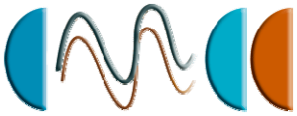
Technical Reports



CENTRO EURO-MEDITERRANEO
PER I CAMBIAMENTI CLIMATICI

SCO – Scientific Computing and Operations Division

CMCC Metadata Agreement

Sandro Fiore
Scientific Computing and Operations, CMCC
University of Salento, Lecce

Salvatore Vadacca
Scientific Computing and Operations, CMCC

Giovanni Aloisio
Scientific Computing and Operations, CMCC
University of Salento, Lecce

Centro Euro-Mediterraneo
per i Cambiamenti Climatici
www.cmcc.it

March 2008 ■ TR18

CMCC Metadata Agreement

Summary

This technical report describes the CMCC Metadata Agreement; the aim of the schema is to classify the CMCC data production (input, intermediate data and output of the experiments), models, services, etc.
The CMCC Metadata Agreement complies with the ISO 19115 standard for Geographic Information.
Metadata entities and code lists are fully described in this technical report.

Keywords: Earth Science, Metadata, schema, ISO 19115, catalogue, XML.

JEL Classification:

Version: 1.0
Release Date: 14.03.2008

Address for correspondence:
Sandro Fiore
Euro-Mediterranean Centre for Climate Change
Viale Gallipoli, 49
73100 Lecce, Italy.
E-mail: sandro.fiore@cmcc.it

Table of Contents

Introduction

Earth Science is strongly becoming a data intensive and oriented activity. Petabytes of data, big collections, huge datasets are continuously produced, managed and stored as well as accessed, transferred and analyzed by several scientists and researchers at multiple sites.

From the data grid perspective, a key element to search, discover, manage and access huge amount of data stored within distributed storages is the related data and metadata framework.

The aim of this document is to define a metadata schema for the Euro-Mediterranean Centre for Climate Change data production.

The outline of this report is as follows: in Section 1 we deal with some useful definitions, whereas in Section 2 we present geospatial metadata. The state of the art is investigated in Section 3, presenting Dublin Core, ISO 19115, NMM, CSML and CF. In Section 4 we describe the CMCC Metadata Agreement, whereas Section 5is devoted to explain the Data Dictionary.

1. Definitions

The term ***Metadata*** has been introduced without a formal definition. That has led today to various definitions. The most common one is the literal translation:

- *metadata is data about data.*

Since the difference between data and information is merely a philosophical one of no relevance in practical use, other definitions are:

- *metadata is information about data.*
- *metadata is information about information.*
- *metadata contains information about that data or other data*

More sophisticated definitions are:

- *metadata is structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities.*
- *metadata is a set of optional structured descriptions that are publicly available to explicitly assist in locating objects.*

Metadata is mainly structured data which describes the characteristics of a resource. It shares many similar characteristics to the cataloguing that takes place in libraries, museums and archives. The term "meta" derives from the Greek word denoting a nature of a higher order or more fundamental kind. A metadata record consists of a number of pre-defined elements representing specific attributes of a resource, and each element can have one or more values.

Each metadata schema will usually have the following characteristics:

- a limited number of elements
- the name of each element
- the meaning of each element

Typically, the semantics is descriptive of the contents, location, physical attributes, type (e.g. text or image, map or model) and form (e.g. print copy, electronic file). Key metadata elements supporting access to published documents include the originator of a work, its title, when and where it was published and the subject areas it covers. Where the information is issued in analog form, such as print material, additional metadata is provided to assist in the location of the information, e.g. call numbers

used in libraries. The resource community may also define some logical grouping of the elements or leave it to the encoding scheme.

Even in the early phases of planning and designing it is necessary to keep track of all metadata created. It is not economical to start attaching metadata only after the production process has been completed. It is necessary for different groups of resource producers to cooperate using compatible methods and standards.

Several standards have been developed in the last years and some will be described in Section 3.

2. Geospatial metadata

Geospatial metadata (also geographic metadata, or simply metadata when used in a geographic context) is a type of metadata that is applicable to objects that have an explicit or implicit geographic extent, in other words, are associated with some position on the surface of the Globe. Such objects may be stored in a geographic information system (GIS) or may simply be documents, datasets, images or other objects, services, or related items that exist in some other native environment but whose features may be appropriate to describe in a (geographic) metadata catalogue (may also be known as a data directory, data inventory, etc.).

The U.S. FGDC (Federal Geographic Data Committee) describes (geospatial) metadata as follows:

A metadata record is a file of information, usually presented as an XML document, which captures the basic characteristics of a data or information resource. It represents the who, what, when, where, why and how of the resource. Geospatial metadata are used to document geographic digital resources such as Geographic Information System (GIS) files, geospatial databases, and earth imagery. A geospatial metadata record includes core library catalogue elements such as Title, Abstract, and Publication Data; geographic elements such as Geographic Extent and Projection Information; and database elements such as Attribute Label Definitions and Attribute Domain Values.

3. State of the art

3.1. Dublin Core

The Dublin Core metadata element set is a standard for cross-domain information resource description. It provides a simple and standardised set of conventions for describing things online in ways that make them easier to find. Dublin Core is widely used to describe digital materials such as video, sound, image, text, and composite media like web pages. Implementations of Dublin Core typically make use of XML and are Resource Description Framework based. Dublin Core is defined by NISO Standard Z39.85-2007.

The "Dublin" in the name refers to Dublin, Ohio, U.S., where the work originated from an invitational workshop hosted in 1995 by OCLC, a library consortium that is based there. The "Core" refers to the fact that the metadata element set is a basic but expandable "core" list.

The semantics of Dublin Core were established and are maintained by an international, cross-disciplinary group of professionals from librarianship, computer science, text encoding, the museum community, and other related fields of scholarship and practice.

The Dublin Core Metadata Initiative (DCMI) is an organization providing an open forum for the development of interoperable online metadata standards that support a broad range of purposes and business models. DCMI's activities include consensus-driven working groups, global conferences and workshops, standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices.

The Dublin Core standard includes two levels: Simple and Qualified. Simple Dublin Core comprises fifteen elements; Qualified Dublin Core includes three additional elements (Audience, Provenance and RightsHolder), as well as a group of element refinements (also called qualifiers) that refine the semantics of the elements in ways that may be useful in resource discovery.

3.2. ISO 19115

The objective of ISO 19115 is to provide a structure for describing digital geographic data.

This International Standard is intended to be used by information system analysts, program planners, and developers of geographic information systems, as well as others in order to understand the basic principles and the overall requirements for standardization of geographic information. This International Standard defines metadata elements, provides a schema and establishes a common set of metadata terminology, definitions, and extension procedures. When implemented by a data producer, ISO 19115 will:

1. Provide data producers with appropriate information to characterize their geographic data properly.
2. Facilitate the organization and management of metadata for geographic data.
3. Enable users to apply geographic data in the most efficient way by knowing its basic characteristics.
4. Facilitate data discovery, retrieval and reuse. Users will be better able to locate, access, evaluate, purchase and utilize geographic data.
5. Enable users to determine whether geographic data in a holding will be of use to them.

