National Spatial Data Infrastructure

Federal Geographic Data Committee Educational Resource:

Preparing for International Metadata

Version 2
October 20, 2011

Federal Geographic Data Committee
Department of Agriculture • Department of Commerce • Department of Defense • Department of Energy
Department of Health & Human Services • Department of Housing and Urban Development • Department of the Interior
Department of Justice • Department of State • Department of Transportation • Environmental Protection Agency
Federal Emergency Management Agency • Library of Congress • National Aeronautics and Space Administration
National Archives and Records Administration • National Science Foundation • Tennessee Valley Authority

Federal Geographic Data Committee

Established by Office of Management and Budget Circular A-16, the Federal Geographic Data Committee (FGDC) promotes the coordinated development, use, sharing, and dissemination of geographic data.

The FGDC is composed of representatives from the Departments of Agriculture, Commerce, Defense, Energy, Health & Human Services, Housing and Urban Development, the Interior, Justice, State, and Transportation; the Environmental Protection Agency; the Federal Emergency Management Agency; the Library of Congress; the National Aeronautics and Space Administration; the National Archives and Records Administration; National Science Foundation, and the Tennessee Valley Authority. Additional Federal agencies participate on FGDC subcommittees and working groups. The Department of the Interior chairs the committee.

FGDC subcommittees work on issues related to data categories coordinated under the circular. Subcommittees establish and implement standards for data content, quality, and transfer; encourage the exchange of information and the transfer of data; and organize the collection of geographic data to reduce duplication of effort. Working groups are established for issues that transcend data categories.

The following is the recommended bibliographic citation for this publication:
Federal Geographic Data Committee, 2011 Preparing for International Metadata, Federal Geographic Data Committee, Washington, D.C.

For more information about the committee, or to be added to the committee's newsletter mailing list, please contact:

Federal Geographic Data Committee Secretariat
c/o U.S. Geological Survey
590 National Center
Reston, Virginia 20192
Telephone: (703) 648-5514
Facsimile: (703) 648-5755
Email: fgdc@fgdc.gov
WWW Home Page: http://www.fgdc.gov

Executive Summary

The Federal Geographic Data Committee has endorsed several new geospatial metadata standards that support the International Organization for Standardization *ISO 19115: Geographic information – Metadata* standard. Most significant among the endorsed standards are:

- *ISO 19115:2003* – the base ISO geospatial metadata standard; a UML representation that specifies the content, conditionality and interrelations of the data documentation elements
- *ISO 19139:2007* – an XML implementation of ISO 19115 that specifies metadata record format
- *ISO 19115-2:2009 Geographic information – Metadata – Part 2: Extensions for imagery and gridded data* – an extension of the base standard that adds element for describing imagery and gridded data
- *North American Profile (NAP) of ISO 19115: Geographic information – Metadata* – a U.S. and Canada specified implementation of the base standard that increases some conditionality, extends some domains and specifies best practices for populating the metadata record.

A revised version of the base standard, *ISO 19115-1: Geographic information – Metadata – Part 1: Fundamentals,* is expected to reach final draft status in 2012. The affiliate standards, *ISO 19139, ISO 19115-2* and *NAP,* are expected to be updated to reflect the changes implemented in *ISO 19115-1*. In addition, the U.S. community has requested that the NAP more fully integrate the geospatial database documentation specified by another affiliate standard, *ISO 19110: Geographic information: Feature Catalogue*.

In the same way that the existing FGDC *Content Standard for Digital Geospatial Metadata* (CSDGM) codified geospatial data documentation for the U.S. geospatial data community in 1994, ISO geospatial metadata extends standardization across national borders. Key features include:

- fewer mandatory elements and more optional elements
- extended elements and new elements to capture more specific information
- a hierarchical structure that creates ‘packages’ of metadata that can be reused and combined to form new metadata records
- support for the documentation of new geospatial data topologies and technologies including geodatabases, web mapping applications, data models, data portals, ontologies, etc.
- suggested best practices for populating metadata elements in a manner that enhances the quality and usefulness of the metadata (NAP feature).

All levels of government, non-governmental organizations (NGOs), and the private sector are strongly encouraged to prepare for the transition to this new, international, approach to metadata by:

- editing existing metadata records to add new content and convert some ‘free text’ elements to a ISO designated code list using either a manual or automated process,
- testing available transform tools for converting records,
- informing management and technical staff of pending changes and
- planning a strategy for transition to international metadata.

This document provides an overview of ISO 19115 and NAP geospatial metadata and specific guidance on preparing for the transition. Additional guidance documents, similar in scope to the CSDGM Workbook, are currently under development and will include detailed explanations of the record structure, individual elements and best practices. GIS vendors and Federal Agencies are actively developing new applications to transform, create, validate, publish and distribute ISO 19115 and NAP metadata. With these resources in hand, geospatial data and service providers can be fully prepared to update and enhance their geospatial metadata to better support data management, discovery, distribution, application and archive both within, and external to, their organization.

Table of Contents

U.S. and ISO Geospatial Metadata Standard Development

1994	CSDGM, Vers. 1.0
1998	CSDGM, Vers. 2.0
2003	ISO 19115, Vers. 1.0
2007	ISO 19139, Vers. 1.0
2009	ISO 19115-2, Vers. 1.0
	North American Profile of 19115, Vers. 1.0
2012	ISO 19115-1, *draft*

U.S. and ISO Geospatial Metadata Development

The Federal Geographic Data Committee (FGDC) approved the first version of the Content Standard for Digital Geospatial Metadata (CSDGM) in June 1994. Executive Order 12906, "Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure," required Federal agencies to use the standard to document data that they produced beginning in 1995. In 1998, the CSDGM was revised and Version 2.0 was published.

Since it was introduced, the CSDGM has been adopted by many organizations external to the U.S. Federal Government including state and local governments, private sector companies and other nations including Canada. The CSDGM is a valuable resource for documenting geospatial data in a standardized manner that facilitates discovery, access, use, and archive. In addition, the standard has been used to document web-mapping applications, data-acquisition projects, data models, and other 'non-traditional' geospatial resources. As these new forms of geospatial resources evolve, new requirements for documentation emerge.

In 1999 the International Standards Organization (ISO) Technical Committee (TC) 211 Geographic Information / Geomatics was tasked with harmonizing the CSDGM with other geospatial metadata standards and a range of de facto standards that had emerged to address new requirements for geospatial documentation. The result was the publication of *ISO 19115: Geographic information – Metadata*. Since then, individual organizations and nations have developed implementation 'profiles' of the standard. The American National Standards Institute (ANSI), the U.S. member body of the ISO, adopted ISO 19115 in December of 2003.

ISO 19115 was developed using Unified Modeling Language (UML). The ISO 19115 model specifies the definition of elements, conditionality, and the relationship among elements but does not provide guidance as to how the content is organized into a formal record and presented to the reader. Without

this critical component, the standard could not be promoted for implementation. A separate ISO TC 211 effort, *ISO 19139 - Geographic information -- Metadata -- XML schema implementation*, was undertaken to provide an Extensible Mark-up Language (XML) implementation schema to specify the ISO 19115 metadata record format and to support the description, validation, and exchange of geospatial metadata.

Even with the XML schema under development, concerns remained as to the adequacy of ISO 19115 to meet the needs of the CSDGM community. One key concern was the lack of elements for the documentation of the geospatial database. Known as 'Entities and Attributes' within the CSDGM, TC 211 had designated the development of the database documentation elements to yet another separate standard, *ISO 19110 - Methodology for Feature Cataloging.*

In addition, members of the remote sensing community found ISO 19115 lacking with respect to the documentation of remotely sensed and gridded data. This community undertook the extension of ISO 19115 to include additional elements needed to capture information about satellite, aerial, monitoring station and other in-situ data. *ISO 19115-2:2009 Geographic information – Metadata – Part 2: Extensions for imagery and gridded data* was finalized in 2009 and includes all ISO 19915 elements and the extended in-situ elements.

