

Geospatial Multistate Archive and Preservation Partnership

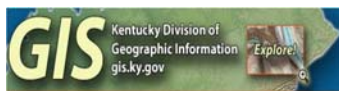
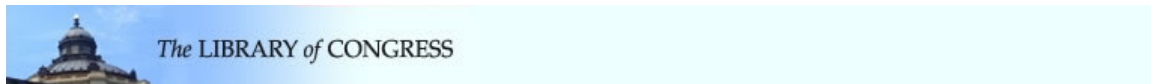
FGDC and Dublin Core Metadata Comparison

In cooperation with the Library of Congress National
Digital Information Infrastructure and Preservation
Program

and

NC Center for Geographic Information & Analysis and North Carolina State
Archives, KY Division of Geographic Information and Department for
Libraries and Archives, NCSU Libraries, and UT Automated Geographic
Reference Center and State Archives

March 2009



The Geospatial Multistate Archive and Preservation Partnership (GeoMAPP) has been tasked with comparing differing types of metadata standards for the purpose of data preservation and discovery. Metadata is data about the data. A metadata standard is language presented using controlled format.¹ The creation of metadata is essential to the quality of the data set.² In order for data to be preserved and discovered, certain criteria need to be met within the metadata:

- Who: creator of data,
- What: title and description of data,
- Where: geographical extent of data,
- Why: reason the data was created,
- When: when the data was created,
- How: how the data was created.³

This paper will compare FGDC and Dublin Core, two metadata formats. Although a third metadata standard, PREMIS, was looked at as a “wrapper” for the data, it was deemed too labor intensive and cost prohibitive at this stage of the GeoMAPP project.

FGDC Metadata Standard

The FGDC or Federal Geographic Data Committee promotes sharing of the nation’s geospatial resources.⁴ The FGDC metadata standard was adopted in 1994. The

¹. Geospatial Metadata Standards. <http://www.fgdc.gov/metadata/geospatial-metadata-standards>.

². Institutionalize Metadata before it Institutionalizes you, http://www.fgdc.gov/metadata/documents/InstitutionalizeMeta_Nov2005.doc.

³. <http://www.gsdi.org/pubs/cookbook/chapter03a.html>.

⁴. Federal Geographic Data Committee. <http://www.fgdc.gov/>.

FGDC is a metadata standard developed to determine the robustness, the method of accessing, and successful transfer of geospatial data.⁵

At this time the FGDC metadata standard does not have a field for the technical and administrative metadata elements needed for archival purposes. Examples of information not supported by this standard include the following: date the data was acquired by the archive, rights of the archive in regards to the data, technical information about the transfer and any future transfers to take place, and the process for guaranteeing the data's integrity long term.⁶

The FGDC metadata standard consists of seven mandatory elements. Each of these seven elements has additional information that can be attached to it.⁷