This International Standard defines general-purpose metadata, in the field of geographic information. More detailed metadata for geographic datatypes and geographic services are defined in other ISO 19100 series standards and user extensions.

3.3. NMM

The Numerical Model Metadata standard is an evolving international metadata standard intended for the exchange of information about numerical model codebases, the associated components and the models/simulations done using them.

By providing a standard convention for metadata describing the numerical codebase and its models/simulations, it is possible to extend and refine the ability of scientists to understand how the model output data was produced. and will provide clear, well-

defined and flexible metadata to enable users to search, retrieve, difference and compare numerical codebases and their models.

3.4. CSML

The Climate Science Modelling Language (CSML) has been developed as a first attempt to define structured semantic data models for the climate sciences. A prototype deployment across the curated data holdings of the British Atmospheric Data Centre and British Oceanographic Data Centre is being undertaken through the UK's NERC DataGrid project. These holdings include a wide range of observational and model data. CSML is expected to evolve and develop with its application; the primary source for up-to-date information is the CSML website.

As its name implies, CSML aims to provide semantically meaningful models of climate science data types. The emphasis on semantics and interoperability distinguish CSML from the XML markup representations of climate data possible with NcML, CDML and ESML.

CSML has been designed explicitly with a dual purpose. In addition to modelling various climate science data types, it provides a wrapper mechanism to encapsulate the representation of those data objects in file-based storage artefacts.

CSML draws on ISO TC211 conceptual schemas for spatiotemporal referencing, geometry, etc., and employs GML for standardised encodings of these wherever possible.

3.5. CF

CF is a standard for "use metadata", whose aim is to distinguish quantities (physical description, units, prior processing, etc.) and to locate the data in space–time and as a function of other independent variables (coordinates). This is the kind of metadata that is used at the time the data is processed and displayed; it can be distinguished from "discovery metadata", which is used in catalogues for identifying datasets. CF provides only rather basic discovery metadata, such as ways to record where and how the file was produced.

CF has been developed over the last few years as a community project mainly by Brian Eaton, Jonathan Gregory, Bob Drach, Karl Taylor and Steve Hankin, with

suggestions and comments from many others. After two years of discussion and improvement on the WWW, it has reached the stage of its first mature release (CF-1.0). Meanwhile it has been adopted as the standard for several international projects such as AMIP, CMIP, ESMF and PRISM, and by various climate centres, including the Hadley Centre (Met Office, Exeter, UK), the National Center for Atmospheric Research (NCAR, Boulder, USA) and the Program for Climate Model Diagnosis and Intercomparison (PCMDI, Livermore, USA). The purpose of this article is to commend it for consideration by other projects.

CF is framed as a standard for data written in netCDF, but most of its ideas relate to metadata design in general, not specifically to netCDF. CF metadata could be contained in other formats, such as XML. The adoption of a common metadata standard makes conversion between different file formats a straightforward task of mapping and translating corresponding concepts.

As a file format for data exchange, netCDF has plenty to recommend it: it is portable, binary, easily translatable to and from an equivalent ASCII format, and supported by a lot of freely available software for processing and graphics, including the netCDF library itself, CDAT, Ferret and NCO. A utility to check conformance of a netCDF file to the CF standard has been made available by the Hadley Centre with PRISM support. At PCMDI some f90 subroutines are being developed to facilitate the writing of CF-conforming netCDF data, with the aim of making it easier for those creating simulated and observational datasets to adopt the format.

CF is intended for use with climate and forecast (hence "CF") data, for atmosphere, surface and ocean. It was designed with model-generated data particularly in mind, but should be equally applicable to observational datasets. Indeed, if observed and simulated data are to be compared, it will be helpful if their metadata describes them in the same way.

The general principles in the design of CF are as follows:

1. Data should be self-describing. No external tables are needed to interpret the file. For instance, CF doesn't use numeric codes, like GRIB does.
2. Conventions have been developed only for things users know they need. Instead of trying to foresee the future, features are added as required.
3. The aim is to avoid being too onerous for data-writers and users of data, as this will make the standard unattractive.

4. The metadata should be readable by humans as well as easily parsed by programs.
5. Redundancy is minimised—a good general principle because it reduces the chance of inconsistency—and developers try also to limit possibilities for making mistakes when writing data.

4. CMCC Metadata Agreement

The following section is devoted to the definition of the CMCC Metadata Agreement. A subset of the ISO 19115 entities has been chosen to model the CMCC data and service production.

Further entities of the ISO 19115 standard for geographical information will be adopted according to the CMCC needs.

4.1. UML notations

The diagrams that appear in this document are presented using the Unified Modelling Language (UML) static structure diagram with the ISO Interface Definition Language (IDL) basic type definitions and the UML Object Constraint Language (OCL) as the conceptual schema language. The UML notations used in this International Standard are described in the Figure 1.