It was not until the finalization of ISO 19110 in 2005 and ISO 19139 in 2007, that the resources were in place to develop a profile of ISO 19115 that provided both the structure and content needed to support migration of CSDGM metadata to the international suite of geospatial metadata standards. However, as work on the profile finalized in 2009, the ISO 19115 base standard began a five year review that resulted in recommendations for significant changes. The revised standard, *ISO 19115-1: Geographic information – Metadata – Part 1: Fundamentals*, is expected to reach final draft status in 2012. The revised standard expands role codes, simplifies paths, adds elements and is intended to improve the applicability and operability of the ISO 19115 standard.

Why Change Standards?
ISO 19115 development was initiated to:

- provide an internationally standardized means of documenting geospatial data resources
- incorporate international references including language and character set
- address new geospatial data structures and models
- include geospatial data applications and services.

As members of the international geospatial community, it is incumbent upon us to employ standards that enhance and support the discovery, access, archive, and application of geospatial data. In 1994, when the CSDGM was published, there was little interest in data sharing beyond the U.S. borders. Since then, the world has come to recognize that pressing environmental, political, health, and financial issues require international and, often, global perspectives. By developing a national profile of ISO 19115, the U.S. joins other nations in a global spatial data infrastructure (GSDI) that shares a common standard for geospatial data documentation and extends our capabilities to document a broader range of geospatial resources.

Coordination with the Canadian Government
In 2005, the U.S. and Canada began independent efforts to develop national profiles of ISO 19115. Since metadata creators and managers in both nations used the CSDGM, it was decided to align profile development efforts. Outreach was made to the Mexican geospatial community to join the cooperative development of a *North American Profile* (NAP) and, while receptive to the concept, the Mexican Government pursued other profile development efforts.

The U.S. and Canada continued with development of the NAP. Profile development requires adoption of 22 core elements, open selection of other elements, options to extend fixed domains and to increase, but not relax, conditionality. Once drafted the NAP was subject to a series of public reviews and revisions. An initial NAP document was approved by ANSI in July of 2009 but work continues on the integration of *ISO 19110 - Methodology for Feature Cataloging*.

Standards Bodies and Processing

OMB Circular A-119 directs Federal participation in voluntary consensus standards bodies including:

- American National Standards Institute (ANSI)
- International Standards Organization (ISO)
- InterNational Committee for Information Technology Standards; Technical Committee L1, Geographic Information Systems (INCITS_L1)
- Open Geospatial Consortium (OGC).

Additional information about standards bodies and standards implementations are maintained at: http://www.fgdc.gov/standards/organization/external-standards-organizations (*accessed October 14, 2011*).

Adoption of Federal and National Standards

When the CSDGM was created it was processed by the FGDC as a Federal standard. As such, only Federal agencies were obligated to create geospatial metadata using the CSDGM. However, as discussed above, many nonfederal organizations realized the benefit of standardizing their geospatial data documentation and chose to formally, and informally, adopt the CSDGM.

ISO 19115, ISO 19115-2 and the NAP were processed by ANSI. ANSI develops 'voluntary national consensus standards' that are promoted for use by the private sector as well as government and non-profit organizations. U.S. OMB Circular A119 [http://standards.gov/standards_gov/a119.cfm (*accessed October 14, 2011*)] states that Federal agencies are required to use voluntary standards for regulatory and procurement purposes when appropriate. In addition, many state and local governments also formally, and informally, adopt ANSI standards.

In September 2010, the FGDC formally endorsed the North American Profile of ISO 19115 and a suite of metadata related ISO standards. Most significant among the endorsed standards are:

- *ISO 19115:2003 Geographic information – Metadata* – the base ISO geospatial metadata standard; a UML representation that specifies the content, conditionality and interrelations of the data documentation elements
- *ISO 19139:2007* – an XML implementation of ISO 19115 that specifies the metadata record format
- *ISO 19115-2:2009 Geographic information – Metadata – Part 2: Extensions for imagery and gridded data* – an extension of the base standard that adds element for describing imagery and gridded data
- *ISO 19110:2005 - Geographic information - Methodology for Feature Cataloguing* – a metadata standard that defines a specific methodology for documenting feature types (entities and attributes)

Which Standard Should I Use?

The current array of geospatial metadata standards and variations of standards has left the community somewhat bewildered as to which geospatial metadata standard/variant they should be utilizing. At this time the FGDC recommends that organizations currently using the CSDGM metadata standard remain to do so *unless there is some compelling reason to change standards.*

The recommendation is based upon:

The continued evolution of ISO-related geospatial metadata standards:

- ISO 19115 five year review revisions and replacement as ISO 19115-1
- NAP revisions to conform to ISO 19115-1 and incorporate ISO 19110: Feature Catalog
- ISO 19115-2 revisions to conform to ISO 19115-1
- ISO 19139 revisions to conform to 19115-1

A lack of resources to adequately support ISO-related geospatial metadata implementation:

- Guidance and reference documents
- Training materials
- Data management and distribution portals
- Community expertise.

However, organization may choose to implement ISO geospatial metadata standards if the they:

- emphasize the documentation of remotely sensed or gridded imagery, including the documentation of monitoring stations and other in-situ data
- are an active participant in the development of data with another country/region that actively utilize~~d~~s ISO 19115
- are in the process of initiating a data management program and do not have significant holdings of metadata created using a different standard.

Objectives of the Standard

ISO 19115 was developed to achieve the following objectives.

- Support geographic data producers in their efforts to:
  - maintain an inventory of geographic data and services,
  - organize and manage geographic data and services, and
  - publish information internationally about available geographic data and services.

- Support geographic data consumers in their efforts to:
  - discover and access needed geographic data and services internationally,
  - assess the fitness for use of available geographic data and services, and
  - apply geographic data and services accessed from other organizations.

- Implement an internationally standardized set of geospatial metadata information and a common set of terminology and definitions for concepts related to metadata, including:
  - the names of data attributes (individual metadata elements) and classes (groups of data attributes) to be used,
  - the relations among attributes and classes,
  - the definitions of attributes and classes, and
  - information about the values that are to be provided for the data attributes.

- Implement an internationally standardized geospatial metadata record format.

- Promote an internationally standardized method for extending the standard to better address the specialized geographic data documentation needs of a specific community.

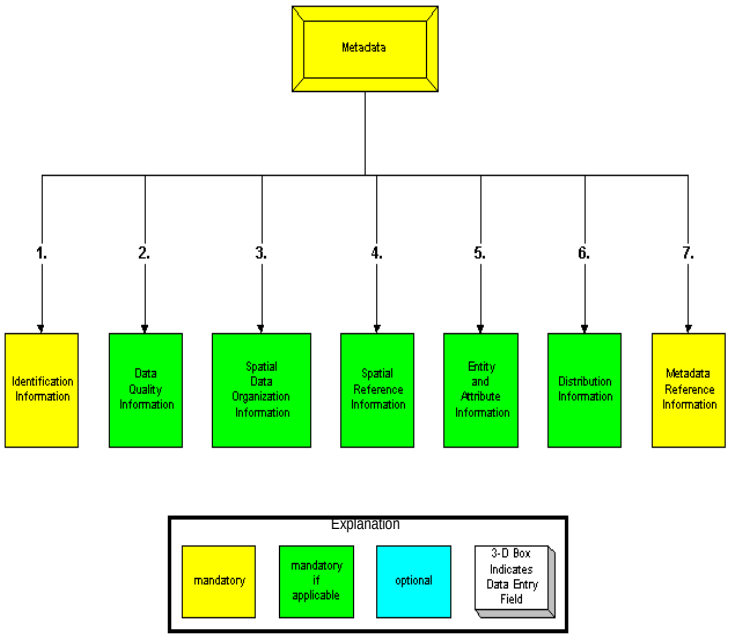
What's New with International Metadata?

