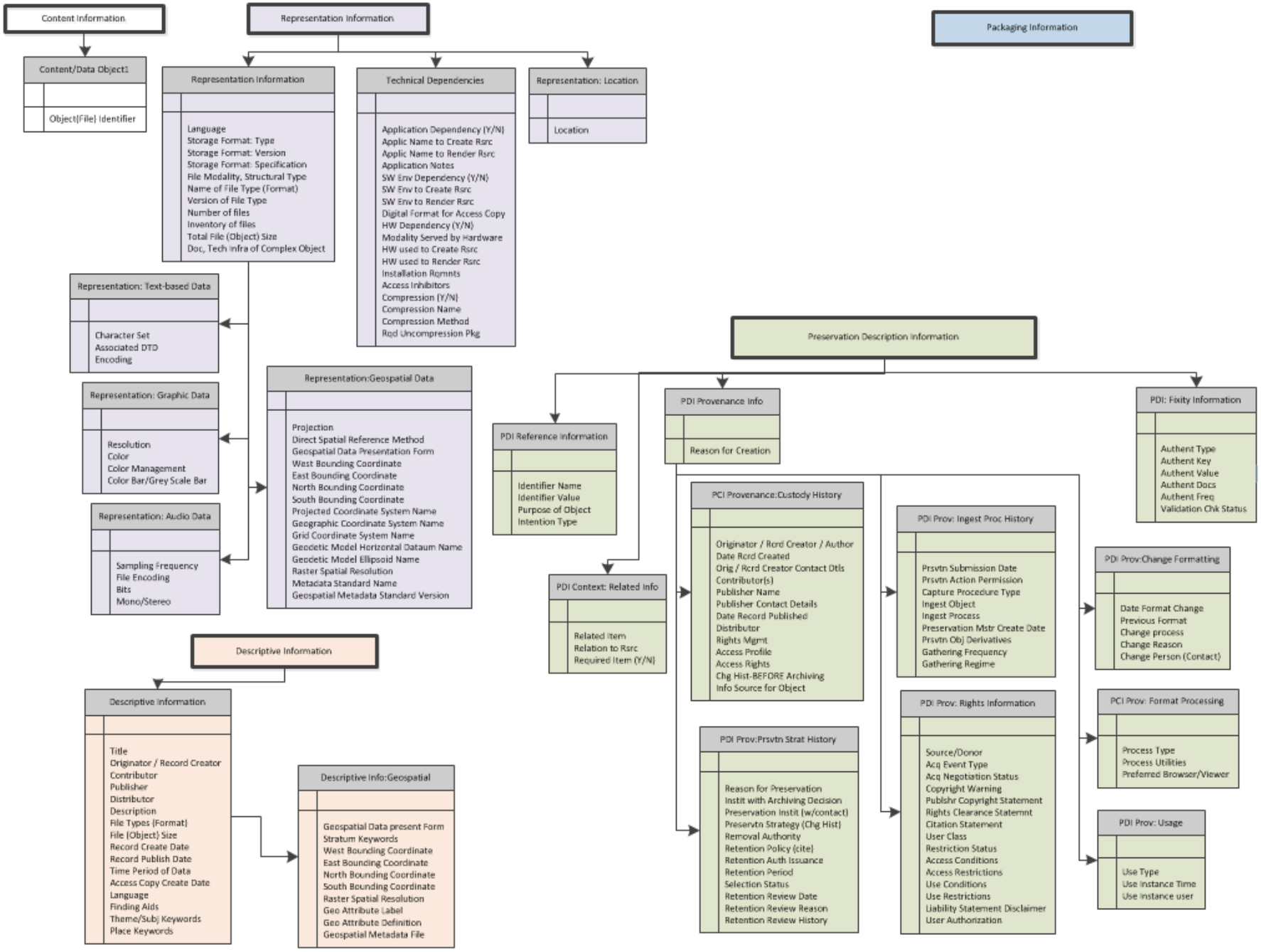
[7] Center for International Earth Science Information Network (CIESIN), Columbia University. "Data Model for Managing and Preserving Geospatial Electronic Records, Version 1.00." (June 2005). http://www.ciesin.org/ger/DataModelV1_20050620.pdf

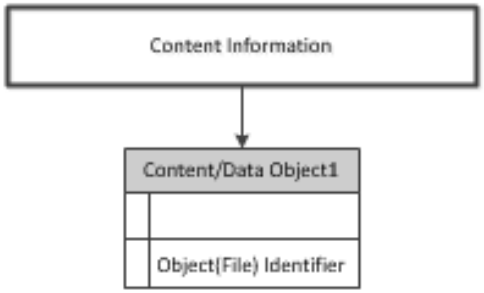
for the metadata value. Finally, the table provides crosswalk references to FGDC[8] and Dublin Core[9] metadata elements, and an additional column for an implementing archive to capture the corresponding metadata fields in the archives' catalog or online digital collection. **Appendix A** offers a summary view of the table that archives may want to use as they are assessing their archival metadata needs. Several FGDC metadata elements have been included throughout the metadata table, but is not all-inclusive. An archives may identify, based on its local implementation, additional FGDC metadata fields the it wants to include in its archival metadata record. Refer to **Appendix B** for a summary view of the FGDC metadata that may be used to review and record the FGDC metadata fields to include in your geospatial metadata record. When Esri shapefiles are being preserved, it may not be necessary to store all of the metadata in the archival metadata record, as it is usually available in a standalone xml-based text geospatial metadata file. This also offers the preservation benefit of a human-readable form that is de-coupled from the dataset. However, for Esri geodatabases, the geospatial metadata for each feature class is embedded in the geodatabase, so archival organizations may want to consider their strategies for application-independent access to those datasets' geospatial metadata.