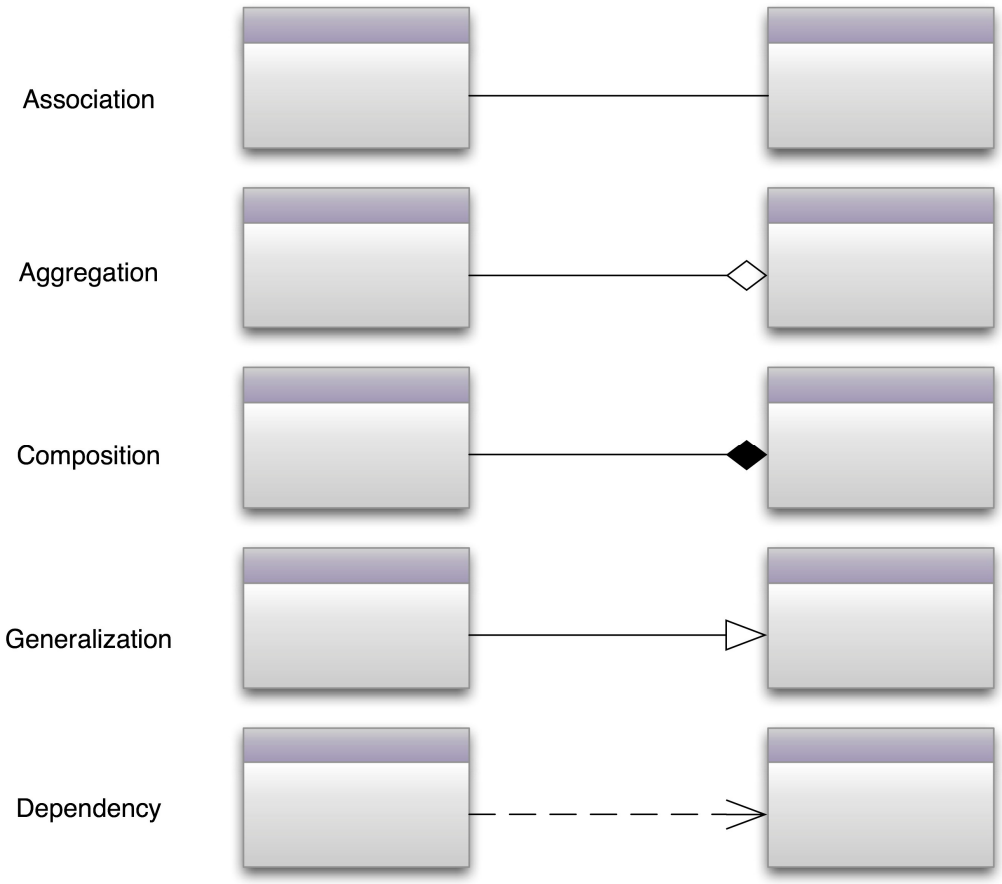


Figure 1: UML notation

4.2. Metadata entity set information

Figure 2 presents the core metadata entity: each metadata is identified by a file identifier, is written using a character set code, should concern datasets, models, series and so on (hierarchy level) and one or more responsibles are defined for the specified resource, with a given role (custodian, author, publisher, etc.).

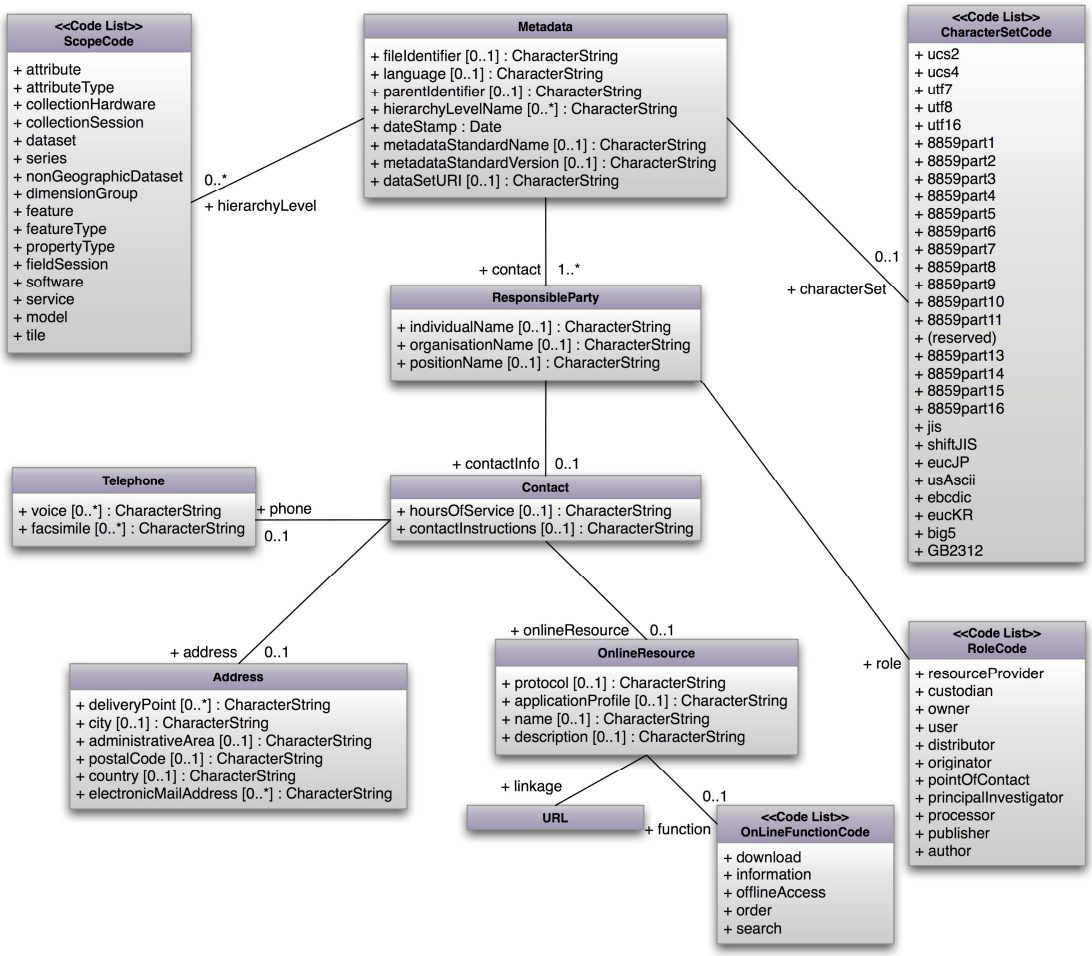


Figure 2: Metadata core entity

4.3. Metadata identification information

Figure 3 shows how to identify a resource: a compulsory abstract, purpose, credits, a status, one or more keywords, citations and point of contacts represent the basic set of information needed to mark resource catalogued by means of metadata.