Design Model
Organization
Best Practices
Multiplicity
Multi-level Metadata
New Content
Eliminated Elements

What's New with International Metadata?

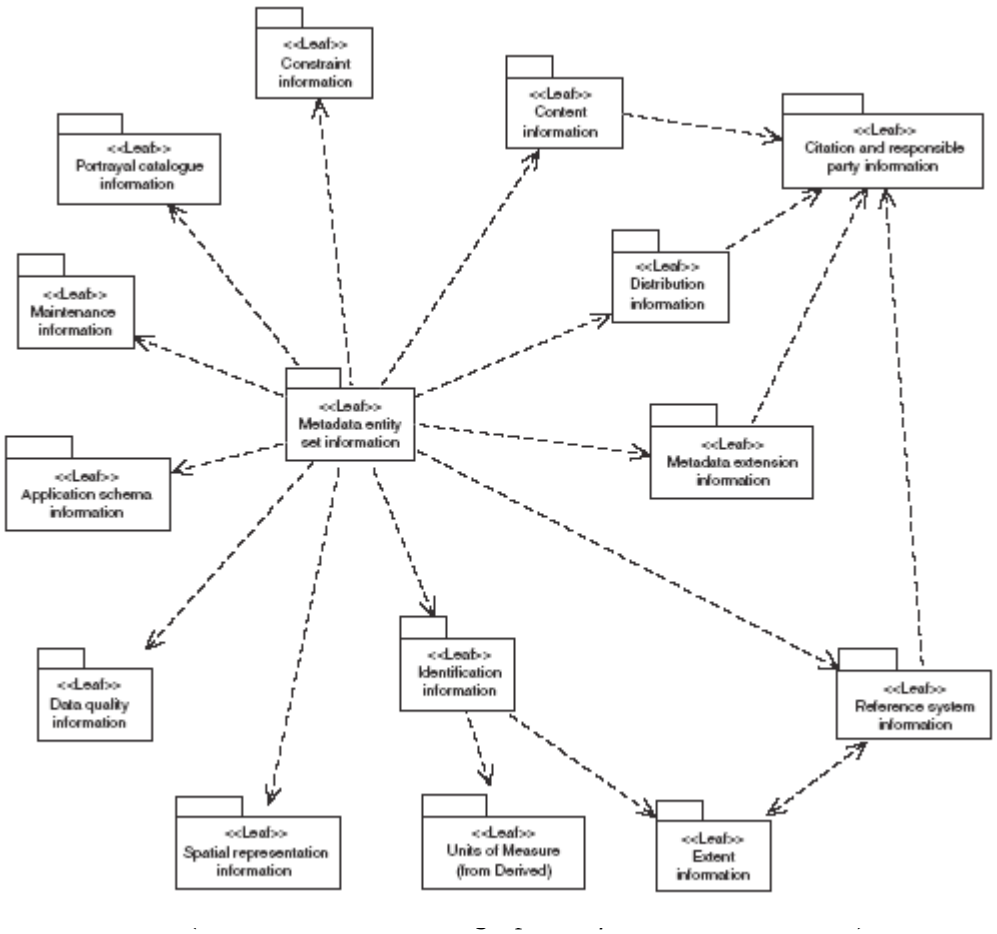
Design Model

The key differences between the CSDGM and the ISO 19115 design models are the manner in which they are arranged. The CSDGM design model is a plain document, also know as a 'flat file', presented as a structured list where all sections are equal as illustrated below.



CSDGM Graphical Representation
(Source: *Content Standard for Digital Geospatial Metadata Workbook*)

ISO 19115 was developed using Unified Modeling Language (UML). The UML data model provides a far more robust, object-oriented structure that helps to visualize more complex relations among the sections, the information contained with the sections, as well as information from related standards.



The UML design model is similar to a bubble diagram that might be used to sketch out the design of project tasks, a landscape plan, or a computer system. For example:

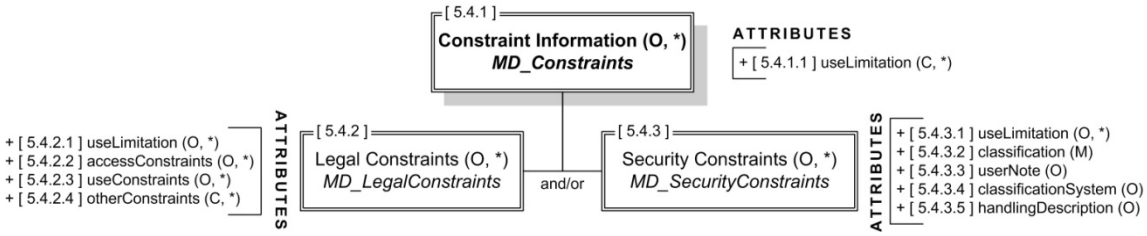
Conceptual designs for a:	might lack:
▪ project task diagram	specific due dates,
▪ landscape plan	plant size and numbers.
▪ computer system requirements.	specifics as to hard-drive capacity or memory

In the same way, the ISO 19115 UML does not provide the specific guidance needed to format a metadata record in a standardized presentation. The XML schema, ISO 19139, was developed to provide a consistent manner for presenting and transferring the metadata among different organizations and applications.

So…what does this mean to the metadata creator? Does one have to learn UML to create ISO metadata? No. The ISO 19139 XML schema provides the structure needed to implement the standard and to develop metadata editors and other resources for ISO 19115 implementation

In the NAP, the ISO 19115 UML diagrams were replaced with new diagrams intended to simplify reader comprehension. The diagrams are supported by a data dictionary and best practice recommendations intended to provide insight and guidance. It should be noted that these diagrams are under review as to their conformance to the ISO 19115 UML and that revisions to the diagrams have been recommended for consideration in future NAP work.

An example of the NAP Vers. 1 (2009) diagram and associated data dictionary complete with best practices are presented below and explained in detail the following section.



NAP Diagram for Constraint Information
(Source: *North American Profile of ISO 19115:2003 – Geographic information – Metadata*)

Clause No	Name/ RoleName	Multiplicity	Type	Description	Best Practices
	Constraint Information	O,Repeatable	MD_Constraints	The limitations, restrictions, or statements on the resource fitness for use.	
1.1.1.1	useLimitation	C,Repeatable	free text (CharacterString)	Statement on the fitness for use or limitations on the use of the resource or metadata.	The attribute *useLimitation* is mandatory unless *Legal Constraints* (5.4.2) or *Security Constraints* (5.4.3) is provided.
	Legal Constraints	O,Repeatable	MD_LegalConstraints	The legal restrictions or prerequisites to using the resource or accessing the metadata.	Legal constraints should be repeated for multiple legal access constraints such as those associated with privacy, sensitivity, and statutory.
1.1.1.2	useLimitation	O,Repeatable	free text (CharacterString)	Statement on the fitness of use or limitations on the use of the resource or metadata.	
1.1.1.3	accessConstraints	O,Repeatable	CodeList napMD_RestrictionCode	Limitations on access to the resource or metadata to protect privacy, intellectual property, or any	Select *accessConstraints* from napMD_RestrictionCode.

Clause No	Name/ RoleName	Multiplicity	Type	Description	Best Practices
				special limitations.	
1.1.1.4	useConstraints	O,Repeatable	CodeList napMD_RestrictionCode	Restrictions or limitations or warnings to protect privacy, intellectual property, or other special restrictions on the resource or the metadata.	Select *useConstraints* from napMD_RestrictionCode.
1.1.1.5	otherConstraints	C,Repeatable	free text (CharacterString)	Other restrictions or legal prerequisites for accessing the resource or metadata.	*otherConstraints* shall be provided if *accessConstraints* (5.4.2.2) or *useConstraints* (5.4.2.3) is set to "otherRestrictions." For an example: "Data only to be used for the purposes for which they were collected."
	Security Constraints	O,Repeatable	MD_SecurityConstraints	Restrictions applied to the resource or metadata to protect security concerns.	
1.1.1.6	useLimitation	O,Repeatable	free text (CharacterString)	Statement on the fitness of use or limitations on the use of the resource or metadata.	
1.1.1.7	classification	M	CodeList napMD_ClassificationCode	Name of the handling restrictions on the resource or the metadata.	Select *classification* from napMD_classificationCode.
1.1.1.8	userNote	O	free text (CharacterString)	An explanation of the classification level applied to the resource or metadata.	
1.1.1.9	classificationSystem	O	free text (CharacterString)	Name of the security classification system.	
1.1.1.10	handlingDescription	O	free text (CharacterString)	Additional information regarding security restrictions on handling the resource or metadata.	*handlingDescription* can serve as a place to state that the data had been reviewed and had been approved for release.