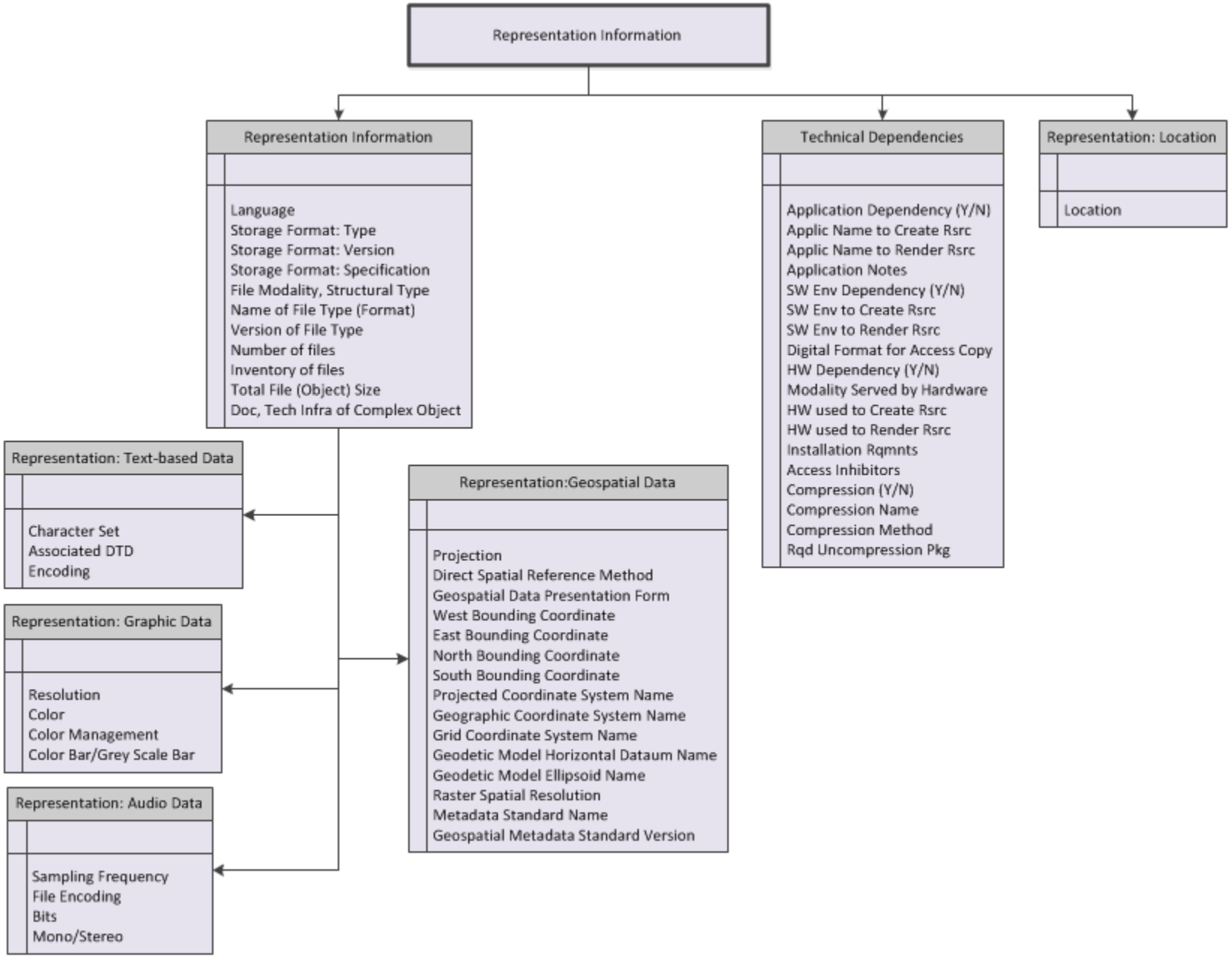
OAIS Preservation Metadata Taxonomy

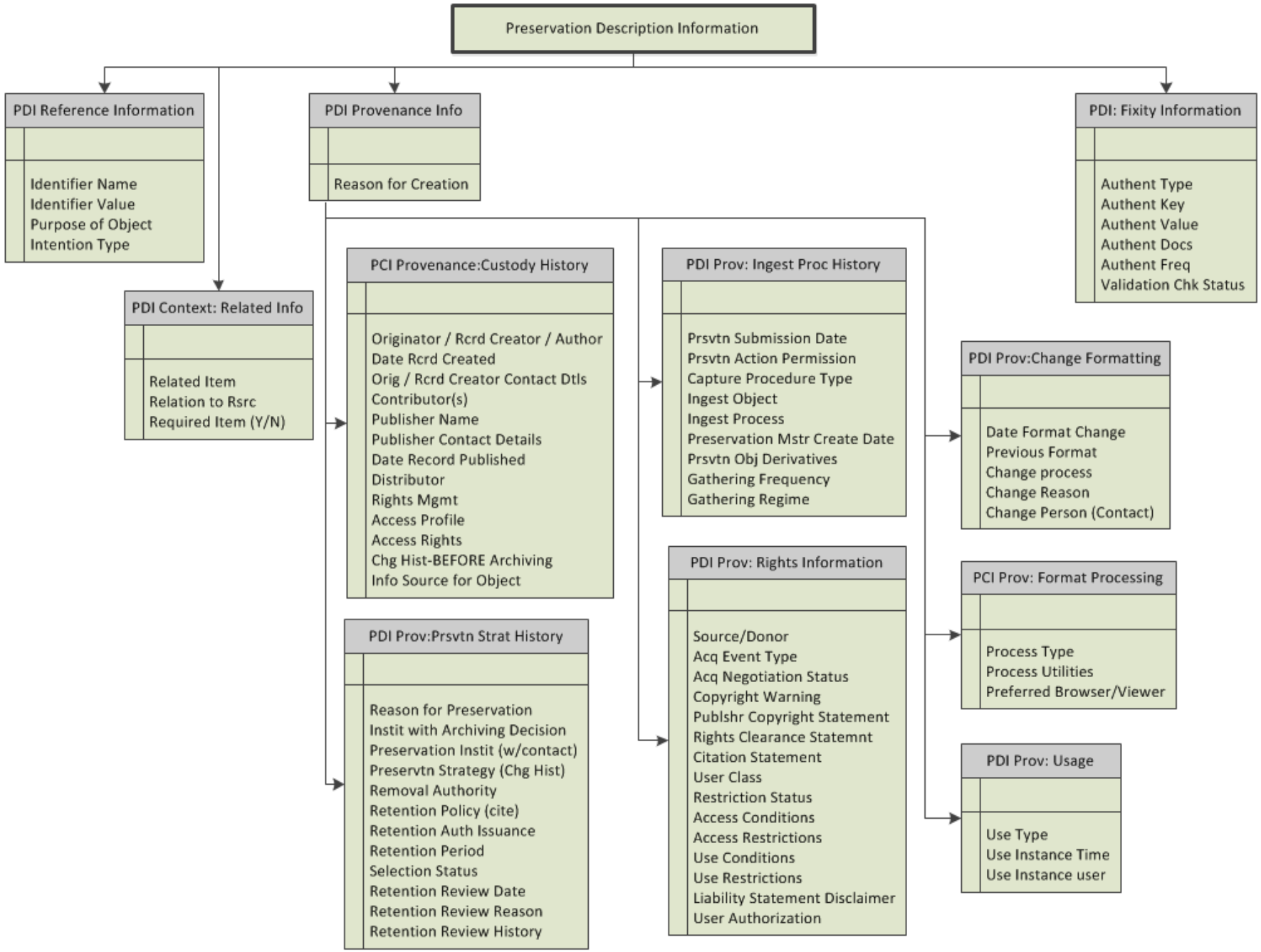


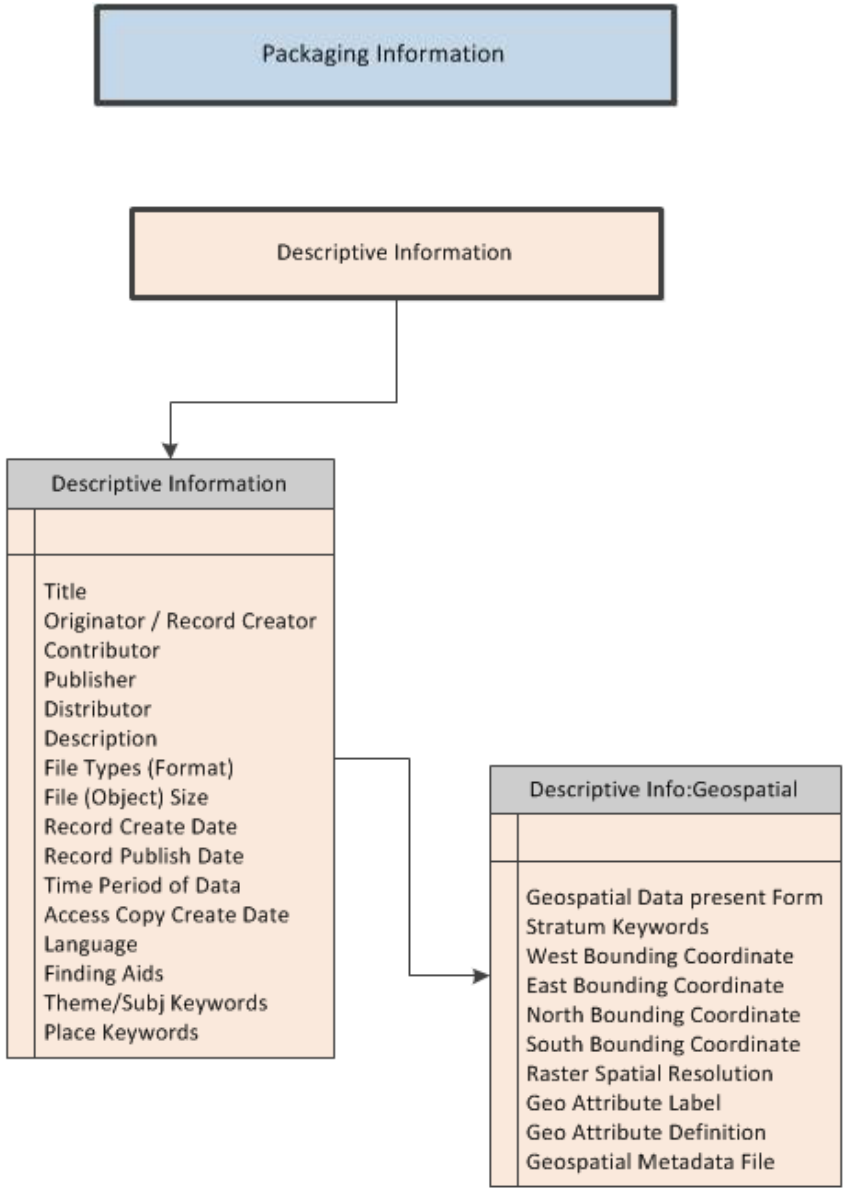
[8] Federal Geographic Data Committee. "Content Standard for Digital Geospatial Metadata FGDC-STD-001-1998." 1998. http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/v2_0698.pdf
[9] Dublin Core Metadata Initiative. "DCMI Specifications." http://www.dublincore.org/specifications/


Content Information
Content/Data Object1
Object(File) Identifier
Representation Information
Representation Information
Language
Storage Format: Type
Storage Format: Version
Storage Format: Specification
File Modality, Structural Type
Name of File Type (Format)
Version of File Type
Number of files
Inventory of files
Total File (Object) Size
Doc, Tech Infra of Complex Object
Technical Dependencies
Application Dependency (Y/N)
Applic Name to Create Rsrc
Applic Name to Render Rsrc
Application Notes
SW Env Dependency (Y/N)
SW Env to Create Rsrc
SW Env to Render Rsrc
Digital Format for Access Copy
HW Dependency (Y/N)
Modality Served by Hardware
HW used to Create Rsrc
HW used to Render Rsrc
Installation Rqmnts
Access Inhibitors
Compression (Y/N)
Compression Name
Compression Method
Rqd Uncompression Pkg
Representation: Location
Location
Packaging Information
Representation: Text-based Data
Character Set
Associated DTD
Encoding
Representation: Graphic Data
Resolution
Color
Color Management
Color Bar/Grey Scale Bar
Representation: Audio Data
Sampling Frequency
File Encoding
Bits
Mono/Stereo
Representation:Geospatial Data
Projection
Direct Spatial Reference Method
Geospatial Data Presentation Form
West Bounding Coordinate
East Bounding Coordinate
North Bounding Coordinate
South Bounding Coordinate
Projected Coordinate System Name
Geographic Coordinate System Name
Grid Coordinate System Name
Geodetic Model Horizontal Datum Name
Geodetic Model Ellipsoid Name
Raster Spatial Resolution
Metadata Standard Name
Geospatial Metadata Standard Version
Preservation Description Information
PDI Reference Information
Identifier Name
Identifier Value
Purpose of Object
Intention Type
PDI Provenance Info
Reason for Creation
PDI: Fixity Information
Authent Type
Authent Key
Authent Value
Authent Docs
Authent Freq
Validation Chk Status
PDI Context: Related Info
Related Item
Relation to Rsrc
Required Item (Y/N)
PCI Provenance:Custody History
Originator / Rcrd Creator / Author
Date Rcrd Created
Orig / Rcrd Creator Contact Dtls
Contributor(s)
Publisher Name
Publisher Contact Details
Date Record Published
Distributor
Rights Mgmt
Access Profile
Access Rights
Chg Hist-BEFORE Archiving
Info Source for Object
PDI Prov: Ingest Proc History
Prsvtn Submission Date
Prsvtn Action Permission
Capture Procedure Type
Ingest Object
Ingest Process
Preservation Mstr Create Date
Prsvtn Obj Derivatives
Gathering Frequency
Gathering Regime
PDI Prov:Change Formatting
Date Format Change
Previous Format
Change process
Change Reason
Change Person (Contact)
PDI Prov:Prsvtn Strat History
Reason for Preservation
Instit with Archiving Decision
Preservation Instit (w/contact)
Preservtn Strategy (Chg Hist)
Removal Authority
Retention Policy (cite)
Retention Auth Issuance
Retention Period
Selection Status
Retention Review Date
Retention Review Reason
Retention Review History
PDI Prov: Rights Information
Source/Donor
Acq Event Type
Acq Negotiation Status
Copyright Warning
Publshr Copyright Statement
Rights Clearance Statemnt
Citation Statement
User Class
Restriction Status
Access Conditions
Access Restrictions
Use Conditions
Use Restrictions
Liability Statement Disclaimer
User Authorization
PCI Prov: Format Processing
Process Type
Process Utilities
Preferred Browser/Viewer
PDI Prov: Usage
Use Type
Use Instance Time
Use Instance user
Descriptive Information
Descriptive Information
Title
Originator / Record Creator
Contributor
Publisher
Distributor
Description
File Types (Format)
File (Object) Size
Record Create Date
Record Publish Date
Time Period of Data
Access Copy Create Date
Language
Finding Aids
Theme/Subj Keywords
Place Keywords
Descriptive Info:Geospatial
Geospatial Data present Form
Stratum Keywords
West Bounding Coordinate
East Bounding Coordinate
North Bounding Coordinate
South Bounding Coordinate
Raster Spatial Resolution
Geo Attribute Label
Geo Attribute Definition
Geospatial Metadata File


Content Information
Content/Data Object1
Object(File) Identifier


Representation Information
Representation Information
Language
Storage Format: Type
Storage Format: Version
Storage Format: Specification
File Modality, Structural Type
Name of File Type (Format)
Version of File Type
Number of files
Inventory of files
Total File (Object) Size
Doc, Tech Infra of Complex Object
Technical Dependencies
Application Dependency (Y/N)
Applic Name to Create Rsrc
Applic Name to Render Rsrc
Application Notes
SW Env Dependency (Y/N)
SW Env to Create Rsrc
SW Env to Render Rsrc
Digital Format for Access Copy
HW Dependency (Y/N)
Modality Served by Hardware
HW used to Create Rsrc
HW used to Render Rsrc
Installation Rqmnts
Access Inhibitors
Compression (Y/N)
Compression Name
Compression Method
Rqd Uncompression Pkg
Representation: Location
Location
Representation: Text-based Data
Character Set
Associated DTD
Encoding
Representation: Graphic Data
Resolution
Color
Color Management
Color Bar/Grey Scale Bar
Representation: Audio Data
Sampling Frequency
File Encoding
Bits
Mono/Stereo
Representation:Geospatial Data
Projection
Direct Spatial Reference Method
Geospatial Data Presentation Form
West Bounding Coordinate
East Bounding Coordinate
North Bounding Coordinate
South Bounding Coordinate
Projected Coordinate System Name
Geographic Coordinate System Name
Grid Coordinate System Name
Geodetic Model Horizontal Dataum Name
Geodetic Model Ellipsoid Name
Raster Spatial Resolution
Metadata Standard Name
Geospatial Metadata Standard Version


Preservation Description Information
PDI Reference Information
Identifier Name
Identifier Value
Purpose of Object
Intention Type
PDI Provenance Info
Reason for Creation
PDI: Fixity Information
Authent Type
Authent Key
Authent Value
Authent Docs
Authent Freq
Validation Chk Status
PDI Context: Related Info
Related Item
Relation to Rsrc
Required Item (Y/N)
PCI Provenance:Custody History
Originator / Rcrd Creator / Author
Date Rcrd Created
Orig / Rcrd Creator Contact Dtls
Contributor(s)
Publisher Name
Publisher Contact Details
Date Record Published
Distributor
Rights Mgmt
Access Profile
Access Rights
Chg Hist-BEFORE Archiving
Info Source for Object
PDI Prov: Ingest Proc History
Prsvtn Submission Date
Prsvtn Action Permission
Capture Procedure Type
Ingest Object
Ingest Process
Preservation Mstr Create Date
Prsvtn Obj Derivatives
Gathering Frequency
Gathering Regime
PDI Prov:Change Formatting
Date Format Change
Previous Format
Change process
Change Reason
Change Person (Contact)
PDI Prov:Prsvtn Strat History
Reason for Preservation
Instit with Archiving Decision
Preservation Instit (w/contact)
Preservtn Strategy (Chg Hist)
Removal Authority
Retention Policy (cite)
Retention Auth Issuance
Retention Period
Selection Status
Retention Review Date
Retention Review Reason
Retention Review History
PDI Prov: Rights Information
Source/Donor
Acq Event Type
Acq Negotiation Status
Copyright Warning
Publshr Copyright Statement
Rights Clearance Statemnt
Citation Statement
User Class
Restriction Status
Access Conditions
Access Restrictions
Use Conditions
Use Restrictions
Liability Statement Disclaimer
User Authorization
PCI Prov: Format Processing
Process Type
Process Utilities
Preferred Browser/Viewer
PDI Prov: Usage
Use Type
Use Instance Time
Use Instance user


Packaging Information
Descriptive Information
Descriptive Information
Title
Originator / Record Creator
Contributor
Publisher
Distributor
Description
File Types (Format)
File (Object) Size
Record Create Date
Record Publish Date
Time Period of Data
Access Copy Create Date
Language
Finding Aids
Theme/Subj Keywords
Place Keywords
Descriptive Info:Geospatial
Geospatial Data present Form
Stratum Keywords
West Bounding Coordinate
East Bounding Coordinate
North Bounding Coordinate
South Bounding Coordinate
Raster Spatial Resolution
Geo Attribute Label
Geo Attribute Definition
Geospatial Metadata File