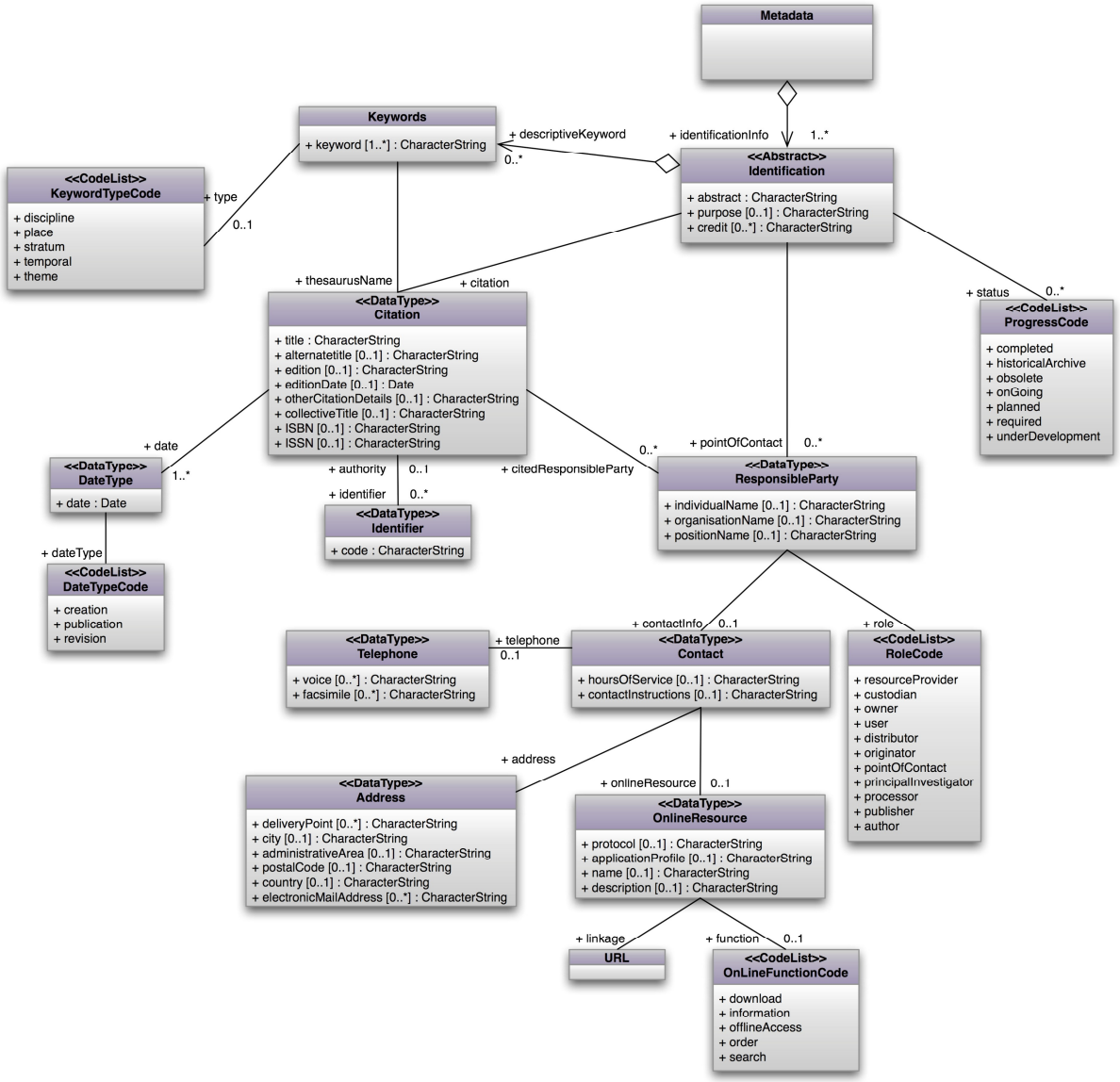


Figure 3: Metadata identification information

4.4. Metadata extent information

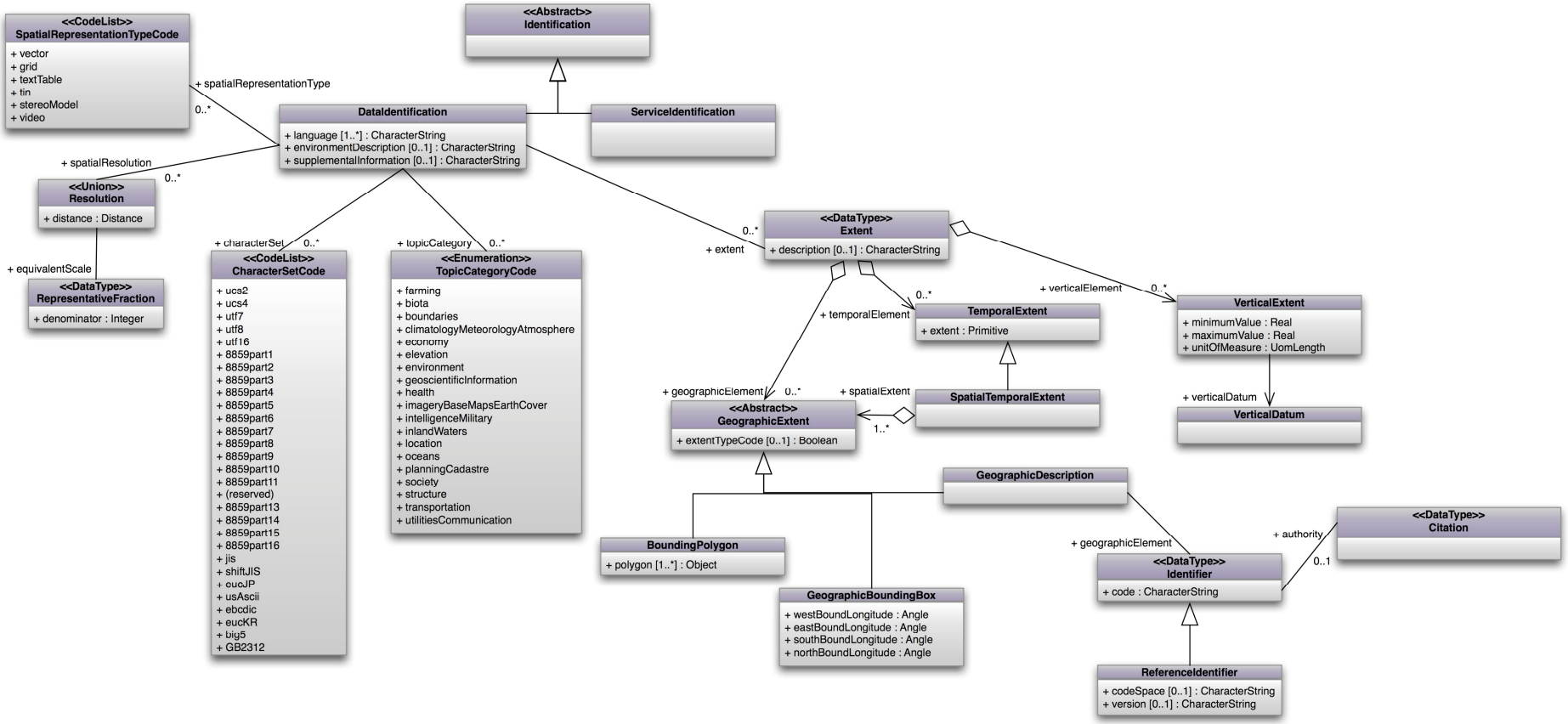


Figure 4: Metadata extent

By means of the generalization notation, the identification entity allows describing both services and data.

In case of data identification, an extent is given for the resource. The extent could be horizontal (bounding box or polygon, both inclusive and exclusive), vertical (altitude) and temporal.

5. Data dictionary

The aim of this section is to present the data dictionary, giving further detail on the entity previously described.

Each UML model class equates to a data dictionary entity. Each UML model class attribute equates to a data dictionary element. The shaded rows define entities. The entities and elements within the data dictionary are defined by six attributes (those attributes are listed below and are based on those specified in ISO/IEC 11179-3 for the description of data element concepts, i.e. data elements without representation). The term "dataset" when used as part of a definition is synonymous with all types of geographic data resources (aggregations of datasets, individual features and the various classes that compose a feature).