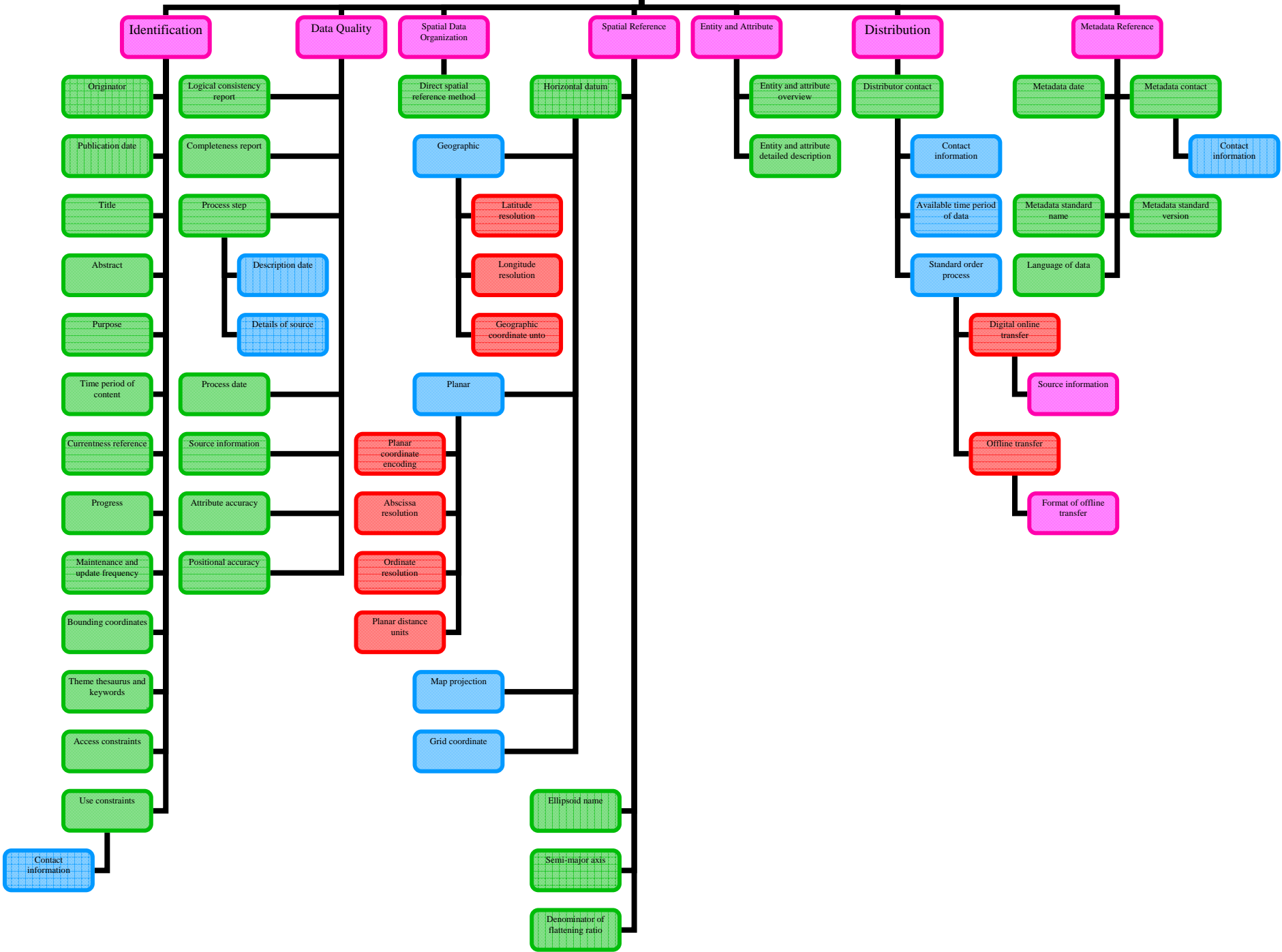
Here is a table showing these elements and their additional attributes:

⁵. Hoebelheinrich, Nancy and John Banning, "An Investigation into Archiving Geospatial data Formats", Stanford University Libraries, March 2008.

⁶. http://www.digitalpreservation.gov/partners/ncgdap/high/NCGDAP_InterimReport_June2008_final.pdf.

⁷. Essential Metadata Elements, http://www.fgdc.gov/metadata/documents/CSDGMEssentialMeta_20080514.pdf.

FGDC Metadata Elements



FGDC allows a user to create templates. These templates allow the organization to pick and choose the elements that stay consistent. “Templates are built by extracting those metadata fields pertinent to the organization and the specific data types and geographies of the organization.”⁸

Dublin Core Metadata Initiative

“The Dublin Core Metadata Initiative (DCMI) is an open organization engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models.”⁹ DC provides a simple and standardized set of rules for describing electronic resources online in ways that make them easier to find.¹⁰ Dublin Core uses Resource Description Framework (RDF) as its markup language. RDF is a language for representing information about resources in the World Wide Web.¹¹ Dublin Core Metadata consists of fifteen simple elements.

Contributor:

An entity responsible for making contributions to the resource. This may include a person or an organization.

Coverage:

This is the spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.

The spatial topic and spatial applicability may be a named place or a location specified by its geographic coordinates. Temporal topic may be a named period, date, or date range. A jurisdiction may be a named administrative entity or a geographic place to which the

⁸. Institutionalize Metadata before it Institutionalizes you, http://www.fgdc.gov/metadata/documents/InstitutionalizeMeta_Nov2005.doc.

⁹. <http://dublincore.org/>.

¹⁰. Wikipedia, http://en.wikipedia.org/wiki/Dublin_Core.

¹¹. RDF Primer, <http://www.w3.org/TR/rdf-primer/>.

resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names (TGN). Where appropriate named places or time periods can be used in preference to numeric identifiers such as sets of coordinates or date ranges. For example: coverage type = spatial, scheme = LATLONG = West = 180, East = 180, North = 90, South = 90.¹²

Creator:

The entity primarily responsible for making the resource.