4. Archival Metadata for the Preservation of Geospatial Datasets Table

Content Information								
Content/Data Object								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Object (File) Identifier	The unique (persistent) identifier which indicates each data object being archived.	This may be a unique internal identifier generated by the archival management system. Alternatively, for file-based assets, it may be characterized as an externally coherent fully qualified path to a file; e.g. HOST:Directory/File Path designation, or use an identification mechanism such as a PURL (Persistent Uniform Resource Locator) or CRID (Content Reference Identifier).	Yes	No	Archival Management System	DC.Identifier		Catalog Id
Representation Information								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Language	The language of the content of the data object being archived. NISO Z39.53 codes could be used.	To be used where language is a significant part of the resource. Important for bibliographic purposes. Assumption: most GIS materials will be in English		Yes		DC.Language		Language
Storage format: Type	The initial storage format of the data object being archived. Examples might be: Diskette 3.5 inch Diskette 5.25 inch CD-ROM Information server HFS path	This may be important since the underlying software may require the emulation of mounting diskettes, for example. This may not be relevant for materials that are being transferred from a portable media to a preservation repository, and the media is just used as a transfer mechanism rather than as a long-term preservation storage format.		Yes	To be supplied as part of the data ingest input.			Format

Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Storage format: Version	The version of storage format being used.	Used to represent the storage format of the physical media the file/object is submitted on or intended as media for long-term storage. Will be a more critical characteristic if the preserved object is to be preserved on its original media.			To be supplied as part of the data ingest input.			
Storage format: Specification	A link to documentation about the source format for the storage media (e.g. a CD specification).	The link would need to be to a reference list of formats that could be derived from the 'type' and 'version' fields		Yes	To be supplied as part of the data ingest input.	DC.Conforms To		
File Modality, Structural Type	This is a higher level of description, which is similar to the Dublin Core 'Type' element, i.e. **text, graphic, sound, video, GIS** etc. Adapted from the Pittsburgh metadata set. The type of object or collection being described using one of the categories (see above): Note: even though an item may consist of multiple component parts, if it's associated with 1 logical item, it would be considered an "object" rather than a collection. A GIS Dataset is a good example of an "object", where a geodatabase might be considered a Collection.	If a type of encoding is not initially recognized, one could be made aware of its generic type so that exploratory work can be done to assess the feasibility of accessing the file. This will be the basis for identifying which subset of format-specific representation fields will be used. For GIS objects, use GIS as a high level designator. The File Type below will be used to further classify the type of GIS object.			To be supplied as part of the data ingest input. or auto-sensed by archival management system based on characteristics of object being ingested	DC.Type - text - image - sound - data - software - interactive - physical object GIS dataset = interactive?	ESRI: <natvform> = Shapefile	

Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Name of File Type (Object Type) (sometimes referred to as "Format")	The name of the file type, or object type. e.g. TIFF HTML Proprietary Collection For GIS objects consider the following: ShapeFile OrthoPhoto Digitized Map GIS Project File Geodatabase Personal File GeoDB ArcSDE DB	A reference list of file formats would need to be maintained; formatting metadata may be required to understand how logical access to the bit-stream takes place. Collection would indicate this particular package contains a collection of files For well-known single file entities being archived, the Mime Type is a common value for this field. However, there are not yet mime types defined for most of the geospatial data formats, so GIS-specific custom-field values may need to be defined to represent the type of the object. Note: as geospatial datasets may contain image files, a hierarchical Representation may need to be constructed that: level 1: identifies the GIS dataset level 2: characterizes the representation of the individual files within the GIS dataset - such as the image files bit depth, resolution, etc. IANA MIME Media Types: http://www.iana.org/assignments/media-types/index.html Dublin Core Type: http://dublincore.org/documents/dcmi-type-vocabulary/		Yes	To be supplied as part of the data ingest input. or auto-sensed by archival management system based on characteristics of object being ingested	DC.Format *Mime Type*	<formname> 6.4.2.1.1	File Characteristics

Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Version of File Type	The version of file type e.g. TIFF Revision 5.0 TIFF Revision 6.0	The reference list of file formats would include the differences between different versions of file types		Yes	To be supplied as part of the data ingest input.		<native> 1.13	
Number of files	The total number of files that comprise the archive object (preservation master)	The preservation master or the archival information package (AIP)		Yes	To be supplied as part of the data ingest input, or introspect pkg.			
Inventory of files For multi-file "objects" (such as GIS datasets) a file inventory of the object	Provide a detailed file inventory or manifest of the preservation master of the archived object. e.g: 1 shapefile 12,983 kb consisting of: county_bndry_2009.dbf 1164kb county_bndry_2009.prj 1 kb county_bndry_2009.sbn 46 kb county_bndry_2009.sbx4 kb county_bndry_2009.shp 11,698 kb county_bndry_2009.shp.xml 34 kb county_bndry_2009.shx 36 kb	For a multi-file "object', this is critical to the preservation of the object as a whole. Provide name, file type description (may be inferred by file extension?), and size of each file that comprises the archived object. If using a tool like BagIt, to virtually package a dataset, you may consider using the BagIt manifest as the basis for the Inventory. Note: the file inventory for the preservation master may be different than the dissemination information package / access copy.			BagIt Manifest			
Total file (object) size	The size of the file (object size - for multi-file dataset packages) in kilobytes, megabytes, etc.	This would give an indication as to how much file space this resource would require to host, and how much data would be transferred from the archive.		Yes	To be supplied as part of the data ingest input, or introspect pkg.	DC.Format.Extent	<transize> 6.4.2.1.7	Size
Documentation, Technical Infrastructure of Complex Object	A link to the documentation **about the file format**; should be derived from type and version) e.g. link to the Shapefile specification	The link would need to be to a reference list of formats that could be derived from the 'type' and 'version' fields		Yes	Mapping Table in archival management system.	DC.Conforms To		

Representation: Text-based Data (content-dependent)								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Character Set	The character set which the data uses. (e.g. ASCII, Unicode)	For text-based data only						
Associated DTD	The Data Type Definition which the text requires for interpretation. The value would be the CRID of DTD	This is only required for text-based data which contain markup tags; i.e. SGML, HTML, XML.						
Encoding		For text-based data only						
Representation: Graphic Data (content-dependent)								
Resolution	The graphics resolution of an image; e.g. 600dpi	For graphics data only						
Color Depth	The color depth of an image; e.g. 1-bit (black and white), 16-bit	This will allow users to be aware of the minimum requirement for viewing graphics.		No	Possibly from metadata within graphics files, or from pre-ingest information supplied			
Color Management	Identify system, if any, that is used to improve consistency of color across capture, display and output of an image. EXAMPLES: Photo CD; OptiCal (color management system); Profile/80 (color sync profile maker); Softproof (Photoshop Plugin)	For graphics data only		No				
Color Bar/Grey Scale Bar	This links to a reference image which would allow colour comparison to take place; e,g, Kodak Q13 or Q14 Colour Separation	An image of the color bar may be required within the archive, in order for comparisons to be made.		No				

	Guide and Grey Scale; Kodak Q60 Colour Input Target	For graphics data only						
Representation: Audio Data (content dependent)								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*		*FGDC*	*Access xwlk*
Sampling Frequency	The sampling frequency at which audio was sampled: e.g. 8kHz, 44.1KhZ	This is only required for audio data		No				
File Encoding	The type of encoding used to create a digital audio file. For example, AIFF, AU, WAV.	This may be the same as file format name. This is only required for audio data.		No				
Bits	The number of bits each audio sample contains; e.g. 8-bit, 16-bit.	This is only required for audio data.		No				
Mono/Stereo		This is only required for audio data		No				
Representation: Geospatial Datasets (content dependent) Note: the following provides some preliminary suggestions on the geospatial metadata that should be recorded as part of the Representation. Most should be available in the FGDC metadata file. You should review the FGDC metadata to identify the specific geospatial representation fields (technical metadata) you want to include in your archival metadata record to facilitate the archival management of your archived geospatial datasets.								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Projection	e.g. Lambert_Conformal_Conic	Recommended initial "validation" task for received shapefiles: open shapefile in ESRI ArcCatalog and validate it has a projection defined. Note: the projection is not stored in the FGDC metadata, but rather is a property of the shapefile.		No	FGDC Metadata		<mapprojn> 4.1.2.1.1	
Direct Spatial Reference Method	The system of objects used to represent space in the data set (per FGDC standard).	Controlled vocabulary is defined by the FGDC standard: - Point - Vector		No	FGDC Metadata		<direct> 3.2	

	- Point - Vector - Raster	- Raster						
Geospatial Data Presentation Form	The mode in which the geospatial data are represented (per the FGDC standard): Use cartographic AACR2 for controlled vocabulary. e.g. “vector digital data”, “raster digital data” Per FGDC standard - controlled vocabulary: atlas, audio, diagram, document, globe, map, model, multimedia presentation, profile, raster digital data, remote-sensing image, section, spreadsheet, tabular digital data, vector digital data, video, view, free text.	This identifies the particular subclass or subcategory of geospatial dataset.		No	FGDC Metadata		<geoform> 8.6	Geospatial Type
West_Bounding_Coordinate	Provides the limits of coverage of a data set.			No	FGDC Metadata		<westbc> 1.5.1.1	West Bounding Coordinate
East_Bounding_Coordinate	Provides the limits of coverage of a data set.			No	FGDC Metadata		<eastbc> 1.5.1.2	East Bounding Coordinate
North_Bounding_Coordinate	Provides the limits of coverage of a data set.			No	FGDC Metadata		<northbc> 1.5.1.3	North Bounding Coordinate
South_Bounding_Coordinate	Provides the limits of coverage of a data set.			No	FGDC Metadata		<southbc> 1.5.1.4	South Bounding Coordinate
Spatial Reference Information - Projected Coordinate System Name - Geographic Coordinate System Name	The description of the reference frame for, and the means to encode, coordinates in the data set. (per the FGDC standard)			Yes	FGDC Metadata		<spref> 4. <horizsys> 4.1 <planar> 4.1.2 <mapprojn> 4.1.2.1	

- Grid Coordinate System Name - Geodetic model Horizontal datum name - Geodetic model Ellipsoid name							<gridsys> 4.1.2.2 <gridsysn> 4.1.2.2.1 <geodetic> 4.1.4 <horizdn> 4.1.4.1 <ellips> 4.1.4.2	
Raster (ortho, elevations) Spatial Resolution	Provides ground resolution of the pixels e.g. 0.25 feet 1 foot 0.08 meters 0.25 meters 1.00 meters	More relevant for raster datasets than vector datasets. Provides ground resolution of the pixels in the raster dataset.		No	FGDC Metadata		<abstract> 1.2.1	Resolution
Geospatial Metadata Standard Name	Metadata standard used to create the geospatial metadata that describes the dataset (per the FGDC standard) e.g. FGDC Content Standards for Digital Geospatial metadata			No	FGDC Metadata		<metstdn> 7.5	
Geospatial Metadata Standard Version	Version of the metadata standard used to create the geospatial metadata e,g, FGDC-STD-001-1998 ISO 19115:2003			No	FGDC Metadata		<metsdv> 7.6	
add your additional geospatial-specific Representation Metadata here								

Technical Dependencies								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Application-Dependency	This is used to indicate whether or not a file, collection, or object is dependent on a specific application. This will be a **Yes/No value**.	Where some resources are given in a stand-alone format, the component data files may be in a proprietary format which requires a specific viewer and is not easily converted.		Yes	To be supplied as part of the data ingest input, or introspect pkg.		<native> 1.13	
Application Name used to **create** the resource - Vendor - Application Name - Applic Version	The name of the application which a file or collection is dependent on; e.g. OpenText, DynaText, MS-Word, ArcGIS 9.3.1. Version of the application which the e.g. version 5.0)	In some collections, an application may not be required to view the files which comprise a collection, but may be required to achieve the functionality of the collection. This would mean while the collection is not dependent on the application, it can be noted that added value can be obtained from it using a particular application. This can inform future preservation decisions.		Yes	To be supplied as part of the data ingest input, or introspect pkg.		<native> 1.13 <techprereq> 6.6	
Application Name used to **render** the resource - Vendor - Application Name - Applic Version	The name of the application which the access copy is dependent on; e.g. PDF, ArcCatalog Version of the application which the e.g. ver 5.0)	In some collections, the access copy may required a different program than what was used to create the file (e.g. Adobe Acrobat to create PDFs, Adobe Reader to read PDFs)		Yes	To be supplied as part of the data ingest input, or introspect pkg.			Software Requirements
Application Notes	Text information which is of importance to the successful use of the application; e.g. "Versions of OpenText later than 5.0 are unable to handle SGML files"	This can also be used to describe the functionality which the application adds to the collection.		No	To be supplied as part of the data ingest input			
Software Environment Dependency	Indicates whether the collection is dependent on a	Dependencies on particular operating systems and		Yes	To be supplied as part of the data			

	particular operating system. **Value: yes/no**	environments.			ingest input, or introspect pkg.			
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Software Environment to **create** resource - Vendor - Name - Version	Name of the operating system which was required to create the digital object (e.g. Windows, Macintosh, UNIX, none) The version of the operating system in question: e.g. Windows 3.1, 95, 98, Mac System 7, System 8.1, AIX, IRIX	In the absence of a particular op sys, an emulator might be used to facilitate use. This could be considered an applic dependency, although this matching could be a distinct function, for ease of maint. A list of operating systems and versions could be kept. This may also help with keeping track of emulator compatibility.		Yes	To be supplied as part of the data ingest input, or introspect pkg.		<native> 1.13 <techprereq> 6.6	
Software Environment used to **render** the resource - Vendor - Name - Version				Yes	To be supplied as part of the data ingest input, or introspect pkg.			Software Requirements
Digital Format for Access Copy	Describes the format the access copy is delivered, could be a single file (e.g. PDF, JPG) or a package of files (e.g. ZIP) e.g. zip file Zipped (7 files) 1,725 KB	Might include the general package format (e.g. zip file), the number of files included in the package (e.g. Zipped (7 files), and the size of the package (e.g. 1,725 KB)		Yes	Archival Management System			Digital Format
Hardware Dependency	Indicates whether the object is dependent on any specific hardware. **Value: yes/no**	Dependency on particular hardware components to create or render the file/object.		Yes	To be supplied as part of the data ingest input, or introspect pkg.			
Modality Served by Hardware	Describes the type of resource the hardware is facilitating; e.g. text, graphics, audio, video, input			Yes	To be supplied as part of the data ingest input, or introspect pkg.			

Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Hardware used to create the file/object - Vendor - Name - Model - Description - Specification -Version	Name of the hardware device required: e.g. Mouse, Joystick, Monitor Type of hardware device; e.g. Pointer Device, Color Monitor Description of why hardware is required. (e.g. mouse needed to interact with software; images served are large, high-quality colour, etc.) **(manufacturer, model, version)**	The examples here are trivial, but could include a specific device This may be audio or visual minimum requirements (e.g. a particular audio player, a graphics board, graphics memory, etc.) The description of the hardware dependency will explain why it has been critical to list it. Some value judgment will need to be made as to what can be regarded as superfluous (e.g. monitor, mouse)		No	To be supplied as part of the data ingest input, or introspect pkg.		<native> 1.13 <techprereq> 6.6	
Hardware used to render the file/object -Vendor - Name - Model - Description - Specification -Version	Name of the hardware device required: e.g. Mouse, Joystick, Monitor Type of hardware device; e.g. Pointer Device, Colour Monitor Description of why hardware is required. (e.g. mouse needed to interact with software; images served are large, high-quality colour, etc.) **(manufacturer, model, version)**			No	To be supplied as part of the data ingest input, or introspect pkg.			Software Requirements
Installation Requirements: Install Process	Any specialized procedures needed to install the file/object			No	To be supplied as part of the data ingest input, or introspect pkg.		<techprereq> 6.6	Software Requirements
Access Inhibitors	Any method used to inhibit			Yes	To be supplied as			

	access which would impact preservation (e.g. encryption, watermarking, etc.)				part of the data ingest input.			
Compression								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*		*FGDC*	*Access xwlk*
Compression	Have files been compressed; value:**yes/no**	With proprietary file formats, it may not be possible to know this. This may reflect the compression of the received file/object, or it may reflect the compression of the archived object (e.g. the access copy may use a compression format to optimize storage utilization)		No	To be supplied as part of the data ingest input, or introspect pkg.			
Compression Name	Name of type of compression used (e.g.; ZIP, TAR, MrSID)			No	To be supplied as part of the data ingest input,			
Compression Method	Description of parameters used, package and process used, etc. e.g. "compressed to 80% quality JPEG files using GraphicConverter on macintosh as a batch process"			No	To be supplied as part of the data ingest input, or introspect pkg.			
Required uncompression package	Name of software package required to uncompress the files; e.g. graphics viewer package, PKUNZIP, etc.	This would be another instance of a file dependency.		Yes	To be supplied as part of the data ingest input, or introspect pkg.			
Location								
Location	physical/logical location of file on archive system - an identifier (i.e. absolute path, PURL, CRID)	This could be the physical location of the file/object in your physical storage system. However, there are logical storage systems that provide physical location transparency to provide more flexibility with			Archival Management System			

		moving digital materials from one physical device to another. May be a logical location in that case.						
Preservation Description Information								
PDI: Context Information								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*		*FGDC*	*Access xwlk*
Identifier Name	Name of **type of identifier**; e.g. CRID,URL, File Path, ISBN, Author, Title, etc.	There are several techniques for identifying a file/object within a system, and several may be used within the same archival system. This identifies the identification mechanism. Preference must be given to persistent and unique identifiers.		No	Archival Management System			
Identifier Value	Value of identifier for this package.	This is the value of the identifier in the form indicated above.			Archival Mgmt System			
Purpose of the Object	What is the file/object's purpose/function? Meaning or essence of the data? What is the significance of the data? Intended community / audience			No	To be supplied as part of the data ingest input, FGDC Metadata	DC.Description	<abstract> 1.2.1 <purpose> 1.2.1	Purpose
Intention Type	Intended use of a particular manifestation (e.g. preservation master, access copy)			No	To be supplied as part of the data ingest input,			
Related Information Describes entities that are related to the Preservation Master file/object								
Related item	Items related to this one: e.g. list of CRID/URNs of items	General archival artifacts: Original object from which preservation master was created.		Yes	To be supplied as part of the data ingest input,	DC.Relation		

		Access Copy derived from preservation master Geospatial archiving may also produce other derivative items such as PDFs of shapefiles, HTML of geo metadata xml file, ZIP for distribution package.			Archival Mgmt System			
Relation to (Preservation Master) Resource	Description of the relation to an item: e.g. it is an element/component of the object (e.g. when describing the individual file types that comprise a geospatial dataset), or it is the access copy of the object, or it is a derivation of the object.	Describes a hierarchy/network of relations: we may decide that a textual description at collection level may suffice, if this is too much to model. Geospatial derivatives may include a geospatial pdf that allows dataset access through Adobe Reader, or a stylized HTML version of the geospatial xml metadata file, or a mosaic of orthoimages.		Yes	To be supplied as part of the data ingest input.			
Is item required?	Will the object in question make sense without it, e.g. links required to understand object within collection. **Yes/No** value.			No	To be supplied as part of the data ingest input.			
PDI Provenance Information This documents the history of the Content Information; i.e. **who has had custody of the Content Information since its origination,** and what was its source. This would include information about the processes that generated it. If the content was not "born digital", information about digitisation processes could be held here. Also, custodial information is also likely to refer to the rights implications regarding access to the content. This section will also record rights information.								
PDI Provenance: Originator Information (History of Origin)								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Reason for Creation	The reason the original record was initially created.	May be similar to the **Purpose of the Object,** but may be different. For example, an object may be created to address legal mandate, but the purpose of the object is to		No	To be supplied as part of the data ingest input. FGDC Metadata	DC.Provenance	<purpose> 1.2.1	

		record and visualize key geospatial information and/or relationships.						
PDI Provenance: Custody History								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Originator / Record Creator / Author	The name of the author(s) or originator of the record	There may be co-authors for a particular work.			FGDC metadata	DC.Creator	<origin> 1.1.8.1	Originator / Author
Originator / Record Creator / Author Contact Details	Contact details (address, email, fax, telephone number) for the originator	Name the Originator is affiliated with - institution/agency originator is affiliated with \- organization / department - address, phone, fax, email		Yes				
Date Record Created	Date the record that is being preserved was originally created	An item may be created and then go through a review/edit cycle before being finalized. This date should reflect the date the object's content was finalized and released.		Yes		DC.Date.Created DC.Date.Modified		
Contributor(s)	Other acknowledged contributors to the record	For example, research assistants, illustrators, etc. may be acknowledged contributors to the work/record, but who are not the author/originator of the record.		Yes	FGDC metadata	DC.Contributor	<procstep> <proccont>	Contributor
Publisher Name	The name of a publisher. e.g. Oxford University Press	The agency, institution, or organization may publish the materials on behalf of the originator. May be the same agency the author is affiliated with, noted above, but may be different.		Yes	FGDC metadata	DC.Publisher	<publish> 1.1.8.8.2	Publisher
Publisher Contact Details	Contact details (address, email, fax, telephone number) for the publisher	Institution Name, Contact - contact name - organization / department - address, phone, fax, email		Yes	FGDC metadata		<pubplace> 1.1.8.8.1	

Date Record Published	Date the record that is being preserved was published	A publisher will go through a process and the publication date will lag the record creation date. For Geospatial datasets, the recorded Publication Date may serve as the Date Record Created		Yes	FGDC metadata	DC.Date.Issued (=publication date) DC.Date.Available	<pubdate> 1.1.8.2	Date Published
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Distributor	The name of the distributor	Geospatial metadata distinguishes the "Publisher" from the "Distributor". An agency may have published the dataset, but a GIS clearinghouse or archives may be the distributor for the dataset. The archives may supersede a Distributor documented in the ingested geospatial metadata file.		No	FGDC metadata		<distrib> 6.1 <cntinfo> <cntorgp> <cntpos> <cntaddr> <addrtype> <address> <city><state> <postal> <cntvoice> <cntemail>	Distributor
Rights Management	Text field or link to rights management information			Yes	Archival Mgmt System	DC.Rights		
Access Profile		Technical or implementation details describing how to interface / access the information in the record		Yes	Archives Policies			
Access Rights	Public Access Restricted access for research purposes only	Type of access allowed, and to whom access is granted. The access constraints documented in a geospatial dataset may be superseded by the Archives' access rights as the dataset is accessioned into the archives.		Yes	Archives Policies	DC.Rights.AccessRights	<accconst> 1.7	Access Constraints
Change History **Before** Archiving	Captures history of the object prior to its submission to the archives. Information about the events, parameters, and source data which constructed the data set prior to archival ingestion,	Geospatial dataset metadata standard includes section to record the lineage, data sources, and processing steps performed to create the dataset.		Yes	FGDC metadata		<procstep> 2.5.2 <procdesc> 2.5.2.1 <sourcecit> 2.5.2.2 <procdate> 2.5.2.3	

	and which need to be retained.							
Info Source for object	Source from which the information was derived	Geospatial datasets are often derived from other data sources, and this should be documented in the geospatial metadata file.		Yes	FGDC metadata		<srcinfo> 2.5.1	
PDI Provenance: Preservation Strategy History (Reason for Preservation)								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Reason for Preservation	- statute / legislation - historical value - research value - evidentiary value	Provide reason this item has been selected for preservation		Yes	Archives Policy			
Institution Responsible for Archiving **Decision**	Name of agency responsible for deciding that this work/ manifestation should be archived.			No	Archives Policy			
Institution with **Preservation Responsibility**	Name of agency responsible for preserving this item Institution name, Contact info (address, email, fax, telephone number)	Implies ongoing management of the object as well		No	Archives Policy			Hosted By
History: Change: Preservation Strategy:	URI/URL/PURL or pointer to the policy document.	This would link to **policies** which would describe/detail the preservation strategy.		No	Archives Policy			
Removal Authority	This should point to a user ID within the preservation system / environment.	Defines who is authorized to purge the record from the archive system		No	Archives Policy			
Retention Policy Citation	Textual information of organization's policy for record retention, e.g. link to relevant document or	Refers to the documentation that documents the decision that a particular class of records should be archived.		No	Archives Retention Schedule			

	inclusion of relevant text in metadata. Could be a URI.							
Retention Authority Issuance	Textual information on legislative /government laws /regulations for record retention	Attribute based on evidential requirements framework.		No	Archives Policy			
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Retention Period - Duration - End Date/Expiry	Date which indicates the end of the retention period (e.g. 2012-03-99 or permanent/indefinite)	May include a link to the retention schedule for the item. Date format should follow ISO 8601 date format		No	Archives Policy, Retention Schedule			
Selection Status	Status of item's selection within archive (e.g. selected, not selected, active, withdrawn, disposed)	Allows the archives to keep track of what records have been withdrawn from archive, while retaining the historical archival record for an item that may have been withdrawn.		No	Archives Selection Processing			
Retention Review - Next Review Date	Date the item will next be reviewed What may be reviewed are aspects of the object such as - retention schedules, usage, legislation/statues to determine if retention should be continued.	Offers next scheduled review date so management reports can be generated to manage review activities		Yes	Archives Policy Archival Mgmt System			
Review Reason	Reason for review			Ye	Archives Policy			
Retention Review - Review History	Dates the item was last reviewed, the reason for the review, and the outcome of the review (e.g. file fixity confirmed, retain until next review date, dispose)	Provides historical retention review trail for the item.		Yes	Archival Mgmt System			
PDI Provenance: Capture Procedure (Ingest Process History)								

Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
History: Preservation Submission: Date	Date when the **digital record** was submitted to the archives by the contributor.	Date format should follow ISO 8601 date format		No	To be determined as part of the data ingest input.			
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Preservation Action Permission	A statement of whether or not permission is held to create copies of the object for preservation purposes.	Useful administrative metadata to document the archives' right to reproduce the record for preservation purposes.		No	Submission Agreement			
Capture Procedure Type	Describes **how the object to be archived was captured**.	Describes the preservation institution's ingest process that transfers the object from the contributor to the archives.		No	To be determined as part of the data ingest process.		<proc step>	
Ingest Object	Record submission package format (e.g. zip, tar, bagit) or file type if an individual file Submission item size	Provide a pointer to the original ingest object that should be preserved within the archival repository.		No	To be determined as part of the data ingest process.			
Ingest Process	All relevant details of any process applied to a digital object or file, including s/w, specific settings, actions rqd to produce the current (preservation) manifestation	- Description of process - Name of Agency responsible for process - Contact information for person who performed the ingest - critical h/w used in the process - critical s/w used in the process - how process was carried out-describe significant steps - description of the ingested object (e.g. zip archive, BagIt package) - date & time of processing - original submission package information - result of process (success/failure,			To be determined as part of the data ingest process.		<procstep> log the archives' processing activities <transize> 6.4.2.1.7	

		file(s) generated) - changes that were made to the collection, object, or file by the process - other (any other misc details) Archives should add a process step to the geospatial metadata file recording the transfer, ingest, and archival processing of the geospatial dataset.						
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Preservation Master Create Date	Date the Preservation Master was produced	Date format should follow ISO 8601 date format		No	To be determined as part of the data ingest process.			
Preservation Object Derivatives	e.g. Access Copy (compressed version of Preservation Master) .PDF of shapefile .ZIP for shapefile access copy .HTML ver of Metadata	Document derivative artifacts created as part of the ingest process - mechanism to create derivative		Yes	To be determined as part of the data ingest process.	DC.relation.isReferencedBy		
Gathering Frequency	Sampling techniques (e.g. weekly, monthly, quarterly, annually, one-off)	Used for archives that have automated metadata harvesting mechanisms.		Yes	Archival Policy	DC.AccrualPeriodicity		
Gathering Regime	e.g. the frequency and extent of capture of a publication	Used for archives that have automated metadata harvesting mechanisms. Geospatial datasets may be published through Geospatial Inventory or Portals, and may be remotely accessed and downloaded. Archives may explore these as a means of automatically harvesting published geospatial datasets.		Yes	Archival Policy	DC.AccrualPolicy		
PDI Provenance: Rights information								

This section records information related to transfer of the material, and rights, from the donor to the preserving organization								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Source / Donor	Name of person/agency responsible for donating this material to the archives. Individual/Instit name, Contact info (address, email, fax, telephone nu)	Who donated the materials to the archives		Yes	Submission Agreement			Source/Donor
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Acquisition Event Type	Type of event in the rights acquisition process (e.g. Publication selected, No publisher response after x weeks, Publisher Permission obtained, Null publisher response)	Provides a management history trail of the acquisition life cycle.		Yes	Archival Submission Management			
Acquisition Negotiation Status	Type of event in the rights negotiation process (e.g. First Contact Made, Follow-up contact made, Agreement Finalized, On hold pending further negotiation, Archived without agreement, De-selected)	Note: event type and negotiation status are interlinked, the PANDORA model examples are cited. Organization (body which negotiates on behalf of publisher – e.g. ASCAP in the case of musicians)		Yes	Archival Submission Management			
Copyright Warning	Copyright statement / restriction for item being submitted.	May require input from organizational representative responsible for overseeing copyright details.		Yes	Submission Agreement	DC.Date.Date Copyrighted		Copyright
Publisher Copyright Statement	Statement from publisher regarding copyright. - Name of publisher - date of publication - place of publication	May require input from organizational representative responsible for overseeing copyright details.		No	Submission Agreement	DC.Publisher		
Rights clearance	Describe how Archives	Record:		No	Submission			

statement	have been granted rights to this object e.g. item not restricted by copyright, established license agreement (provide license information), permitted by statute	Who checked the license clearance. Date license clearance was determined Date license clearance was established Or reference boilerplate that indicates that there are no copyright restrictions on the archives to create copies of the object for preservation and/or access purposes.			Agreement			
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Citation Statement	Statement directing archives' end user how to cite this item	NOTE: while the submitted geospatial metadata may include a citation statement, the archives will likely provide guidance for an end user on how s/he may cite the use of materials accessed through the archives, as the ownership of the materials has transferred to the archives.		No	Archival Policy		<useconst> 1.8	Citation
User Class	e.g. from a list of designated user types – t	This would need to be developed in conjunction with the representative in the archives that oversees the definition and implementation of user rights and user access.		No	Archival Policy	DC.Audience		
Restriction Status	(Terms & Conditions – e.g. restricted only to archivists, i.e. no public access)	This would need to be developed in conjunction with the representative in the archives that oversees the definition and implementation of user rights and user access. At a minimum – a boilerplate statement indicating item is available for free public access with no restrictions.		No	Submission Agreement Archival Policy	DC.Rights.AccessRights		
Access Conditions	e.g. may only be accessed only by users affiliated with an authorized universities;	This would need to be developed in conjunction with the representative in the archives that oversees the		No	Submission Agreement Archival Policy	DC.Rights.AccessRights		

	may require special software to access	definition and implementation of user rights and user access. May use this to document special software needed to access the geospatial datasets. At a minimum – a boilerplate statement indicating item is available for free public access with no special access conditions.						
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Access Restrictions	e.g. embargoed for # years	NOTE: while the submitted geospatial metadata may include an access constraint statement, the archives will likely supersede with archives-specific statement based on the archives' policies and procedures, as the ownership of the materials has transferred to the archives. At a minimum – a boilerplate statement indicating item is available for free public access with no special access restrictions.		No	Submission Agreement Archival Policy	DC.Rights.AccessRights	<accconst> 1.7	Access Restrictions
Use Conditions	special software required to use	This would need to be developed in conjunction with the representative in the archives that oversees the definition and implementation of user rights and user access. Geospatial Datasets will often require some sort of Use statement because it often requires special software to access/manipulate the dataset. At a minimum – a boilerplate		No	Submission Agreement Archival Policy			

		statement indicating item is available for free public access with no special use conditions.						
Use Restrictions	e.g. may be used only for research or educational purposes	NOTE: while the submitted geospatial metadata may include a use constraint statement, the archives will likely supersede with archives-specific statement based on the archives' policies and procedures, as the ownership of the materials has transferred to the archives. At a minimum – a boilerplate statement indicating item is available for free public access with no special use restrictions.		No	Submission Agreement Archival Policy		<useconst> 1.8	Use Restrictions
(Archives) Distribution Liability Statement Disclaimer	include any necessary liability disclaimer statement	NOTE: while the submitted geospatial metadata may include a distributor liability statement, the archives will likely supersede with archives-specific statement based on the archives' policies and procedures, as the ownership of the materials has transferred to the archives. At a minimum – a boilerplate statement indicating no liability.		No	Submission Agreement Archival Policy		<distliab> 6.3	Use Restrictions
User Authorization	Groups / Individuals Access Rights	Relates to: - Original submission - Preservation Master - Copies of Preservation Master - Access Copies At a minimum – a boilerplate statement indicating access for "public" or "anonymous" users.		Yes	Submission Agreement Archival Policy			
PDI Provenance: History: Change: Formatting (since item has come to Archives and after original Preservation Master created):								

The RLG metadata refers to "Change History" - changes made to the formatting of the data under a migration strategy, or to decisions regarding its distinct preservation. If the Current File/Format fields have a "no" value, then information about the file/format type (and version) converted to; reason for conversion; date converted; and conversion agent is also captured. It can be envisaged that this information will be retained as conversions take place over time. This facilitates a history trail to be kept of files/formats. (e.g. handle upgrades to newer versions, transformations to other more reliable formats)								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
FormattingChange: date	Date of change, in ISO 8601 format: (CCYY-MM-DD). Example: 1985-04-12			Yes	Archival Preservation Reformatting Policy		<procdate> 2.5.2.3	
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
FormattingChange: previous format	The previous format which data was stored in.	This may be a redundant field; if you have date and formats held, could be inferred through a query.		Yes	Archival Mgmt System			
FormattingChange: process	Process undergone to change formats.			Yes	Archival Preservation Reformatting Process		<procstep> 2.5.2 <procdesc> 2.5.2.1 <sourcecit> 2.5.2.2	
FormattingChange: reason	Reason for changing format.	This could be a change in storage policy, or file format in danger of being outdated.		Yes	Archival Preservation Reformatting Policy		<procdesc> 2.5.2.1	
FormattingChange: conversion agent	Details of who carried out the work.	Name, Address, Phone		Yes	Reformatting Process input.		<proccont> 2.5.2.6	
PDI Provenance: Format Processing Records the formal states a tangible format undergoes before and/or during each time its content is **made available to user**; e.g. testing by installation and virus checking, conversion from TIFF to JPEG format, rendering from SGML to HTML.								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Process Type	The processes a given tangible format may undergo	Changes in format imply changes required in format processing. Format support and (Re)Format processing trends and options		Yes	Archival Preservation Reformatting Process			

		would need to be monitored, a "technology watch" facility would be useful to update this.						
Process Utilities	Documents what particular utilities have been used (e.g. Dr Solomons, Norton, TIFF2JPG utility, MS PhotoShop)	Also, changes in format imply changes required in format processing utilities.		Yes	Archival Preservation Reformatting Process			
Preferred Browser or Viewer for the Format	Recommended version of browser or software to view the format; e.g. Internet Explorer v9.0, Adobe Reader, JPEG viewer			Yes	Archival Preservation Reformatting Process			
PDI Provenance: Usage This section records data related to the end user usage of the record, to maintain a usage history as part of provenance.								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Use-Type	Description of usage of the item - e.g. view, download, copy, edit, file, index, classify, send, dispose – identifies usage types permitted.	This will require a set of keywords to be identified		Yes	Archival Policy			
Use Instance Time	Records when the data was used (ISO 8641), e.g. 09/02/99 15:45.	This looks to be elements that would be tied to some sort of online usage tracking application.		Yes	Archival Access System			
Use Instance User	who used the data – although this could be just at the level of organisation; e.g. user ID number, guest, email of user	The strategy for tracking the "user" and what constitutes the "user" would need to be developed in conjunction with the representative in the archives that oversees the definition and implementation of user rights and user access. At what level would users be defined (individuals, groups, geography, anonymous? Need to		Yes	Archival Access System			

		check with Rights (Content Issues) Working Group). Usage may be of interest if it is a contributing factor to continuing preservation of the object.						
PDI: Fixity / Validity Information This documents authentication mechanisms, and provides any authentication keys which may be needed, for example: Checksum, Validation Keys, Encryption, Digital watermarks								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Authentication Type	The authentication/ validation mechanism in use: e.g. checksum, hash, digital watermark			N	To be determined as part of the data ingest process.			
Authentication Key	The authentication key (if appropriate)			N	To be determined as part of the data ingest process.			
Authentication Value	The value e.g. a hash value			N	To be determined as part of the data ingest process.			
Authentication Documentation	Information on, or link to information, on the authentication mechanism in use			N	Archives			
Frequency of Validation Status Check	Number of days cycle to validate the files' fixity values.	For systems with the ability to automate a validation check. Designate the frequency fixity		N	Archival Policy			

		validation checks should be performed.						
Validation Check Status	Status of Validation Check - Date - Validation Result - Remediation performed if failed	Holds the status of each fixity validation status check. Time to do validation status check is defined by the Frequency of Validation Status Check field.		Y	Archival Mgmt System			
Packaging Information This is defined as information which, logically or actually, binds and relates the components of the package into an identifiable entity on specific media, e.g. on a CD-ROM, packaging info may include the ISO-9660 volume/file structure of a CD-ROM. These choices are the subject of local archive definitions or conventions.								
Descriptive Information Descriptive metadata is used to facilitate end user and use of the archived records. Several of these fields may be derived from the Representation information - so file identification/location, and other descriptive information will be replicated here, and will be noted as a <DUPLICATE>.								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Title	Title of archived information object			Yes	FGDC Metadata	DC.Title	<title> 1.1.8.4	Title
Originator / Record Creator / Author <DUPLICATE>	Creator of intellectual content of the archived information object	- Institution Name - Contact Information		Yes	Archival Mgmt System sourced from FGDC Metadata	DC.Creator	<origin> 1.1.8.1	Originator / Author Index Terms - CorpNames
Contributor(s) <DUPLICATE>	Other contributors to the intellectual content of the archived information object.	FGDC process steps may note other contributors.		Yes	Archival Mgmt System sourced from FGDC Metadata	DC.Contributor	<procstep> <proccont>	Contributor
Publisher <DUPLICATE>	Publisher of the archived object	Geospatial metadata distinguishes the “Publisher” from the “Distributor”. An agency may have published the dataset, but a GIS clearinghouse or archives may be		No	Archival Mgmt System sourced from FGDC Metadata		<publish> 8.8.2	Publisher