NAP Data Dictionary for Constraint Information

(Source: *North American Profile of ISO 19115:2003 – Geographic information – Metadata*)

Organization of ISO 19115 and the NAP

ISO 19115 metadata is organized into:

Sections

Classes/Subclasses

Attributes

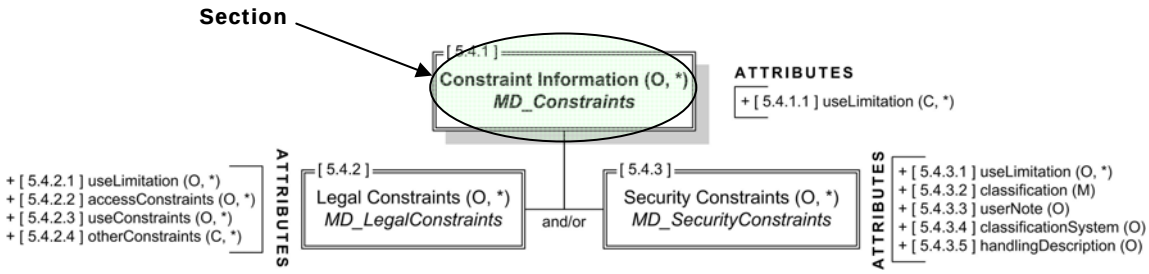
Domains

Code Lists

In *general* terms, ISO 19115 organizational components can be related to CSDGM components in the following manner:

ISO 19115	CSDGM
Section	Section
Class/Subclass	Compound Element
Attribute	Element
Domain	Domain
Code List	Fixed Domain Values

Sections



Example ISO 19115 Section

(Source: *North American Profile of ISO 19115:2003 – Geographic Information – Metadata*)

Sections are the uppermost component of the metadata organization. The eleven main sections are described below. Note that since ISO 19115 applies to both geospatial data and services the term 'resource' is used as a collective reference to both.

Metadata Record Information
: Information about the metadata record

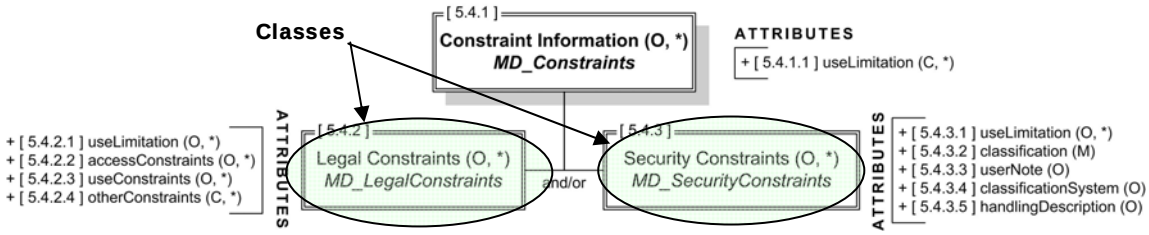
Identification Information
: Information needed to uniquely identify the resource (data or service)

Constraint Information
: Information about use limitations, legal constraints and security constraints

Data Quality Information
: Information about completeness, thematic accuracy, logical consistency, positional accuracy, temporal accuracy, sources used, and processing steps

Maintenance Information
Information about how the resource is updated
Spatial Representation Information
Information about the grid or vector models used to represent features
Reference System Information
Information about the coordinate system used to georeference the resource
Content Information
Description of the content of the database (feature catalog), discrete grid cells (coverage), and continuous grid cells (image)
Portrayal Catalogue Information
Information of the symbols used to depict features on a map
Distribution Information
Information about how one can access the resource
Application Schema
Information about published and unpublished data and process models that were applied to the resource

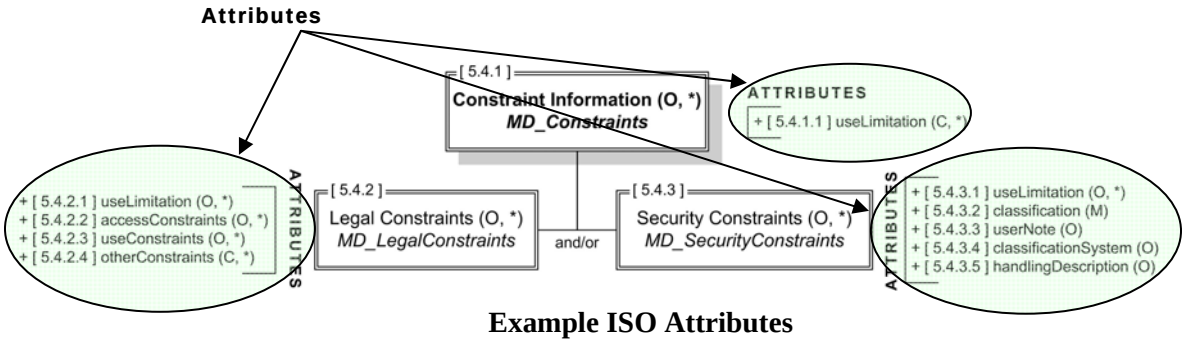
Classes



Example ISO Classes
(Source: *North American Profile of ISO 19115:2003 – Geographic Information – Metadata*)

Classes are the secondary component of the metadata organization. Classes are used to organize related metadata information and may contain both subordinate 'sub classes' and attributes.

Attributes



Example ISO Attributes
(Source: *North American Profile of ISO 19115:2003 – Geographic Information – Metadata*)

Attributes are the most primary component of the metadata organization. Unlike Sections and Classes, attributes contain actual data about the resource. Attributes can be associated with Sections and Classes and the initial letters of attribute names are written in lower case.

Attribute Types
Compared to the CSGDM, the ISO 19115 provides an expanded list of designated data types that includes:

- Boolean (true/false)
- Date / Time
- Distance
- Free Text
- Generic Name
- Integer
- Measure
- Real Number
- URL

In addition, the ISO 19115 employs a significant number of code lists that serve as fixed domains. Each code list includes a set of values and definitions for each value. The values are referenced, internally, by a numeric value. By employing code lists, ISO 19115:

- enhances the ability to locate specific content by standardizing the descriptors, and
- enhances the international value of the content by the association of each value with a numeric code that requires no language translation.

For example, the CSDGM element ‘Originator’ is defined as ‘the name of an organization or individual that developed the data set’. This presented a challenge to organizations that procure data development from outside vendors but assumed ownership upon completion. The CSGDM did allow for the inclusion of multiple ‘Originators’ and the formal designation of the ‘Originator’ as the data set editor (ed.) or compiler (comp.) but neither option fully addressed the problem and metadata creators were left to generate their own solutions to the dilemma.

ISO 19115 addresses this problem by providing a code list of contact ‘roles’ that can be used to more fully describe the relationship of the individual or organization to the resource.