5.1. Name/role name

A label assigned to a metadata entity or to a metadata element. Metadata entity names start with an upper case letter. Spaces do not appear in a metadata entity name. Instead, multiple words are concatenated, with each new subword starting with a capital letter (example: XnnnYmmm). Metadata entity names are unique within the entire data dictionary of this metadata schema. Metadata element names are unique within a metadata entity, not the entire data dictionary. Metadata element names are made unique, within an application, by the combination of the metadata entity and metadata element names.

5.2. Definition

The metadata entity/element description.

5.3. Obligation/Condition

This is a descriptor indicating whether a metadata entity or metadata element shall always be documented in the metadata or sometimes be documented (i.e. contains value(s)). This descriptor may have the following values: M (mandatory), C (conditional), or O (optional).

- **Mandatory (M):** The metadata entity or metadata element shall be documented.
- **Conditional (C):** Specifies an electronically manageable condition under which at least one metadata entity or a metadata element is mandatory. 'Conditional' is used for one of the three following possibilities:
    1. Expressing a choice between two or more options. At least one option is mandatory and must be documented.
    2. Documenting a metadata entity or a metadata element if another element has been documented.
    3. Documenting a metadata element if a specific value for another metadata element has been documented.
- **Optional (O):** The metadata entity or the metadata element may be documented or may not be documented. Optional metadata entities and optional metadata elements have been defined to provide a guide to those looking to fully document their data. (Use of this common set of defined elements will help promote interoperability among geographic data users and producers world-wide.) If an optional entity is not used, the elements contained within that entity (including mandatory elements) will also not be used. Optional entities may have mandatory elements; those elements only become mandatory if the optional entity is used.

5.4. Maximum occurrence

Specifies the maximum number of instances the metadata entity or the metadata element may have. Single occurrences are shown by "1"; repeating occurrences are represented by "N". Fixed number occurrences other than one are allowed, and will be represented by the corresponding number (i.e. "2", "3"…etc).

5.5. Data type

Specifies a set of distinct values for representing the metadata elements; for example, integer, real, string, DateTime, and Boolean. The data type attribute is also used to define metadata entities, stereotypes, and metadata associations.

5.6. Domain

For an entity, the domain indicates the line numbers covered by that entity.

For a metadata element, the domain specifies the values allowed or the use of free text. “Free text” indicates that no restrictions are placed on the content of the field. Integer-based codes shall be used to represent values for domains containing codelists.

5.7. Metadata entities

A thorough overview of the metadata entities and attributes is given, following the outline detailed in the previous sections.

	Name / Role Name	Definition	Obligation / Condition	Max Occurrence	Data Type	Domain
1	Metadata	Root entity which defines metadata about a resource or resources	M	1	Class	Lines 2 - 13
2	fileIdentifier	Unique identifier for this metadata file	O	1	CharacterString	Free Text
3	language	Language used for documenting metadata	C / not defined by encoding?	1	CharacterString	ISO 639-2
4	characterSet	Full name of the character coding standard used for the metadata set	C / ISO/IEC 10646-1 not used and not defined by encoding?	1	Class	CharacterSetCode
5	parentIdentifier	File identifier of the metadata to which this metadata is a subset (child)	C / hierarchyLevel is not equal to "dataset"?	1	CharacterString	Free Text
6	hierarchyLevel	Scope to which the metadata applies	C / hierarchyLevel is not equal to "dataset"?	N	Class	ScopeCode
7	hierarchyLevelName	Name of the hierarchy levels for which the metadata is provided	C / hierarchyLevel is not equal to "dataset"?	N	CharacterString	Free Text
8	contact	Party responsible for the metadata information	M	N	Class	ResponsibleParty (80)
9	dateStamp	Date that the metadata was created	M	1	Class	Date (ISO/TS 19103)
10	metadataStandardName	Name of the metadata standard (including profile name) used	O	1	CharacterString	Free Text
11	metadataStandardVersion	Version (profile) of the metadata standard used	O	1	CharacterString	Free Text
12	dataSetURI	Uniformed Resource Identifier (URI) of the dataset to which the metadata applies	O	1	CharacterString	Free Text
13	identificationInfo	Basic information about the resource(s) to which	M	N	Association	Identification (14)

		the metadata applies				
14	Identification	Basic information required to uniquely identify a resource or resources	Use obligation from referencing object	Use maximum occurrence from referencing object	Aggregated Class	Lines 15 – 21
15	citation	Citation data for the resource(s)	M	1	Class	Citation (66)
16	abstract	Brief narrative summary of the content of the resource(s)	M	1	CharacterString	Free Text
17	purpose	Summary of the intentions with which the resource(s) was developed	O	1	CharacterString	Free Text
18	credit	Recognition of those who contributed to the resource(s)	O	N	CharacterString	Free Text
19	status	Status of the resource(s)	O	N	Class	ProgressCode
20	pointOfContact	Identification of, and means of communication with, person(s) and organization(s) associated with the resource(s)	O	N	Class	ResponsibleParty (80)
21	descriptiveKeywords	Provides category keywords, their type, and reference source	O	N	Association	Keywords (32)
22	DataIdentification	Information required to identify a dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Specified Class	Lines 15-21 & 23-30
23	spatialRepresentationType	Method used to spatially represent geographic	O	N	Class	SpatialRepresentation TypeCode

		information				
24	spatialResolution	Factor which provides a general understanding of the density of spatial data in the dataset	O	N	Class	Resolution (38)
25	language	Language(s) used within the dataset	M	N	CharacterString	ISO 639-2
26	characterSet	Full name of the character coding standard used for the dataset	C / ISO/IEC 10646-1 not used?	N	Class	CharacterSetCode
27	topicCategory	Main theme(s) of the dataset	C / if hierarchyLevel equals "dataset"?	N	Class	TopicCategoryCode
28	environmentDescription	Description of the dataset in the producer's processing environment, including items such as the software, the computer operating system, file name, and the dataset size	O	1	CharacterString	Free Text
29	extent	Extent information including the bounding box, bounding polygon, vertical, and temporal extent of the dataset	C / if hierarchyLevel equals "dataset"? either extent.geographicElement.GeographicBoundingBox or extent.geographicElement.GeographicDescription is required	N	Class	Extent (41)
30	supplementalInformation	Any other descriptive information about the dataset	O	1	CharacterString	Free Text
31	ServiceIdentification	Identification of capabilities which a service provider makes available to a service user through a set of interfaces that define a behaviour - See ISO 19119 for further information	Use obligation from referencing object	Use maximum occurrence from referencing object	Specified Class	Lines 15-21