		the distributor for the dataset.						
Distributor <DUPLICATE>	Distributor of the archived object, and contact information	Geospatial metadata distinguishes the "Publisher" from the "Distributor". An agency may have published the dataset, but a GIS clearinghouse or archives may be the distributor for the dataset. The archives may supersede a Distributor documented in the ingested geospatial metadata file.		No	Archival Policy		<distrib> 6.1 <distrib> 6.1 <cntinfo> <cntorgp> <cntpos> <cntaddr> <addrtype> <address> <city> <state> <postal> <country> <cntvoice> <cntemail> <hours>	Distributor
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Description	Brief textual description of the archived information object.	This could contain an abstract.		Yes	Archival Mgmt System sourced from FGDC Metadata	DC.Description DC.Description.Abstract	<purpose> 1.2.2 <abstract> 1.2.1	Description Abstract
File Type <DUPLICATE> (sometimes referred to as Format)	The name of the file type. e.g. TIFF HTML Proprietary Collection GIS Specific data file types: ShapeFile OrthoPhoto Digitized Map GIS Project File Geodatabase Personal File GeoDB ArcSDE DB	A reference list of file formats would need to be maintained; formatting metadata may be required to understand how logical access to the bit-stream takes place. Collection would indicate this particular package contains a collection of files		Yes	Archival Mgmt System	DC.Format *Mime Type*		Format
Object (File) Size	The size of the file in			No	Archival Mgmt		<transize>	

<DUPLICATE>	kilobytes, megabytes, etc.				System		6.4.2.1.7	
Date Record Created (as opposed to the date the ACCESS copy of a preserved record was created) <DUPLICATE>	Date the record that is being preserved was originally created. ISO 8601 format.	An item may be created and then go through a review/edit cycle before being finalized. This date should reflect the date the object's content was finalized and released.		Yes		DC.Date.Created DC.Date.Modified	Not Available	
Date Record Published <DUPLICATE>	Date the record that is being preserved was published, ISO 8601 format.	A publisher will go through a process and the publication date will lag the record creation date. For Geospatial datasets, the recorded Publication Date may serve as the Date Record Created		Yes	FGDC metadata	DC.Date.Issued (=publication date) DC.Date.Available	<pubdate> 1.1.8.2	Date Published
Time Period of Data	Provides the time period for the data	Form could be in: - Single Date - Multiple Dates - Range of Dates Many geospatial data sets include data gathered over a period of time, rather than on a specific date. The time period of the data may not be the same as the date the dataset was published. For example a dataset may cover 2001-2005 and not be published until 2006.		No	FGDC metadata	DC.Coverage. Temporal	<timeperd> 1.3 <timeinfo> 1.3.9 <rngdates> 1.3.9.3 <begdate> 1.3.9.3.1 <enddate> 1.3.9.3.2	Years
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Date Access Copy created	Date the access copy was created	May just be creating a 'dark archive' in which case an access copy is not created.		Yes	Archival Mgmt System	DC.Date.Created DC.Date.Modified		
Language <DUPLICATE>	Language of resource. Probably in NISO X39.53.			Yes		DC.Language	ESRI: <langdata> = en	Language
Finding and Searching	References to any system			Yes				

Aids	or method used to enhance access to information within the digital object, which need to be maintained in successive generations.							
Theme / Subject Keywords	Enables search for datasets by theme, subject	Geospatial Data should include keywords that represent the RAMONA Data Categories and Data Layers Reference: http://www.sco.wisc.edu/wisclinc/Layer_Assignments_060206.pdf		Yes	FGDC metadata	DC.Subject	<themekt> 1.6.1.1 <themekey> 1.6.1.2	Index Terms - Subjects - CorpNames Subjects (LOC Subject headings)
Place Keywords	Enables search for datasets by place	e.g. state, county, city / municipality		Yes	FGDC metadata	DC.Coverage. Spatial	<placekt> 1.6.2.1 <placekey> 1.6.2.2	Coverage Index Terms -GeoNames
Descriptive: Geospatial Datasets (content dependent descriptive fields)								
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Geospatial Data Presentation Form <DUPLICATE>		e.g. “vector digital data” Aerial Photographs (AAT) Shapefile (non-AAT) MrSID (non-AAT), GeoPDF					<geoform> 8.6	Physical Desc File Characteristics
Stratum Keywords		Additional Geospatial-specific keywords that can inform search.					<stratumkt> 1.6.3.1 <stratumkey> 1.6.3.2	
West_Bounding_Coordinate		May facilitate map-oriented search & access of geospatial datasets					<westbc> 1.5.1.1	West Bounding Coordinate
East_Bounding_Coordinate		May facilitate map-oriented search & access of geospatial datasets					<eastbc> 1.5.1.2	East Bounding Coordinate
North_Bounding_Coordinate		May facilitate map-oriented search & access of geospatial datasets					<northbc> 1.5.1.3	North Bounding Coordinate
South_Bounding_Coor		May facilitate map-oriented search					<southbc>	South Bounding

dinate		& access of geospatial datasets					1.5.1.4	Coordinate
Point of Contact	Point of contact for information on dataset Name, address, phone, email	Archives may consider superseding the metadata-provided point of contact with an archives-staffed contact.					<ptcontac> 1.9	
Raster (ortho, elevations) Spatial resolution <DUPLICATE>	Provides ground resolution of the pixels e.g. 0.25 feet 1 foot 0.08 meters 0.25 meters 1.00 meters	More relevant for raster datasets than vector datasets. Provides ground resolution of the pixels in the raster dataset.		No	Archival Mgmt System sourced from FGDC Metadata		<abstract> 1.2.1	Resolution
Element	*Examples/Definitions*	*Comments*	*Req.*	*Rep.*	*Source*	*Dublin Core*	*FGDC*	*Access xwlk*
Attribute Label Attribute Definition	label: DISTRICT def: District Identification number label: DNCCON_ label: DNCCON_ID label: SHAPE_Area label: CO_NAME label: CO_ABBR	Geospatial datasets contain an underlying database of fields (columns) and values for those fields. As users are searching for datasets and evaluating search results, they will often want to explore the data fields to understand the data that is held within the datset. Note: There are no naming standards for field names, and many fields have non-intuitive names and may not include an accompanying attribute definition. Therefore, it would be very difficult for a user to conduct a search for datasets based on field names. For example, a researcher may be looking for datasets that		Yes	FGDC Metadata		<attrlabl> 5.1.2.1 <attrdef> 5.1.2.2	Attributes

		include an attribute that identifies the counties in a state. One dataset may call the field “COUNTIES”, another may call it “CO_,” or CO_NAME, and may or may not include an attribute definition that indicates this field represents county names. Yet another may use numerical County IDs to identify the fields. We recommend you provide an HTML-stylized version of the xml metadata file through the dataset access interface so that searchers can browse the metadata file without having to download the entire dataset to assess their data needs.						
Geospatial Metadata File	Link to file containing user-friendly manifestation of the geosopatial metadata	Provides end user with access to all of the geospatial metadata.					<geoform> 8.6	Geospatial Metadata File

Appendix A: Geospatial Archival Metadata - summary list

geospatial-specific information is indicated in green font color

	Metadata Field	FGDC	Required Y/N
1	**Content Information**		
2	Object (File) Identifier		Y
3			
4			
	Representation Information		
	General		
5	Language expect most of our data will be English based.		
6	Storage (media) format (of the original data): Type Generally not expected to be necessary for geospatial, assuming transfer either via network, or with a portable disk whose format won't impact the long term perservability of the data set. Assume the transfer medium is not holding the preservation copy.		
7	Storage format: Version	<native>	
8	Storage format: Specification (documentation for source format)		
9	"Object" File Modality, Structural Type (text, graphic, sound, video, GIS)	<natvform>	
10	Name of File Type (Object Type) (TIFF, HTML, Proprietary, Collection) Also known as Format IANA MIME Media Types: http://www.iana.org/assignments/media-types/index.html Dublin Core Type: http://dublincore.org/documents/dcmi-type-vocabulary/ GIS possible File Types: Shapefile, Orthophoto, Digitized Map, GIS Project, Geodatabase, File Geodatabase, Personal File Geodatabase		
11	Version of File Type		
12	Number of Files (in preservation master record) CONDITIONAL if multiple files		
13	Inventory of Files (in object) CONDITIONAL if multiple files		
14	Total File (Object) Size	<transize>	
15	Documentation, Technical Infrastructure of Complex Object		
16			
17			
18			
19			
20			
	Text-based Data (content dependent)		
21	Character Set		
22	Associated DTD		
23	Encoding		
	Graphic Data (content dependent)		
24	Resolution		
25	Color Depth		
26	Color Management		
27	Color Bar / Grey Scale Bar		
	Audio Data (content dependent)		
28	Sampling Frequency		
29	File Encoding		
30	Bits		
31	Mono/Stereo		
	Geospatial Data (content dependent)		
32	Projection	<mapprojn>	
33	Direct Spatial Reference Method controlled vocabulary defined by the FGDC standard:	<direct>	

	Metadata Field	FGDC	Required Y/N
	Point, Vector, Raster		
34	Geospatial data presentation form controlled vocabulary per FGDC: Anglo-American Committee on Cataloging of Cartographic Materials, 1982: interpretation of AACR2. ALA *atlas, audio, diagram, document, globe, map, model, multimedia presentation, profile,* ***raster digital data****,* ***remote-sensing image****, section, spreadsheet, tabular digital data,* ***vector digital data,*** *video, view, free text.*	<geoform>	
35	West Bounding Coordinate	<westbc>	
36	East Bounding Coordinate	<eastbc>	
37	North Bounding Coordinate	<northbc>	
38	South Bounding Coordinate	<southbc>	
39	Spatial Reference Information - Projected Coord System Name - Geographic Coord System Name - Geodetic model Horizontal Datum Name - Geodetic model Ellipsoid name	<spref> <horizsys> <planar> <gridsys> <geodetic> <horizdn> <ellips>	
40	Raster (ortho, elevations) Spatial Resolution	<abstract>	
41	Geospatial Metadata Standard Name	<metstdn>	
42	Geospatial Metadata Standard Version	<metsdv>	
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
	Technical Dependencies		
61	Application Dependency (Yes/No)	<native>	
62	Application Name used to **create** resource (vendor, app name, version)	<native> <techprereq>	
63	Application Name used to **render** resource (vendor, app name, version)		
64	Application Notes		
65	Software Environment Dependency (Yes/No)		
66	Software Environment to create resource (vendor, sw name, version) CONDITIONAL if SW dep=Y	<techprereq>	
67	Software Environment to render the resource (vendor, sw name, version) CONDITIONAL if SW dep=Y		
68	Digital Format for access copy		
69	Hardware Dependency (Yes/No)		
70	Modality Served (text, graphics, audio, video, input) CONDITIONAL if HW dep=Y		
71	Hardware to **create** resource (vendor, name, model, desc, spec, ver) CONDITIONAL if HW dep=Y	<techprereq>	

	Metadata Field	FGDC	Required Y/N
72	Hardware to **render** the resource (vendor, name, model, desc, spec, ver) CONDITIONAL if HW dep=Y	<techprereq>	
73	Installation requirements / Installation Process	<techprereq>	
74	Access inhibitors (encryption, watermarking, etc.) CONDITIONAL if applicable		
75			
76			
77			
78			
79			
80			
	Compression		
81	Compression (Yes/No)		
82	Compression Name CONDITIONAL if Compression=Y		
83	Compression Method CONDITIONAL if Compression=Y		
84	Required uncompression package CONDITIONAL if Compression=Y		
	Location		
85	Location (physical location of file on archival system, i.e. CRID, PURL)		
86			
87			
88			
	Preservation Description Information		
	PDI: Context Information		
89	Identifier Name (e.g. CRID, URL, File, Path, ISBN, Purl, etc.)		
90	Identifier Value		
91	Purpose of the resource What is the file/object's purpose/function? Meaning/essence of the data?	<abstract> <purpose>	
92	Intention Type (preservation master, access copy)		
93			
	Related Information		
94	Related item		
95	Relation to resource CONDITIONAL if related resources exist		
96	Is the related item required (Yes/No) CONDITIONAL if related resources exist		
	PDI Provenance Information		
	Publisher Information (History of Origin)		
97	Reason for Creation May be different than Purpose of the Resource (#91). e.g. Resource may have been created to satisfy a legal mandate.	<purpose>	
	Custody History		
98	Originator / Record Creator / Author	<origin>	
99	Originator / Record Creator / Author - Contact Details		
100	Originator / Record Creator Contact Details		
102	Date Record Created - date record was originally created that has been submitted		
103	Contributor(s)	<proccont>	
104	Publisher Name	<publish>	
105	Publisher Contact Details	<publish>	
106	Date Record Published	<pubdate>	
107	Distributor	<distrib>	
108	Rights Management		
109	Access Profile		