ISO Section: Contact
Attribute: role
Type: Code List: Role Code
resource provider
custodian
owner
user
distributor
originator
point of contact
principal investigator

processor
publisher
author
collaborator
editor
mediator
rights holder

ISO 19115 code lists are extendible. A registry of the NAP code lists and user extensions is maintained at: http://www.fgdc.gov/nap/metadata/register/index.html *(accessed October 14, 2011).*

Best Practices

Because ISO 19115 is a more flexible document than the CSDGM with fewer mandatory elements, the NAP effort included the development of 'best practices' that provide guidance on:

- the intended purpose of the component
- the level of detail that best supports metadata function
- available thesauri, and other standardized resources that can be used
- suggested values
- conditions
- other guidance that may facilitate stronger content or utilization of the standard.

Multiplicity

'Multiplicity' is an indication of the number of times a component must or can be used. The CSDGM includes a 'Conditionality' feature to indicate whether an element is: 'mandatory', 'mandatory if applicable' or 'optional' and a 'Repeatability' feature to indicate whether an element can be used more than once. 'Multiplicity' combines these features into a single feature to indicate if an element is:

- M – Mandatory, must be provided, one value only
- C – Conditional, must be provided based on some described condition. Conditions may include the requirement that related attributes are also provided or that the attribute itself becomes 'mandatory' when related attributes are provided
- O – Optional, at the discretion of the provider
- 'Repeatable'

The conditions are combined if multiple values are allowed. . For example, 'O, Repeatable'.

Multilevel Metadata

Metadata creators have long struggled with the concept of data granularity, The CSDGM was designed to document a single 'data set' defined as "a collection of related data". This vague description becomes increasingly problematic with the advent of new data and feature models and variations among geospatial software data organization and management applications.

ISO 19115 introduces a hierarchy for the documentation of related 'levels' of data including:

- collection
- series
- dataset
- feature
- attribute.

When creating a metadata record, the author indicates the level of data that is being documented as well as related data levels. In this manner, the author records the metadata content specific to the specified data level. For example, data quality measurements vary among features but the values generally are aggregated when reported for the data set as a whole. To extend the example, if a single instance of an attribute, such as one out of a group of salinity measurements, varied widely from the others, the 'attribute' level of metadata would enable the author to report the anomaly and give context to the reported 'average salinity value'.

While the CSDGM supported the inclusion of multiple 'quantitative' accuracy reports, the reports were contained within the broader metadata record. By creating metadata files specific to the level of data, the metadata is tied to the attribute, feature, data set or series and can be 'inherited' when compiled into a new data product.

Multilevel metadata facilitates the:

- creation of highly specific metadata
- reuse of metadata
- distribution of metadata creation among data contributors.

Implementation of hierarchical metadata will be a challenge. Data managers will need to determine the optimal level of documentation to support their project and organization and to identify the metadata content that applies to each level. In addition, as metadata is reused, as in the case of a road feature that is integrated into multiple datasets, the metadata will be duplicated and potentially altered presenting additional concerns about metadata lineage and accuracy. These challenges serve as great opportunities for geospatial software developers to create applications that effectively incorporate hierarchical metadata, facilitate user implementation and ease the burden on the metadata creator to scope the effort.

New Content

The general content of ISO 19115 correlates closely to that of the CSDGM. Metadata records developed using the CSDGM can be manually crosswalked to ISO 19115. XML stylesheets (XSLTs) have also been developed to automatically translate CSDGM metadata records to ISO 19115 and ISO 19115-2. However, because ISO 19115 and ISO 19115-2 include new elements, extended elements, changed domains, and fixed code lists, translated records may not result in complete or compliant ISO records. Additional information may be required. New content was developed in ISO 19115 to:

- address the international scope of the standard
- expand existing elements to provide more specific information
- change domains and incorporate code lists to standardize information
- address the documentation of new technologies and data models, and
- create new elements to facilitate discovery and access.

International Scope

To accommodate the writing of metadata for multiple nations and international organizations, ISO 19115 includes attributes to record:

- *Dataset Language*
- *Dataset Character Set*
- *Metadata Language*
- *Metadata Character Set.*

While most U.S. applications will include default values, in Canada there is significant, and sometimes required, use of multiple languages. In addition, the attributes will prove useful to U.S. researchers searching for data that extend across international boundaries.

Expansion of Existing Elements

The CSDGM was designed for use by a broad spectrum of geospatial data developers and consumers. When applying the CSDGM, some metadata creators have found it limiting with respect to the level of detail they are able to capture. ISO 19115 addresses many of these shortcomings by expanding some CSDGM elements to include a more robust set of attributes. The following elements have been significantly expanded.

Data Quality

Like the CSDGM, ISO 19115 Data Quality section includes attributes to describe both data assessment reports and lineage (sources and processing). The ISO 19115 Lineage class was extended to include information about:

- source reference systems
- source extent
- process step rationale.

The ISO 19115 Assessment Report classes were extended to include:

- thematic accuracy
- temporal accuracy.

The ISO 19115 Data Quality reporting section also provides a standardized set of attributes for use within all Assessment Report types: completeness, thematic accuracy, logical consistency, positional accuracy, attribute accuracy, and temporal accuracy. The Data Quality attribute set includes:

- name of measures
- measure identification (for registered standard data quality procedures)
- measure description
- evaluation method type
- evaluation procedure citation
- date and time
- result.

It should be noted that significant changes to the Data Quality section are proposed in the pending five year update to ISO 19115, ISO 19115-1.

Since the attributes set is more robust than the CSDGM and can be repeated multiple times, it enables a more detailed and consistent documentation of assessment reporting.

Citations

Similar to the use of a consistent set of attributes to describe Data Quality assessments, the use of Citations has been expanded within ISO 19115 to document external references and authorities for a broader range of metadata components including:

- Keyword thesauri
- Data Quality evaluation procedures
- Data Quality conformance specifications
- Spatial Representation grid referencing parameters

- Reference System authorities
- Data Content feature catalogs (data dictionaries)
- Portrayal catalogs (symbology sets)
- Application schemas (models and analyses)
- Metadata Identifier (unique file name/number) authorities.

Online Resource
Given the expanded number of available online resources and the array of functions provided by these resources, the following attributes are provided to better describe online linkages referenced within the metadata:

- linkage (URL)
- connection protocol (http, ftp, etc.)
- applications used by the resource (web mapping, decision support, etc.)
- name
- description
- function (code list).

Maintenance Information
Metadata can serve as a geospatial data management resource. If the metadata information is robust and kept current, it can be used to identify data sets in need of archive or update, to schedule updates, and to inform data implementers of pending updates. ISO 19115 enhances metadata data management capabilities by providing the following attributes:

- maintenance and update frequency
- date of next update
- user defined maintenance frequency
- scope of the update (code list)
- description of the scope of the update
- requirements for maintaining the resource
- maintenance contact.

Changed Domains
Changes to the domains of several metadata elements represent the greatest challenge facing the translation of CSDGM metadata to ISO 19115 format. In those cases where domains have simply been extended, the metadata information will translate with the original value and it will be dependent on the metadata manager to change to a newer, more specific value. However, in those cases where CSDGM domains values were dropped, converted from free text to fixed domain (‘Code list’), or the format changed, translation will be dependent on either manual input or the use of smart strategies (look up tables, if/then programming, etc.) by translation applications. Problems with the fixed domain ‘Code lists’ should be minimal since only eight of the twenty-nine ISO 19115 code lists relate to existing CSDGM elements and, in those eight cases, the CSDGM domain values, with the exception of the ‘free text’ values, are either fully integrated or can be mapped to a value in the related ISO 19115 code list.

Date elements present another challenge. The CSDGM allows the use of text values such as ‘unknown’, ‘present’, and ‘unpublished’ in date fields. ISO 19115 restricts dates to numeric values. CSDGM-to-ISO translation applications can provide options for replacing text strings in date fields by either prompting for a value or utilizing xml capabilities such as ‘nilReason’ to clarify why a required item is missing, e.g. nilReason=”unknown”.

New Technologies

Many new geospatial data technologies have emerged or expanded since the 1998 publication of the CSDGM Version 2.0 including, but not limited to:

- web mapping
- data models
- online data catalogs, warehouses and sales
- online data processing
- ontologies, thesauri, data hierarchies and classification systems.