32	Keywords	Keywords, their type and reference source	Use obligation from referencing object	Use maximum occurrence from referencing object	Aggregated Class	Lines 33-35
33	keyword	Commonly used word(s) or formalised word(s) or phrase(s) used to describe the subject	M	N	Character String	Free Text
34	type	Subject matter used to group similar keywords	O	1	Class	KeywordTypeCode
35	thesaurusName	Name of the formally registered thesaurus or a similar authoritative source of keywords	O	1	Class	Citation (66)
36	RepresentativeFraction	Serived from ISO 19103 Scale where RepresentativeFraction. denominator = 1 / Scale.measure and Scale.targetUnits = Scale.sourceUnits	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Line 37
37	denominator	the number below the line in a vulgar fraction	M	1	Integer	Integer > 0
38	Resolution	Level of detail expressed as a scale factor or a ground distance	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 39-40
39	equivalentScale	Level of detail expressed as the scale of a comparable hardcopy map or chart	C / distance not documented?	1	Class	RepresentativeFraction (36)
40	distance	Ground sample distance	C / equivalentScale not documented?	1	Class	Distance (ISO/TS 19103)

41	Extent	Information about horizontal, vertical, and temporal extent	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 42-45
42	description	Spatial and temporal extent for the referring object	C / geographicElement and temporalElement and verticalElement not documented?	1	CharacterString	Free Text
43	geographicElement	Provides geographic component of the extent of the referring object	C / description and temporalElement and verticalElement not documented?	N	Association	GeographicExtent (46)
44	temporalElement	Provides temporal component of the extent of the referring object	C / description and geographicElement and verticalElement not documented?	N	Association	TemporalExtent (57)
45	verticalElement	Provides vertical component of the extent of the referring object	C / description and geographicElement and temporalElement not documented?	N	Association	VerticalExtent (61)
46	GeographicExtent	Geographic area of the dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Aggregated Class	Line 47
47	extentTypeCode	Indication of whether the bounding polygon encompasses an area covered by the data or an area where data is not present	O	1	Boolean	0 – exclusion 1 – inclusion
48	BoundingPolygon	Boundary enclosing the dataset, expressed as the closed set of (x,y) coordinates of the polygon (last point replicates first point)	Use obligation from referencing object	Use maximum occurrence from referencing object	Specified Class	Lines 47 & 49
49	Polygon	Sets of points defining	M	N	Class	Object (ISO 19107)

		the bounding polygon				
50	GeographicBoundingBox	Geographic position of the dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Specified Class	Lines 47 & 51-54
51	westBoundLongitude	Western-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)	M	1	Class	Angle (ISO/TS 19103) -180,0 <= Value <= 180,0
52	eastBoundLongitude	Eastern-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)	M	1	Class	Angle (ISO/TS 19103) -180,0 <= Value <= 180,0
53	southBoundLatitude	Southern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north)	M	1	Class	Angle (ISO/TS 19103) -90,0 <= Value <= 90,0 Value <= North Value
54	northBoundLatitude	Northern-most, coordinate of the limit of the dataset extent expressed in latitude in decimal degrees (positive north)	M	1	Class	Angle (ISO/TS 19103) -90,0 <= Value <= 90,0 Value >= South Value
55	GeographicDescription	Description of the geographic area using identifiers	Use obligation from referencing object	Use maximum occurrence from referencing object	Specified Class	Lines 47 & 56
56	geographicIdentifier	Identifier used to represent a geographic area	M	1	Class	Identifier (116)

57	TemporalExtent	Time period covered by the content of the dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Aggregated Class	Line 58
58	extent	Date and time for the content of the dataset	M	1	Class	Primitive (ISO 19108)
59	SpatialTemporalExtent	Extent with respect to date/time and spatial boundaries	Use obligation from referencing object	Use maximum occurrence from referencing object	Specified Class	Lines 58 & 60
60	spatialExtent	Spatial extent component of composite spatial and temporal extent	M	N	Association	GeographicExtent (46)
61	VerticalExtent	Vertical domain of dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Aggregated Class	Lines 62-65
62	minimumValue	Lowest vertical extent contained in the dataset	M	1	Real	Real
63	maximumValue	Highest vertical extent contained in the dataset	M	1	Real	Real
64	unitOfMeasure	Vertical units used for vertical extent information	M	1	Class	UoMLength (ISO/TS 19103)
65	verticalDatum	Provides information about the origin from which the maximum and minimum elevation values are measured	M	1	Association	VerticalDatum (ISO 19111)
66	Citation	Standardized resource reference	Use obligation from referencing object	Use maximum	Class	Lines 67 – 79

				occurrence from referencing object		
67	title	Name by which the cited resource is known	M	1	CharacterString	Free Text
68	alternateTitle	Short name or other language name by which the cited information is known	O	N	CharacterString	Free Text
69	date	Reference date for the cited resource	M	N	Class	DateType (99)
70	edition	Version of the cited resource	O	1	CharacterString	Free Text
71	editionDate	Date of the edition	O	1	Class	Date (ISO/TS 19103)
72	identifier	Value uniquely identifying an object within a namespace	O	N	Class	Identifier (116)
73	citedResponsibleParty	Name and position information for an individual or organization that is responsible for the resource	O	N	Class	ResponsibleParty (80)
74	presentationForm	Mode in which the resource is represented	O	N	Class	PresentationFormCode
75	series	Information about the series, or aggregate dataset, of which the dataset is a part	O	1	Class	Series (109)
76	otherCitationDetails	Other information required to complete the citation that is not recorded elsewhere	O	1	CharacterString	Free Text
77	collectiveTitle	Common title with holdings note	O	1	CharacterString	Free Text
78	ISBN	International Standard Book Number	O	1	CharacterString	Free Text

79	ISSN	International Standard Serial Number	O	1	CharacterString	Free Text
80	ResponsibleParty	Identification of, and means of communication with, person(s) and organizations associated with the dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 81- 85
81	individualName	Name of the responsible personsurname, given name, title separated by a delimiter	C / organisationName and positionName not documented?	1	CharacterString	Free Text
82	organisationName	Name of the responsible organization	C / individualName and positionName not documented?	1	CharacterString	Free Text
83	positionName	Role or position of the responsible person	C / individualName and organisationName not documented?	1	CharacterString	Free Text
84	contactInfo	Address of the responsible party	O	1	Class	Contact (93)
85	role	Function performed by the responsible party	M	1	Class	RoleCode
86	Address	Location of the responsible individual or organization	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 87- 92
87	deliveryPoint	Address line for the location	O	N	CharacterString	Free Text
88	city	City of the location	O	1	CharacterString	Free Text
89	administrativeArea	State, province of the location	O	1	CharacterString	Free Text
90	postalCode	ZIP or other postal code	O	1	CharacterString	Free Text
91	country	Country of the physical address	O	1	CharacterString	ISO 3166-3
92	electronicMailAddress	Address of the electronic mailbox of the responsible organization	O	N	CharacterString	Free Text