Examples of a creator include a person, an organization, or a service. Typically, the name of a creator should be used to indicate the entity.

Date:

A point or period of time associated with an event in the life cycle of the resource.

Date may be used to express temporal information at any level of granularity.

Description:

The explanation of the resource.

Description may include: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.

Format:

The file format, physical medium, or dimensions of the resource.

Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types (MIME).

¹². OCLC/NCSA Metadata Workshop, <http://www.ifla.org/documents/libraries/cataloging/oclcmeta.htm>.

Identifier:

An unambiguous reference to the resource within a given context.

Recommended best practice is to identify the resource by means of a string conforming to a formal identification system.

Language:

Language of the resource.

Publisher:

An entity responsible for making the resource available.

Relation:

A related resource.

Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.

Rights:

Information about rights held in and over the resource.

Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.

Source:

A related resource from which the described resource is derived.

The described resource may be derived from the related resource in whole or in part.

Subject:

The topic of the resource.

Typically, the subject will be represented using keywords, key phrases, or classification codes.

Title:

A name given to the resource.

Typically, a title will be a name by which the resource is formally known.

Type:

The nature or genre of the resource.

Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary.¹³

Because the elements can be used more than once, there may be a need to differentiate between the multiple values in each element.¹⁴ This is the reason qualifiers were created. There are two classes of qualifiers; the element refinement and the encoding scheme. Element refinement is used to narrow an element down to a more specific value. Encoding scheme is used to interpret a more specific value.¹⁵ For example:

<u>Element</u>	<u>Element refinement</u>	<u>Encoding scheme</u>
Coverage	Spatial	Lat/Long Coordinates

¹³. Dublin Core Metadata, <http://www.dublincore.org/documents/dces/>.

¹⁴. Dublin Core Qualifiers, <http://dimes.lins.fju.edu.tw/dimes/meta-ref/DC-SubElements.html>.

¹⁵. Using Dublin Core, <http://drott.cis.drexel.edu/1664/Dublinusageguide.html#qualifiers>.

Crosswalk¹⁶

Categories	FGDC	Dublin Core	Qualifiers
Mandatory	Title	Conditional Title	main, short, alternative, subtitle, part title, spine, translated creation, current, modified, valid to and from
Mandatory	Publication date	Conditional Date	
Mandatory	Language	Conditional Language	
Mandatory	Subject	Conditional Subject	
Mandatory	Abstract	Conditional Description	abstract, freetext contact information freetext, abstract temporal freetext, abstract freetext, abstract
Mandatory	Originator	Conditional Creator	
Mandatory	Purpose	Conditional Description	
Conditional	Time period of content	Conditional Coverage	
Conditional	Currentness reference	Conditional Description	
Conditional	Progress	Conditional Description	
Conditional	Maintenance and update frequency	Conditional	
Conditional	Bounding coordinates	Conditional Coverage	spatial controlled vocabularies, LCSH, etc.
Conditional	Theme thesaurus and keywords	Conditional Subject	
Mandatory	Access constraints	Conditional Rights	
Mandatory	Use constraints	Conditional Rights	
Conditional	Publisher	Conditional Publisher	contact information primary, copy relationship of
		Conditional Identifier	
		Conditional Relation	
		Conditional Source	
		Contributor	contact information, role of
		Format	extent, medium
		Type	

¹⁶ <http://www.coris.noaa.gov/backmatter/examples/MetadataCrosswalk.pdf>.

Conclusion

The two metadata formats compared in this paper, FGDC and Dublin Core, capture information for GIS metadata. However, Dublin Core was created to be a broad metadata capture tool and does not necessarily capture all of the fields needed to complete a GIS metadata record. Dublin Core works well for data discovery, whereas FGDC contains the more in-depth information for research¹⁷ and was created specifically for GIS. It is more robust in capturing the specifics needed for a complete metadata record. Because of the flexibility of Dublin Core it can be used in conjunction with the FGDC to supplement those fields needed for archival purposes.¹⁸ Below is a simplified model listing the optimal metadata of the FGDC Standard and Dublin Core combined.

¹⁷ . <http://www.gisnetwork.org/Documents/GISINProc2004HTML/GISINProc2004213.html>.

¹⁸ . http://www.digitalpreservation.gov/partners/ncgdap/high/NCGDAP_InterimReport_June2008_final.pdf.

Optimal Metadata
FGDC Standard and Dublin Core Combined