While not 'datasets', these 'services' are valuable geospatial data resources that are frequently coupled with geospatial data. For example, a particular data set may be available for use in a specific web mapping application, available for download from a data warehouse, or developed using a specific data model. Even when not coupled directly with a data set, data developers benefit greatly from the ability to discover, access, assess, apply, and archive these services.

The NAP expands on the ISO 19115 'service' metadata content and incorporates additional attributes from *ISO 19119:2005 Geographic Information – Services*. Similar to the 'Data Identification' class, the 'Service Identification' includes many of the same attributes but provides additional service specific attributes such as:

- service type, for example, 'OGC Catalog'
- coupled data resources
- operations the service can perform.

Expansion of the service metadata elements is also expected in the pending five year update of ~~to~~ ISO 19115 to ~~,~~ ISO 19115-1.

New Metadata Elements

After ten years of application, it is inevitable that users would discover documentation needs simply not perceived earlier. ISO 19115 includes the following new elements to address those shortcomings.

Vertical Extent

While the CSDGM provided element to document the vertical coordinate system of the data, there was no means to express the vertical *extent* of the data. In the same way that geographic extent establishes a footprint on the earth and temporal extent sets a time period for the content, ISO 19115 provides new attributes to establish the maximum and minimum elevations of the data; characteristics common to atmospheric, bathymetric, subsurface geology and elevation data.

Topic Categories

One of the most common requests presented to metadata managers and trainers is the need for a single 'thesaurus' for theme keywords. The CSDGM provides no specific guidance as to a standard list of the key words but managers and trainers have long recommended that metadata creators use published thesauri such as the *Global Change Master Directory* (GCMD) and the *Canadian Core Subject Thesaurus* (CST) as well as discipline-specific thesauri including *Cowardin Wetlands Classification System* and the *Glossary of Geologic Terms*.

ISO 19115 provides the 'Topic Categories' element to serve as a single, high level set of subject headings that can be used to coarsely sort data into logical thematic

categories. Metadata creators are required to select one or more of the Topic Categories. These concatenated terms with their unusual capitalization appear odd because they reflect the UML origins of the standard. Metadata editors and other applications can present the terms in a more easily read style but the content should remain as presented in the table below.

farming	intelligenceMilitary
biota	inlandWaters
boundaries	location
climatologyMeteorologyAtmosphere	oceans
economy	planningCadastre
elevation	society
environment	structure
geoscientificInformation	transportation
health	utilitiesCommunications
imageryBaseMapsEarthCover	

ISO 19115 Topic Categories
(Source: *ISO 19115 Geographic Information – Metadata*)

Online Linkage Function Code
As discussed earlier, the CSDGM Online Linkage option has been expanded to include a number of descriptors including a 'Function Code'. This code list is used to qualify the operations of any given URLs. By providing this information, the user has better insights as to 'what happens' when they select a dynamic link from the metadata record. ISO 19115 online linkage functions include:

- download
- information
- offline access
- order
- search
- upload
- web service
- email service
- browsing
- file access
- web mapping service

Portrayal Catalog
In the same way that Topic Categories encourage the use of a standardized set of subject headings, 'Portrayal Catalog' enables metadata creators to provide a citation for the use of standard symbologies. This may include externally published symbologies, such as the *Anderson Land Use Land Cover Mapping Standard*, or internally developed symbologies. While ISO 19115 only supports reference to a citation, a related TC 211 Standard, ISO 19117 Portrayal Catalog, outlines the methodology for describing symbols.

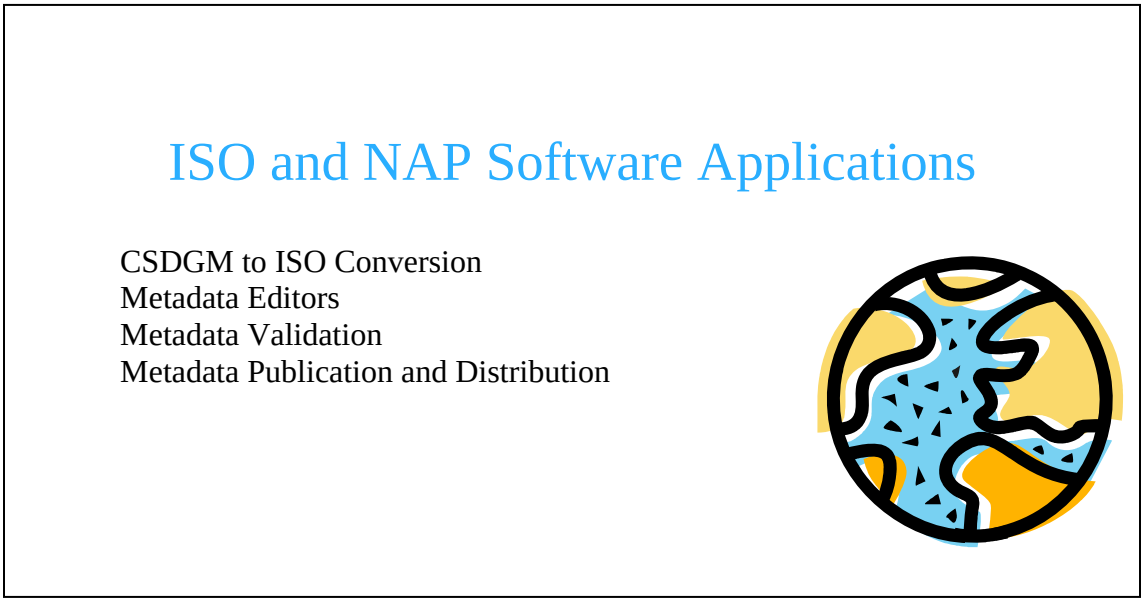
Application Schema

In the same spirit of standardization, the Application Schema provides a method to describe the use of standard software-applications such as a hydrology data model, a process methodology for extracting impervious services from imagery or delineating a watershed, or a land suitability analysis. In addition to a citation of the application schema, ISO 19115 provides attributes to describe the schema language, constraints to the application, software dependencies as well as the ability to fully embed the schema as an ASCII file, graphic depiction or operational software development file.

Eliminated Elements

CSDGM content correlates closely to the ISO 19115 content. Though some CSDGM elements have been expanded into a broader set of attributes in ISO 19115 and will require manual review/edits, most content can be transferred directly from one standard to the next. A few CSDGM elements, however, were deemed unnecessary and not included in ISO 19115. Items were deprecated for the following reasons:

- Based on use of old technology:
  - Dial-up instruction for the use of a modem to transfer data
- Information is apparent from the values provided
  - Planar Coordinate Encoding Method
  - Distance and Bearing Representation
- Rarely applied or applied as intended
  - Landsat satellite and path number: intended to ‘define the space oblique mercator projection’ and not intended for documenting data collected via a remote sensing application
- Already specified by ISO 19115 or some affiliate ISO standard
  - Metadata Time Convention.



ISO and NAP Software Applications

A number of software applications have been developed to assist in ISO 19115, ISO 19115-2, and NAP implementation. Links to resources referenced below are provided in **Appendix C:** ISO and NAP References and Resources.

CSDGM to ISO Conversion

GIS vendors, Federal agencies, and other organizations have developed various transforms to support CSDGM to ISO 19115 and 19115-2 conversion. Some transforms are intended as internal resources to convert application or organization specific metadata but others have been developed to support a broader community.

Working with examples and guidance from the FGDC metadata community, Intergraph, under contract to the FGDC, has developed a spreadsheet (.xsl) mapping of all FGDC content to elements of ISO-19115 (Metadata), ISO-19110 (Feature Catalogue), and related standards. Once the mapping of Feature Catalogue elements is complete, this spreadsheet will be implemented as an XSLT transformation file that can be used to convert an FGDC-compliant XML file into an ISO-19139-compliant XML file. It will be provided with a simple interface and user help program. The transformation will process most FGDC elements, and others will be included as "supplemental" elements as part of the XML output. Users will be able to copy and modify the XSLT for their own purposes.