	Metadata Field	FGDC	Required Y/N
110	Access Rights	<accconst>	
111	Change History **Before** Archiving	<procstep>	
112	Info Source for object (geospatial datasets are often derived from other data sources -- but this will be documented in the geospatial metadata file)	<srcinfo>	
113			
	Preservation Strategy History (Reason for Preservation)		
114	Reason for Preservation		
115	Institution Responsible for Archiving Decision - name of agency		
116	Institution with Preservation Responsibility - Instit Name - Contact - Contact name - address - phone - email - fax		
117	History: Change: Preservation Strategy - link to changes in policy which affect the preservation strategy		
118	Removal Authority - who is authorized to purge item from archive		
119	Retention Policy Citation - text info of org's policy for record retention, could be a URL		
120	Retention Authority Issuance: text info on leg/gov laws/regs for rcd retention		
121	Retention Period - Duration (e.g. permanent, 10 years, 50 years, 100 years) AND/OR - End Date/Expiry (e.g. Dec 31, 2050)		
122	Selection Status (e.g. withdrawn, selected, not selected)		
123	Retention Review Date - date item will next be reviewed CONDITIONAL if to be reviewed		
124	Retention Review Reason - what may be reviewed CONDITIONAL if to be reviewed		
125	Retention Review History		
126			
127			
128			
	Capture Procedure		
129	History: Preservation Submission - Date (date digital prsvtn copy was submitted to the archives by the contributor)		
130	Preservation Action Permission - statement of whether or not permission is held to create copies of the object for preservation purposes		
131	Capture Procedure Type - describes how obj to be archived was originally created	<procstep>	
132	Ingest Object - Description (Record Submission Pckg format (e.g. zip. tar, bagit) or file type of an individual file		
133	Ingest Process - all relevant details of any process applied to a digital obj or file, including s/w, specific settings, actions rqd to produce the current preservation manifestation Archives may consider augmenting the geospatial metadata file and adding their ingest processing steps.	<procstep> <transize>	
134	Preservation Master Create Date		
135	Preservation Object Derivatives - document derivative artifacts created as part of the ingest process, and mechanism to create derivative There may be no derivatives		
136	Gathering frequency - sampling rate (weekly, monthly, quarterly, annually, once) - used for archives that have automated harvesting mechanisms		
137	Gathering regime - what/how materials are being captured - used for archives that have automated harvesting mechanisms Useful for automating harvests of materials. Consider web archiving, archiving datasets from published geo WMS servers.		

	Metadata Field	FGDC	Required Y/N
138			
139			
140			
141			
	Rights Information		
142	Source/Donor		
143	Acquisition Event Type (tracks the communication exchange with the item source (e.g. pub selected, no publisher response after x weeks, pub permission obtained, no publisher response))		
144	Acquisition Negotiation Status (current state in workflow (e.g. first contact made, follow-up contact made, agreement finalized, on hold pending further negotiation, archived without agreement, deselected)		
145	Copyright Warning		
146	Rights Clearance Statement		
147	Citation Statement - statement directing user how to cite this item	<useconst>	
148	User Class (controlled vocabulary of user groups to whom rights can be granted)		
149	Restriction Status - terms & conditions (e.g. restricted only to archivists, public access under xyz legislation, no public access)		
150	Access Conditions - terms & conditions (e.g. may only be accessed by universities)		
151	Access Restrictions - terms & conditions (e.g. available, embargoed until x years after death, embargoed for x years)	<accconst>	
152	Use Conditions - terms & conditions (e.g. only for research purposes)		
153	Use Restrictions - special software required to use	<useconst>	
154	(Archives) Distribution Liability Statement Disclaimer	<distliab>	
155	User Authorization - groups/individuals access rights relates to original mtls submitted, preservation master, preservation master copies, access copies. At a minimum a “boiler plate” statement indicating access for “public” or anonymous users.		
156			
157			
158			
159			
160			
	History: Change formatting (since item has come to archives and original preservation master created		
161	Formatting Change: Date - date of change CONDITIONAL if reformatted	<procdate>	
162	Formatting Change: Previous format		
163	Formatting Change: Process	<procstep>	
164	Formatting Change: Reason	<procdesc>	
165	Formatting Change: Conversion Agent - details of who carried out the work	<proccont>	
166			
167			
168			
169			
170			
171			
172			
	Format Processing		
173	Process Type - the processes a given tangible format may undergo		
174	Process Utilities - documents what particular utilities have been used		
175	Preferred Browser or Viewer for the Format		
176	Use Type - description of usage of the item (e.g. view, download, copy, edit, index, classify, send)		
177	Use Instance time - records when the data was used place to record resource usage w/in the archival data management system.		

	Metadata Field	FGDC	Required Y/N
178	use Instance user - who used the data (user id name, user id number, "guest", email of user, etc.) place to record resource usage w/in the archival data management system.		
179			
180			
181			
182			
183			
	PDI: Fixity / Validity Information		
184	Authentication type - mechanism (e.g. checksum, hash, digital watermark)		
185	Authentication key - if applicable to authentication mechanism		
186	Authentication value (e.g. the value of the hash) There may be many files associated with a single resource, so there may be multiple authentication values associated with a single resource		
187	Authentication documentation - info or link to the authent mechanism used		
188	Frequency of Validation Status Check - basis to update the Preservation Strategy: Review Date Field		
189	Validation Check Status - Date & Time checked - Validation Result (success, failure w/error msg or error number) - Remediation performed if failed This field provides the Fixity Log for the resource. Note: there may be many files within a single resource that have to have their fixity value checked.		
190			
191			
192			
193			
194			
	Packaging Information		
195			
196			
197			
198			
199			
200			
	Descriptive Information **Note:** There are duplicates in some of the following fields, as a metadata field may be used for archival management as well as access.		
201	Title	<title>	
202	Originator / Record Creator / Author <duplic 98>	<origin>	
203	Contributor(s) <duplic 103>	<proccont>	
204	Publisher Name <duplic 104>		
205	Distributor <duplic 107>	<distrib>	
206	Name of File Type (also known as Format) <duplic 10>		
206	Object (File) Size <duplic 3>	<transize>	
207	Date Record Created <duplic 102 >		
208	Date Record Published <duplic 106>	<pubdate>	
209	Time Period of Data	<timeperd> <begdate> <enddate>	
210	Date Access Copy Created		
211	Language <duplic 5>		
212	Finding and Searching Aids		

	Metadata Field	FGDC	Required Y/N
213	Theme / Subject Keywords		
214	Place Keywords GIS records are usually associated with a place, and so should be recorded for access	<placekey>]	
215			
216			
217			
218			
219			
220			
221			
222			
223			
224			
225			
226			
	Descriptive Info for Geospatial Datasets		
227	Geospatial data presentation form <duplic 34> Controlled vocabulary Per FGDC: Anglo-American Committee on Cataloging of Cartographic Materials, 1982: interpretation of AACR2. ALA *atlas, audio, diagram, document, globe, map, model, multimedia presentation, profile,* ***raster digital data****,* ***remote-sensing image****, section, spreadsheet, tabular digital data,* ***vector digital data,*** *video, view, free text.*	<geoform>	
228	Stratum keywords	<stratumkey>	
229	West Bounding Coordinate <duplic 35>	<westbc>	
230	East Bounding Coordinate <duplic 36>	<eastbc>	
231	North Bounding Coordinate <duplic 37>	<northbc>	
232	South Bounding Coordinate <duplic 38>	<southbc>	
233	Point of Contact from metadata file	<ptcontac>	
234	Raster Spatial Resolution	<abstract>	
235	Attribute Label	<attrlabl>	
236	Attribute Definition	<attrdef>	
237	Geospatial Metadata File - link or pathname to human-readable manifestation of the geospatial metadata file - XML - styled HTML		
238			
239			
240			
241			
242			

Appendix B: FGDC Metadata - summary list

Contact Identification Information included in geospatial metadata file:

FGDC standard includes several contacts throughout the metadata record. Some of these reference the originators of the geospatial dataset, while others reference the processors, and others reference the distributor. Archives will need to consider which of these contacts you will want to carry through into your archival metadata record, and which of these you may want to present in your archived geospatial data access interface. Additionally, as the geospatial dataset is being transferred to the archives, the archives may want to consider creating an updated version of the geospatial dataset that reflects the archival organization as the Distributor for this dataset. This will not be as much of an issue for geospatial datasets distributed directly from the archives, where end users know that they are retrieving the dataset from the archival organization, but may be relevant if the archives chooses to publish the availability of their archived datasets through one of the inventory systems such as RAMONA GIS Inventory (http://gisinventory.net/) or Geospatial Onestop (http://www.geodata.gov/).

1. Citation::Originator

FGDC Defn: the name of an organization or individual that developed the dataset. May be multiple creators (p 54).[10]

2. Citation::Publisher / Publication Place

FGDC Defn: the name of the individual or organization that published the dataset.

Note: Originator and Publisher do not necessarily have to be the same, but often are.

3. Point of Contact

FGDC defn: contact information for an individual or organization that is knowledgeable about the dataset (p. 8).[9]

FGDC FAQ: What's the difference between "originator" in compound element "Citation" (1.1) and the "Point of Contact" element?
The "Originator" is the person(s) and organization(s) that developed the dataset.
The "Point of Contact" is the person(s) and organization(s) that can be contacted if questions arise about the dataset.
The "Originator" and the "Point of Contact" may be the same.
Thought Question for Archives: Should the archives publish the original Point of Contact in their data access record? Store the original Point of Contact in the archival metadata record?
Or does the archives become the "Point of Contact" as a dataset is accessioned (as **ownership of the dataset is transferred**)?
Will the archives field technical questions about the dataset?

4. Process Step Process Contact

FGDC defn: the party responsible for the processing step information (p. 15).[9]

5. Distributor Information

FGDC defn: Information about the distributor and options for obtaining the dataset (p. 42).[9]

Note: With the transfer of ownership of the dataset to the archives for long-term preservation, an archival organization may want to consider creating a new version of the geospatial metadata file that designates the archives as the Distributor contact, and includes the archives' distribution liability statement

!! RECOMMENDATION !!

If you are going to modify the originally received metadata file, be sure to copy and save the original as the original historical record.

The following table also indicates which FGDC fields were identified as important for the geospatial professionals to populate to support the long-term preservation of geospatial datasets. As these metadata values are provided in the geospatial metadata record, it is not necessarily required that they be replicated in the geospatial **archival** metadata record. Only the subset of geospatial metadata fields that have been identified as useful in supporting either the archival management or end user access have been identified as needing to be extracted from the geospatial metadata record and loaded into the archival metadata record. Regardless of their anticipated archival management / access use, archival organizations may alternatively

[10] Federal Geographic Data Committee. "Content Standard for Digital Geospatial Metadata FGDC-STD-001-1998." 1998.

choose to load the entire geospatial metadata record into the archival metadata record. The impact of importing all geospatial metadata into the archival metadata repository will be an increased archival database size.