		or individual				
93	Contact	Information required to enable contact with the responsible person and/or organization	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 94- 98
94	phone	Telephone numbers at which the organization or individual may be contacted	O	1	Class	Telephone (113)
95	address	physical and email address at which the organization or individual may be contacted	O	1	Class	Address (86)
96	onlineResource	On-line information that can be used to contact the individual or organization	O	1	Class	OnlineResource (102)
97	hoursOfService	Time period (including time zone) when individuals can contact the organization or individual	O	1	CharacterString	Free Text
98	contactInstructions	Supplemental instructions on how or when to contact the individual or organization	O	1	CharacterString	Free Text
99	DateType	Reference date and event used to describe it	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 100-101
100	Date	Reference date for the cited resource	M	1	Class	Date (ISO/TS 19103)
101	dateType	Event used for reference	M	1	Class	DateTypeCode

		date				
102	OnlineResource	Information about on-line sources from which the dataset, specification, or community profile name and extended metadata elements can be obtained	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 103-108
103	linkage	Location (address) for on-line access using a Uniform Resource Locator address or similar addressing scheme	M	1	Class	URL
104	protocol	Connection protocol to be used	O	1	CharacterString	Free Text
105	applicationProfile	Name of an application profile that can be used with the online resource	O	1	CharacterString	Free Text
106	Name	Name of the online resource	O	1	CharacterString	Free Text
107	description	Detailed text description of what the online resource is/does	O	1	CharacterString	Free Text
108	function	Code for function performed by the online resource	O	1	Class	OnLineFunctionCode
109	Series	Information about the series, or aggregate dataset, to which a dataset belongs	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 110-112
110	Name	Name of the series, or aggregate dataset, of which the dataset is a part	O	1	CharacterString	Free Text
111	issueIdentification	Information identifying the issue of the series	O	1	CharacterString	Free Text

112	page	Details on which pages of the publication the article was published	O	1	CharacterString	Free Text
113	Telephone	Telephone numbers for contacting the responsible individual or organization	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 114-115
114	voice	Telephone number by which individuals can speak to the responsible organization or individual	O	N	CharacterString	Free Text
115	facsimile	Telephone number of a facsimile machine for the responsible organization or individual	O	N	CharacterString	Free Text
116	Identifier	Value uniquely identifying an object within a namespace	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 117-118
117	authority	Person or party responsible for maintenance of the namespace	O	1	Class	Citation (66)
118	code	Alphanumeric value identifying an instance in the namespace	M	1	CharacterString	Free Text
119	ReferenceIdentifier	Identifier used for reference systems	Use obligation from referencing object	Use maximum occurrence from referencing object	Class	Lines 117-118 & 120-121
120	codeSpace	Name or identifier of the	O	1	CharacterString	Free Text

		person or organization responsible for namespace				
121	version	Version identifier for the namespace	O	1	CharacterString	Free Text

5.8. Code lists and enumerations

The stereotype classes <<CodeList>> and <<Enumeration>> can be found below. These two stereotype classes do not contain “obligation / condition”, “maximum occurrence”, “data type” and “domain” attributes.

These two stereotype classes also do not contain any “other” values as <<Enumeration>>s are closed (not extendable) and <<CodeList>>s are extendable.

Name	Domain code	Definition
DateTypeCode	-	Identification of when a given event occurred
creation	001	Date identifies when the resource was brought into existence
publication	002	Date identifies when the resource was issued
revision	003	Date identifies when the resource was examined or re-examined and improved or amended

OnLineFunctionCode	-	Function performed by the resource
download	001	Online instructions for transferring data from one storage device or system to another
information	002	Online information about the resource
offlineAccess	003	Online instructions for requesting the resource from the provider
order	004	Online order process for obtaining the resource
search	005	Online search interface for seeking out information about the resource

PresentationFormCode	-	Mode in which the data is represented
documentDigital	001	Digital representation of a primarily textual item (can contain illustrations also)
documentHardcopy	002	Representation of a primarily textual item (can contain illustrations also) on paper, photographic material, or other media
imageDigital	003	Likeness of natural or man-made features, objects, and activities acquired through the sensing of visual or any other segment of the electromagnetic spectrum by sensors, such as thermal infrared, and high resolution radar and stored in digital format
imageHardcopy	004	Likeness of natural or man-made features, objects, and activities acquired through the sensing of visual or any other segment of the electromagnetic spectrum by sensors, such as thermal infrared, and high resolution radar and reproduced on paper, photographic material, or other media for use directly by the human user
mapDigital	005	Map represented in raster or vector form
mapHardcopy	006	Map printed on paper, photographic material, or other media for use directly by the human user
modelDigital	007	Multi-dimensional digital representation of a feature, process, etc.
modelHardcopy	008	3-dimensional, physical model
profileDigital	009	Vertical cross-section in digital form

profileHardcopy	010	Vertical cross-section printed on paper, etc.
tableDigital	011	Digital representation of facts or figures systematically displayed, especially in columns
tableHardcopy	012	Representation of facts or figures systematically displayed, especially in columns, printed on paper, photographic material, or other media
videoDigital	013	Digital video recording
videoHardcopy	014	Video recording on film

RoleCode	-	Function performed by the responsible party
resourceProvider	001	party that supplies the resource
custodian	002	party that accepts accountability and responsibility for the data and ensures appropriate care and maintenance of the resource
owner	003	party that owns the resource
user	004	party who uses the resource
distributor	005	party who distributes the resource
originator	006	party who created the resource
pointOfContact	007	party who can be contacted for acquiring knowledge about or acquisition of the resource
principalInvestigator	008	key party responsible for gathering information and conducting research
processor	009	party who has processed the data in a manner such that the resource has been modified
publisher	010	party who published the resource
author	011	party who authored the resource

CharacterSetCode	-	name of the character coding standard used for the resource
ucs2	001	16-bit fixed size Universal Character Set, based on ISO/IEC 10646
ucs4	002	32-bit fixed size Universal Character Set, based on ISO/IEC 10646
utf7	003	7-bit variable size UCS Transfer Format, based on ISO/IEC 10646
utf8	004	8-bit variable size UCS Transfer Format, based on ISO/IEC 10646
utf16	005	16-bit variable size UCS Transfer Format, based on ISO/IEC 10646
8859part1	006	ISO/IEC 8859-1, Information technology – 8-bit single-byte coded graphic character sets – Part 1: Latin alphabet No. 1
8859part2	007	ISO/IEC 8859-2, Information technology – 8-bit single-byte coded graphic character sets – Part 2: Latin alphabet No. 2
8859part3	008	ISO/IEC 8859-3, Information technology – 8-bit single-byte coded graphic character sets – Part 3: Latin alphabet No. 3
8859part4	009	ISO/IEC 8859-4, Information technology – 8-bit single-byte coded