NOAA developed a set of transforms (XSLT's) to support CSDGM to ISO 19115 and 19115-2 conversion within its MERMAid metadata editor. However, These transforms have been developed as public resources and NOAA is soliciting public input to refine the transforms. The transforms include certain levels of logic that are able to interpret content to ensure minimal, if any, manual edits. Users are encouraged to apply and customize the transforms to support their own data and workflow.

Metadata Editors

Several editors have been developed to support the creation of ISO 191** series metadata. Most editors are designed to support the generic standard with options to support various national and community profiles including the NAP.

The FGDC maintains an online review of ISO Metadata Editors that includes information provided by the application developers about editor capabilities and profile support as well as independent user reviews. In the 2009 version of the review, approximately one third of the respondents indicated that they were in the process of developing NAP metadata creation capabilities. The CSDGM community is encouraged to monitor the ISO Metadata Editor Review and the FGDC Metadata website, http://www.fgdc.gov/metadata/iso-metadata-editor-review (*accessed October 14, 2011*), for information about available ISO metadata editors. Users are encouraged to explore the ISO editors to learn more about ISO 19115 metadata.

Many metadata creators have experienced challenges using ~~as~~ their traditional editors to implement ISO. These challenges include the loss of CSDGM content, format and auto population features. Most editors have worked hard to re-establish these functions and the U.S. National Park Service (NPS), Environmental Protection Agency (EPA) and other agencies have developed applications to support these efforts.

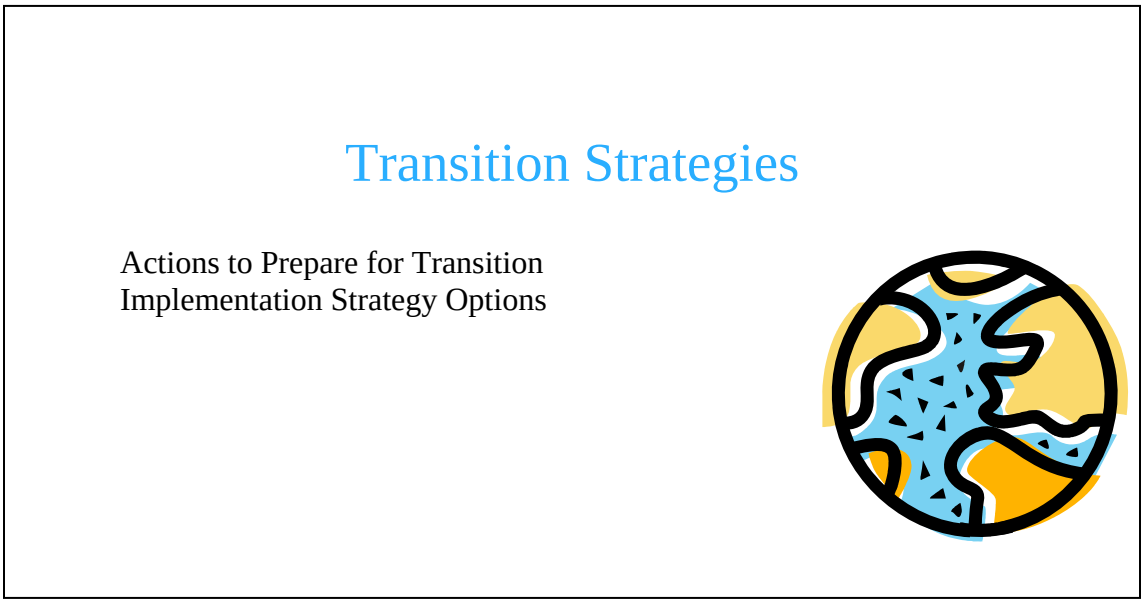
Metadata Validation

Because ISO 19115 metadata are presented using XML, the record can be easily validated using schemas and schematrons. Schemas (.xsd files) define the XML rules for structure, content, and semantics and define the objects and attributes of an XML document. Schematrons are stylesheets (XSLTs) that further enforce rule-based validation not addressed by the schemas, for example, 'end' dates cannot occur before 'begin' dates. In addition, schematrons can be applied to present cryptic error messages using more descriptive, 'human', language.

Current ISO 191** series XSDs do not support NAP-specified domains and conditionality. An 'authoritative' NAP XSD is currently under development by the NAP development team that will use the ISO 191** series schemas as the base. Several GIS vendors support ISO 19139 compliant metadata validation.

Metadata Publication and Distribution

ISO 19115, ISO 19115-2 and NAP metadata can be published and made available via web folders and other application that do not require specific file content or format. Formal metadata publication applications that standardize metadata to support enhanced search and discovery, including Z39.50 Clearinghouses and the Data.gov geospatial data catalog at http://www.data.gov/catalog/geodata (*accessed October 14, 2011*), require new utilities and validation routines to enable the publication of ISO formatted metadata. There are numerous applications that can validate and publish ISO 19115 metadata but, as discussed above, they may not be specific to the NAP or ISO 19115-2.



Transition Strategies

As stated earlier, ISO 19115, ISO 19115-2 and NAP are national (ANSI) and Federal (FGDC) standards ~~and while~~are ~~-~~not currently *required* for use, agencies and organization should prepare for implementation and consider the following actions and implementation strategy options.

Actions to Prepare for Transition

Organizations can take actions now that will facilitate future implementation of ISO metadata. The following steps outline a strategy for updating the content and format of your CSDGM metadata to ease translation and engaging colleagues so that they are informed and ready.

1. Assess current data holdings to determine:
   - current value and frequently accessed
   - current value and infrequently accessed
   - historic value
   - none or limited value.

2. Delete or archive data and metadata considered of none or limited value.

3. Archive data of historic value to remote drive or media.

4. Convert the metadata records of current value data to XML if stored in other format.
   CSDGM records that are in text (.txt) or SGML format can be converted to XML via using the:
   - 'output format' function of most metadata editors
   - Metadata Parser (mp) utility developed by Peter Schweitzer (USGS) and available for free from: http://geology.usgs.gov/tools/metadata/ *(accessed October 14, 2011)*
   - Metadata Enterprise Resource Management Aid (MERMAid) available from: http://www.ncddc.noaa.gov/metadata-standards/mermaid/ *(accessed October 14, 2011)*

CSDGM metadata records that are stored as HTML cannot be directly convert to XML. However, ESRI outlines a detailed procedure for converting HTML to txt that can then be converted to XML using another Peter Schwitzer (USGS) utility, 'Chew N Spit' (cns). An article describing the process is available from:
http://support.esri.com/index.cfm?fa=knowledgebase.techarticles.articleShow&d=23071 *(accessed October 14, 2011).*

5. Add one or more Topic Categories to the CSDGM metadata Theme_Keywords of data deemed of current value.
   To ensure that CSDGM to ISO translation applications are able to identify the Topic Categories, it is recommended that you enter the Topic Category name with the concatenation and capitalization presented in the standard, for example, 'geoscientificInformation', 'imageryBaseMapsEarthCover', 'intelligenceMilitary', etc

   Starting with most frequently accessed holdings and distributing the effort over a reasonable period of time, consider one of the following methods:
   - **Manual Edit of Theme_Keywords**
     For small collections of metadata records, applicable Topic Categories can be added to the Theme_Keywords of each metadata record using a metadata editor. The Topic Categories should be added using 'ISO 19115 Topic Category' as the Theme_Keyword_Thesaurus.
   - **Automated Edit of Theme_Keywords**
     For large collections of metadata, it may be more efficient to 'map' your existing Theme_Keywords to the ISO Topic Categories and automate the addition of the Topic Categories using either a Look-up Table (LUT) or Stylesheet (XSLT).
     Resources for each are included in Appendix A.