	Included in above Geospatial Archival Metadata table	MT GIS Standard P=Portal access R=Rqd Fields I=Important Fields	GeoMAPP GIS Preservation Whitepaper**
Global Unique ID		P	
1. IDENTIFICATION INFORMATION			
CITATION			
Originators	Y	R	X
Title	Y	P date+region+subj	X
Geospatial data presentation form (e.g. vector digital data)	Y		X
Publisher	Y	R	X
Publication Date	Y	R	X
Publication Place			X
Other citation details will likely be superseded by Archives' citation statement, so original citation not needed in the archival metadata record. However, the original citation statement is retained in the originally received geo metadata file.			
Online Linkage		P	
DESCRIPTION			
Abstract	Y	P	X
Purpose	Y	I	X
TIME PERIOD OF CONTENT	Y	P	X
Beginning Date	Y		X
Ending Date	Y		X
Currentness Reference			
STATUS			
Progress		I	
Update Frequency		I	X
SPATIAL DOMAIN			
West Bounding Coordinate	Y	P (adv search)	X
East Bounding Coordinate	Y	P (adv search)	X
North Bounding Coordinate	Y	P (adv search)	X
South Bounding Coordinate	Y	P (adv search)	X
KEYWORDS			
Theme Keyword Thesaurus		P ISO 19115 Topic Category	X ISO 19115 Topic Category
Theme Keywords	Y	P cntrld vocab=layer name	X cntrld vocab=layer name
Place Keyword Thesaurus		P	X
Place Keywords	Y	P	X
Stratum Keyword Thesaurus			
Stratum Keywords	Y		
Temporal Keyword Thesaurus			
Temporal Keywords			
ACCESS CONSTRAINTS		I	X

	Included in above Geospatial Archival Metadata table	MT GIS Standard P=Portal access R=Rqd Fields I=Important Fields	GeoMAPP GIS Preservation Whitepaper**
will likely be superseded by Archives' policy, so original constraint not likely needed in the archival metadata record. **However, the original constraint statement should be retained in the originally received geo metadata file.**			
USE CONSTRAINTS will likely be superseded by the Archives' policy, so original constraint may not be needed in the archival metadata record. **However, the original constraint statement should be retained in the originally received geo metadata file.**	?	I	X
POINT OF CONTACT ??supersede original dataset value with Archives' contact info as the "Point of Contact" for the dataset? QUESTION: Are there consistent/commonly known GIS Professionals who are identified as the Point of Contact across multiple datasets? Might somebody be interested in searching for all of the datasets that "Jim Mapmaker" is the contact point for?? If so -- then Yes we should include this in the archival metadata record.	?		
Contact Person		P (search field)	X
Contact Organization		P	X
Contact Address		P	
Contact Voice Telephone		P	X
Contact Facsimile			
Contact email		P	X
Hours of Service			
Contact Instructions			
BROWSE GRAPHIC (repeatable)			
DATA SET CREDIT			
SECURITY INFORMATION	if applicable		
NATIVE DATA SET ENVIRONMENT e.g. Microsoft Windows XP Version 5.1 (Build 2600) Service Pack 3; ESRI ArcCatalog 9.3.1.3000	Y		X
Native Dataset Format* (controlled vocab) e.g. Personal GeoDatabase Feature Dataset, Layer, Map, Raster Dataset, Shapefile, etc.	Y		
2. DATA QUALITY INFORMATION			
ATTRIBUTE ACCURACY report		I	
LOGICAL CONSISTENCY REPORT			
COMPLETENESS REPORT		I	
POSITIONAL ACCURACY			
Horizontal Positional Accuracy		I	
Vertical Positional Accuracy		I	
LINEAGE			
Source Citation Information (Title, originator, pub	Y	I	X

	Included in above Geospatial Archival Metadata table	MT GIS Standard P=Portal access R=Rqd Fields I=Important Fields	GeoMAPP GIS Preservation Whitepaper**
date, pub info)			
Source scale denominator			
Type of source media			
Source citation abbreviation			
Source contribution			X
Source time period of content			
Source currentness			X
Process Step (repeatable)		I	
Process Description To maintain the ongoing lineage in the geospatial metadata file, the archives may consider adding an additional processing step to the geospatial metadata record to record the archival ingest processing.	Y archivist's processing steps should be recorded in archival record ?add to Geospatial Metadata File?		X
Process Date The archival record should need only the record of the archivists' actions -- to establish the record of the archives' actions on the record.	Y archivist's processing steps will be recorded in archival record ?add to Geospatial Metadata File?		X
Process Contact (name, addr, ...) Probably local policy if the archives wants to record the name of the processing archivist in the archival record. Provide institution name, address, and institution phone/email contact points.	Y at least institution information if not individual contact information		X
3. SPATIAL DATA ORGANIZATION INFORMATION *varies based on dataset type*			
INDIRECT SPATIAL REFERENCE	Y		
DIRECT SPATIAL REF METHOD (controlled vocabulary) e.g. Point, Vector, Raster	Y		
POINT & VECTOR OBJECT INFO	Y		
OR			
RASTER OBJECT INFO	Y		
4. SPATIAL REFERENCE INFORMATION *varies based on type of coord systm*			
HORIZONTAL COORD SYSTEM DEF	Y	R	
Geographic	Y		
OR			
Planar	Y		
Map Projection Name	Y		X

	Included in above Geospatial Archival Metadata table	MT GIS Standard P=Portal access R=Rqd Fields I=Important Fields	GeoMAPP GIS Preservation Whitepaper**
OR			
Grid Coordinate System Name	Y		X
Planar Coordinate Information	Y		X
Planar Distance Units	Y		X
Geodetic Model			
Horizontal datum name	Y		X
Ellipsoid name	Y		X
Semi-major axis	Y		X
Denominator of flattening ratio	Y		X
VERTICAL COORD SYSTEM DEF		R	
Altitude System Definition	Y		
Altitude Datum Name	Y		
Altitude Resolution	Y		
Altitude Distance Units	Y		
Altitude Encoding Method	Y		
Depth System Definition	Y		
Depth Datum Name	Y		
Depth Resolution	Y		
Depth Distance Units	Y		
Depth Encoding Method	Y		
5. ENTITY & ATTRIBUTE INFORMTION			
DETAILED DESCRIPTION (repeatable)		R	
Entity type		R	
Entity label			X
Attribute (repeatable)			
Attribute label	Y	R	X
Attribute definition	Y	R	X
Attribute definition source			
OVERVIEW DESCRIPTION (repeat)			
Entity and Attribute Overview			
Entity and Attribute Detail Citation			
6. DISTRIBUTION INFORMATION			
DISTRIBUTOR		R	
Contact information (name, addr,...) Question: The archives essentially becomes the distributor for the "archived" version of the dataset. Will the archives publish an archives version of the geospatial metadata (e.g. to the national geospatial inventory systems) that reflects the archives ownership. The Archives distributor contact information would then supersede the original distributor information in the metadata file. **However, the original Distributor information should be retained in the originally received geospatial metadata file.**	Maintain original Distributor in archival record? or replace with Archival Contact info?	R	
RESOURCE DESCRIPTION		P (cntrld vocab) Adv Search	

	Included in above Geospatial Archival Metadata table	MT GIS Standard P=Portal access R=Rqd Fields I=Important Fields	GeoMAPP GIS Preservation Whitepaper**
DISTRIBUTION LIABILITY will likely supersede original distribution liability statement with Archives' policy, so the original Distribution Liability statement would not be required in or transferred to the archival meta data record. However, the original statement is retained in the originally received geo metadata file.	Record original Distribution Liability Statement in archival record? or will this be the archives' liability statement?		
STANDARD ORDER PROCESS will likely supersede with Archives' policy. However, the original statement is retained in the originally received geo metadata file.			
Non-digital form			
Digital	Y		
Digital Transfer Format Name (controlled vocab: see values at: http://www.fgdc.gov/csdgmgraphical/distr/storder/digform/traninf/fname.htm)			
Digital Transfer Info: Transfer size Archives's distribution package may be different than the original dataset that was received - so original metadata transfer size would no longer be relevant.			
Fees			
Ordering Instructions			
Turnaround			
CUSTOM ORDER PROCESS will likely supersede original custom order process with Archives' process, so the original Custom Order Process statement would not be required in or transferred to the archival meta data record. However, the original statement is retained in the originally received geo metadata file.			
TECHNICAL PREREQUISITES	Y		
AVAILABLE TIME PERIOD			
7. METADATA REFERNC INFORMTION			
METADATA DATE will likely supersede when the creates an updated version of the geo metadata file reflecting the Archives processing steps, distribution information, distribution liability statement, constraints, etc. However, for historical searching purposes, would end users may want to search on the metadata update date for the original dataset??	Record original Metadata Date in archival record? or will this be the date the archives updated the metadata record?	R	
METADATA REVIEW DATE			
METADATA FUTURE REVIEW DATE			
METADATA CONTACT (name ,addr,..) will likely supersede original Metadata Contact with Archives' processing archivist contact information.	Record original Metadata Contact in archival	R	

	Included in above Geospatial Archival Metadata table	MT GIS Standard P=Portal access R=Rqd Fields I=Important Fields	GeoMAPP GIS Preservation Whitepaper**
However, the original metadata contact information is retained in the originally received geo metadata file.	record? or will this be the date the archives updated the metadata record?		
METADATA STANDARD NAME important for archival record to know the format the metadata is formed in.	Y		X
METADATA STANDARD VERSION likewise, useful to know which version of the metadata standard was used to form the archival metadata record.	Y		X
METADATA TIME CONVENTION			
METADATA ACCESS CONSTRAINTS			
METADATA USE CONSTRAINTS			
METADATA SECURITY INFORMTION			
METADATA EXTENSIONS (repeat) If metadata file includes extensions (e.g. Esri additional fields), that likewise should be stored in the archival metadata record to provide the full picture of the geospatial metadata structure.			
Online linkage http://www.esri.com/metadata/esriprof80.html			
Profile name e.g. ESRI Metadata Profile			
Language of metadata*			

*Esri-specific metadata field

** The GeoMAPP whitepaper, "Utilizing Geospatial Metadata to Support Data Preservation Practices" (http://www.geomapp.net/docs/GeoMetadata_Items_for_Preservation_2011_0110.pdf) identifies the FGDC geospatial metadata fields that are important in facilitating the long-term preservation of geospatial data. As the metadata is either embedded within the geospatial dataset, or exported to an external geospatial metadata (xml) file, it is not necessarily required to load all of the geospatial metadata into the archival metadata record. Alternatively, you may consider loading just the subset of geospatial metadata fields into the archival metadata record that support either the ongoing management and/or access of the geospatial dataset. For example, the original record owner's Custom Order Process, while possibly interesting from an historical information perspective, will not necessarily facilitate the management or access to the geospatial record. However, this information should be available to end users/researchers through the availability of the original metadata file, so would not necessarily need to be included in the archival metadata record.

Appendix C: References

The following resources were reviewed and informed the development of this metadata data model:

Articles:

Caplan, Priscilla. (2006). DCC Digital Installment on "Preservation Metadata". *Curation Reference Manual.* Document Link: http://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/preservation-metadata/preservation-metadata.pdf

Lavoie, Brian. (2005). Technology Watch Report: Preservation Metadata. http://www.dpconline.org/index.php?option=com_docman&task=doc_download&gid=88

Development Initiatives:

CEDARS (CURLS Exemplars in Digital Archives) – nice description of OAIS model, developed preservation metadata data model based on OAIS: http://www.ukoln.ac.uk/metadata/cedars/papers/aiw02/

Center for International Earth Science Information Network (CIESIN) Geospatial Electronic Records (GER): http://www.ciesin.org/ger/

Guide to Managing Geospatial Electronic Records v1.00. Center for International Earth Science Information Network (CIESIN), Columbia University. (June 2005). http://www.ciesin.org/ger/GuideToManagingGERv1Final.pdf (accessed 12/20/2010)
GER Data Model: http://www.ciesin.org/ger/DataModelV1_20050620.pdf (see page 10 for a graphic depiction)

Digital Curation Center's Curation Reference Manual:
DCC Curation Reference Manual: http://www.dcc.ac.uk/resources/curation-reference-manual

Stanford

National Geospatial Digital Archive Project: http://www.digitalpreservation.gov/partners/ngda/ngda.html
Paper: Hoebelheinrich, Nancy. An Investigation into Metadata for Long-lived Geospatial Data Formats (March 2008): http://www.digitalpreservation.gov/news/events/ndiipp_meetings/ndiipp08/docs/session7_hoebelheinrich_paper.doc
Hoebelheinrich, Nancy. An Investigation into Metadata for Long-Lived Geospatial Data Formats. (NDIIP – Stanford University). (March 2008) http://www.ngda.org/reports/InvestigateGeoDataFinal_v2.pdf (accessed 12/20/2010)

New Zealand: Knight, Steve. (2005). Preservation Metadata: National Library of New Zealand Experience. Library Trends. 54(1), pp91-110.
(includes diagram of their metadata data model)
New Zealand Develops New Tools, Policies for Digital Preservation: http://www.digitalpreservation.gov/edge/edge_nlnz.html

National Library of Australia – Pandora
Preservation Metadata for Digital Collections. National Library of Australia. (Oct. 1999) http://www.nla.gov.au/preserve/pmeta.html (accessed 12/20/2010)
(nice example of technology-specific metadata definitions)

Standards:

OAIS: Consultative Committee for Space Data Systems. *Reference Model for an Open Archival Information System (OAIS); CCSDS 650.0-B-1 Blue Book.* January 2002.
http://public.ccsds.org/publications/archive/650x0b1.pdf

PREMIS: http://www.loc.gov/standards/premis/
PREMIS Data Dictionary for Preservation Metadata version 2.1:
http://www.loc.gov/standards/premis/v2/premis-2-1.pdf

Ramona Data Categories and Layers: http://www.sco.wisc.edu/wisclinc/Layer_Assignments_060206.pdf