		graphic character sets – Part 4: Latin alphabet No. 4
8859part5	010	ISO/IEC 8859-51, Information technology – 8-bit single-byte coded graphic character sets – Part 5: Latin/Cyrillic alphabet
8859part6	011	ISO/IEC 8859-6, Information technology – 8-bit single-byte coded graphic character sets – Part 6: Latin/Arabic alphabet
8859part7	012	ISO/IEC 8859-7, Information technology – 8-bit single-byte coded graphic character sets – Part 7: Latin/Greek alphabet
8859part8	013	ISO/IEC 8859-8, Information technology – 8-bit single-byte coded graphic character sets – Part 8: Latin/Hebrew alphabet
8859part9	014	ISO/IEC8859-9, Information technology – 8-bit single-byte coded graphic character sets – Part 9: Latin alphabet No. 5
8859part10	015	ISO/IEC 8859-10, Information technology – 8-bit single-byte coded graphic character sets – Part 10: Latin alphabet No. 6
8859part11	016	ISO/IEC 8859-11, Information technology – 8-bit single-byte coded graphic character sets – Part 11: Latin/Thai alphabet
(reserved for future use)	017	A future ISO/IEC 8-bit single-byte coded graphic character set (e.g. possibly 8859 part 12)
8859part13	018	ISO/IEC 8859-13, Information technology – 8-bit single-byte coded graphic character sets – Part 13: Latin alphabet No. 7
8859part14	019	ISO/IEC 8859-14, Information technology – 8-bit single-byte coded graphic character sets – Part 14: Latin alphabet No. 8 (Celtic)
8859part15	020	ISO/IEC 8859-15, Information technology – 8-bit single-byte coded graphic character sets – Part 15: Latin alphabet No. 9
8859part16	021	ISO/IEC 8859-16, Information technology – 8-bit single-byte coded graphic character sets – Part 16: Latin alphabet No. 10
jis	022	Japanese code set used for electronic transmission
shiftJIS	023	Japanese code set used on MS-DOS based machines
eucJP	024	Japanese code set used on UNIX based machines
usAscii	025	United states ASCII code set (ISO 646 US)
ebcdic	026	Ibm mainframe code set
eucKR	027	Korean code set
big5	028	Traditional Chinese code set used in Taiwan, Hong Kong of China and other areas

GB2312	029	Simplified Chinese code set

KeywordTypeCode	-	Methods used to group similar keywords
discipline	001	Keyword identifies a branch of instruction or specialized learning
place	002	Keyword identifies a location
stratum	003	Keyword identifies the layer(s) of any deposited substance
temporal	004	Keyword identifies a time period related to the dataset
theme	005	Keyword identifies a particular subject or topic

ProgressCode	-	Status of the dataset or progress of a review
completed	001	Production of the data has been completed
historicalArchive	002	Data has been stored in an offline storage facility
obsolete	003	Data is no longer relevant
onGoing	004	Data is continually being updated
planned	005	Fixed date has been established upon or by which the data will be created or updated
required	006	Data needs to be generated or updated
underDevelopment	007	Data is currently in the process of being created

ScopeCode	-	Class of information to which the referencing entity applies
attribute	001	Information applies to the attribute class
attributeType	002	Information applies to the characteristic of a feature
collectionHardware	003	Information applies to the collection hardware class
collectionSession	004	Information applies to the collection session
dataset	005	Information applies to the dataset
series	006	Information applies to the series
nonGeographicDataset	007	Information applies to non-geographic data
dimensionGroup	008	Information applies to a dimension group
feature	009	Information applies to a feature
featureType	010	Information applies to a feature type
propertyType	011	Information applies to a property type
fieldSession	012	Information applies to a field session
software	013	Information applies to a computer program or routine
service	014	Information applies to a capability which a service provider entity makes available to a service user entity through a set of interfaces that define a behaviour, such as a use case

model	015	Information applies to a copy or imitation of an existing or hypothetical object
tile	016	Information applies to a tile, a spatial subset of geographic data

SpatialRepresentationTypeCode	-	Method used to represent geographic information in the dataset
vector	001	Vector data is used to represent geographic data
grid	002	Grid data is used to represent geographic data
textTable	003	Textual or tabular data is used to represent geographic data
tin	004	Triangulated irregular network
stereoModel	005	Three-dimensional view formed by the intersecting homologous rays of an overlapping pair of images
video	006	Scene from a video recording

TopicCategoryCode	-	High-level geographic data thematic classification to assist in the grouping and search of available geographic data sets. Can be used to group keywords as well.
farming	001	Rearing of animals and/or cultivation of plants
biota	002	Flora and/or fauna in natural environment
boundaries	003	Legal land descriptions
climatologyMeteorologyAtmosphere	004	Processes and phenomena of the atmosphere
economy	005	Economic activities, conditions and employment
elevation	006	Height above or below sea level
environment	007	Environmental resources, protection and conservation
geoscientificInformation	008	Information pertaining to earth sciences
health	009	Health, health services, human ecology, and safety
imageryBaseMapsEarthCover	010	Base maps
intelligenceMilitary	011	Military bases, structures, activities
inlandWaters	012	Inland water features, drainage systems and their characteristics
location	013	Positional information and services
oceans	014	Features and characteristics of salt water bodies (excluding inland waters)
planningCadastre	015	Information used for appropriate actions for future use of the land
society	016	Characteristics of society and cultures
structure	017	Man-made construction
transportation	018	Means and aids for conveying persons and/or goods
utilitiesCommunication	019	Energy, water and waste systems and communications infrastructure and services
…	…	…

…	…	…
…	…	…

Bibliography

[1] ISO 19115:2003 – Geographic Information – Metadata

[2] ISO/TS 19139:2007 – Geographic Information – Metadata – XML Schema Implementation

[3] Dublin Core Metadata Initiative – http://dublincore.org

[4] Numerical Model Metadata - http://ncas-cms.nerc.ac.uk/NMM/

[5] Woolf et. al. (2005): "Climate Science Modelling Language: Standards-based markup for metocean data", 85th meeting of American Meteorological Society, San Diego, Jan 2005

[6] NetCDF Climate and Forecast Metadata Convention - http://cf-pcmdi.llnl.gov/documents/cf-conventions/1.1/cf-conventions.pdf

[7] D. C. A. Bultermann, Is It Time For a Moratorium on Metadata?, IEEE MultiMedia, Oct-Dec 2004

[8] Federal Geographic Data Committee (FGDC) Metadata - http://www.fgdc.gov/

[9] Woolf et al.: “Data Virtualisation in the NERC DataGrid”

[10] Lawrence et al.: “The NERC DataGrid Prototype”

[11] Deelman et al. (2004): “Grid-based metadata services” - Proceedings of the 16th International Conference on Scientific and Statistical Database Management, 2004 – pp. 393-402

[12] Kindermann et al.: “C3Grid ISO Metadata Profile”