6. Educate and inform management and technical staff
   Both management and technical staff should be made aware of the new standard and options for implementation. This can be achieved through discussion, meetings, and presentations. An individual or working group should be tasked to lead the effort and to consider the following recommendations.
   - Review this document, referenced sources and other available information to determine and document the ISO and NAP features and issues of key significance to the organization.
   - Meet with management to discuss options for implementation and determine the appropriate level of implementation.
   - If management recommends ISO or NAP implementation, meet with technical staff to discuss preparation strategies and outline a series of implementation tasks complete with assigned roles and due dates.
   - Manage the implementation effort and participate quarterly telecons of the FGDC Metadata Working Group (info at: http://www.fgdc.gov/participation/working-groups-subcommittees/mwg/index_html) *(accessed October 14, 2011).*
   - Participate in national dialogs regarding ISO and NAP implementation at conferences, meetings and the metadata list server at http://lists.geocomm.com/mailman/listinfo/metadata *(accessed October 14, 2011).*

Implementation Strategy Options

Organizations that decide to move forward with ISO or NAP are encouraged to consider the following implantation strategy options.

1. Continue with CSDGM metadata creation with no ISO or NAP implementation
   Benefits:
   - continued return on investment of existing CSDGM expertise, capabilities, and resources

   Considerations:
   - no increases in metadata capabilities and richness
   - challenges using ISO and NAP metadata created by others including need for ISO and NAP-to-CSDGM translation application and staff time and expertise to convert
   - eventual 'phasing out' of CSDGM support by metadata applications

2. Begin creating new metadata using ISO or NAP and convert CSDGM records as used or distributed
   Benefits:
   - greater flexibility in developing custom metadata templates because of fewer mandatory elements and more robust optional elements
   - improved discovery and use of geospatial services because of added elements for the documentation of geospatial services'
   - more efficient metadata compilation using code lists and multilevel metadata features
   - translation effort is limited to active datasets
   - improved data portal participation and discovery utilizing Topic Categories
   - improved international interoperability

   Considerations:
   - increased time and expense required for ISO or NAP metadata training and software implementation
   - effective multilevel metadata performance requires development and implementation of a strategic approach

3. Full implementation of ISO or NAP for new and existing metadata records
   Benefits:
   - similar to partial implementation benefits listed above
   - no need to translate CSDGM metadata at time of use or distribution
   - opportunity to establish a fresh approach to metadata that encourages new ways of thinking about metadata and broader participation.

   Considerations:
   - similar to partial implementation concerns listed above
   - translation of all metadata requires development and implementation of a strategic approach and significant effort

Appendix A

Automation of Topic Categories as Theme_Keywords

Appendix A: Automation of Topic Categories as Theme_Keywords

Directions for scripting metadata to extract keywords using Tcl, created by Peter Schweitzer
http://geology.usgs.gov/tools/metadata/tools/doc/mq.html *(accessed October 14, 2011)*

Use of ISITE software to build a Theme Keyword LUT
http://registry.gsdi.org/statuschecker/counselorTools.php *(accessed October 14, 2011)*

A Topic Category Thesaurus for Theme Keywords created by Bruce Westcott
http://registry.fgdc.gov/reports/docs/smms_thesaurus.txt *(accessed October 14, 2011)*

How to use XML templates and Stylesheets to transform data
http://en.wikipedia.org/wiki/XSL_Transformations *(accessed October 14, 2011)*

Appendix B

Glossary of Terms and Acronyms

Appendix B: Glossary of Terms and Acronyms (as used in this publication)

ANSI- The American National Standards Institute.

Attribute – Most basic metadata element, requires a data value.

Class / Sub Class – Description of a set/ subset of metadata elements.

Code Lists – Fixed list (domain) of values for a specified metadata element.

CSDGM – The Content Standard for Digital Geospatial Metadata.

Domain – Description of a set of valid values for a specified metadata element.

INCITS L1 – The InterNational Committee for Information Technology Standards; Technical Committee L1, Geographic Information Systems.

ISO – The International Standards Organization.

Ontology – The organization of information into a relationship-based hierarchy such as a subject index or disciplines within a field of knowledge.

Section – Category of metadata information including classes, sub-classes and attributes.

Services – The online functions provided through a user interface <adapted from ISO 19915 definition>.

UML – The Unified Markup Language – a technical language used to define entities and visualize the relationship among entities.

XML – The Extensible Markup Language – a technical language used to describe data in a way that supports data transfer and storage. Differs from HTML in that HTML is used to describe how to display data.

XSD – The XML Schema Definition.

XSLT – The Extensible Stylesheet Language Transformations (XSLT), or ‘transform’ is an XML-based language used for the transformation of XML documents into other XML or "human-readable" documents. The original document is not changed; rather, a new document is created based on the content of an existing one. XSLT is the language for transforming the XML documents. XPath is a query language used to select nodes from an XML document. XPath is used by XSLT as a mechanism to access and refer to a place within the XML document in order for it to be transformed. XPath and XSLT share the same data model library of functions and operators.

Appendix C

ISO and NAP References and Resources

Appendix C: ISO and NAP References and Resources

ANSI publication: *North American Profile of ISO 19115:2003 Geographic Information - Metadata*
http://webstore.ansi.org/RecordDetail.aspx?sku=INCITS+453-2009 *(accessed October 14, 2011)*

CSDGM to ISO/NAP Conversion xslt
http://www.ncddc.noaa.gov/metadata-standards/metadata-xml/ (*accessed October 14, 2011*)

EPA Metadata Editor (EME)
https://edg.epa.gov/EME/ (*accessed October 14, 2011*)

FGDC ISO/NAP Webpage
http://www.fgdc.gov/standards/projects/incits-l1-standards-projects/NAP-Metadata *(accessed October 14, 2011)*

FGDC ISO Metadata Editor Review
http://www.fgdc.gov/metadata/iso-metadata-editor-review *(accessed October 14, 2011)*

FGDC Online Training Materials
http://www.fgdc.gov/training/training-materials *(accessed October 14, 2011)*

GeoCommunities Metadata Listserver
http://lists.geocomm.com/mailman/listinfo/metadata *(accessed October 14, 2011)*

ISO Publication: *ISO 19110: Geographic Information – Methodology for Feature Cataloging* Factsheet
http://www.isotc211.org/Outreach/Overview/Factsheet_19110.pdf *(accessed October 14, 2011)*

ISO Publication: *ISO 19115:2003 Geographic Information – Metadata* Factsheet
http://www.isotc211.org/Outreach/Overview/Factsheet_19115.pdf *(accessed October 14, 2011)*

ISO Publication: *ISO 19115-2:2009 Geographic Information – Metadata = Part 2: Extensions for Imagery and Gridded Data* Factsheet
http://www.isotc211.org/Outreach/Overview/Factsheet_19115-2.pdf *(accessed October 14, 2011)*

ISO Publicationt: *ISO 19119:2005 Geographic Information – Services* Factsheet
http://www.isotc211.org/Outreach/Overview/Factsheet_19115.pdf *(accessed October 14, 2011)*

ISO/NAP Metadata Register (includes code lists)
http://www.fgdc.gov/nap/metadata/register/index.html *(accessed October 14, 2011)*

National Park Service FGDC Synchronize Tool For ArcGIS 10
http://resources.arcgis.com/gallery/file/geoprocessing/details?entryID=8DA58FD4-1422-2418-8836-2FE83C67EF7D
(accessed October 14, 2011)

Unified Modeling Language (UML) Tutorial by Exforsys
http://www.exforsys.com/tutorials/uml.html *(accessed October 14, 2011)*

USGS (Pulford) Updates on ArcGIS 10 Implementation
ftp://ftpext.usgs.gov/pub/nonvisible/cr/cprice/MetadataAt10.pdf *(accessed October 14, 2011)*

(XML) Extensible Markup Language Tutorial by w3schools
http://www.w3schools.com/xml/default.asp *(accessed October 14, 2011)*