EVK2-CT-2002-00177

A pre-standardisation development for marine data interoperability using XML

MarineXML

D6 Standards Report

Catalogue

Title:	**Analysis of marine and XML data standards**
Creator:	Torill Hamre, Torill.Hamre@nersc.no
Creator:	Keiran Millard, k.millard@hrwallingford.co.uk
Creator:	Brian Matthews, B.M.Matthews@rl.ac.uk
Date Issued:	17/08/2004
Identifier:	MarineXML_D6_v5
Replaces:	MarineXML_D6_v4 (07/06/2004) MarineXML_D6_v3 (23/04/2004) MarineXML_D6_v2 (11/12/2003)
Approved:	Approved for release by Keiran Millard (07/06/04)
Accepted:	Accepted for release by Steering Committee at Oxford Progress Meeting subject to minor amendments to v4 (15/06/04)
Description of document:	This document identifies relevant standards, emanating form both from marine applications as well as from outside the marine domain. About 50 standards have been proposed for inclusion in the study and thereafter analysed to decide whether and how they should be used in the MarineXML development. Of these standards, many were international *de jure* standards like S-57, GML (Geographic Mark-Up Language) and GeoTIFF, while other standards have smaller user groups but were included because they covered an important marine application domain or general concept that could be useful in multiple domains.
Statement of use	The Copyright of this document is the property of MarineXML. This document may be reproduced in whole or in part (excluding logos) free of charge in any format or medium provided that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as '© MarineXML' with the title and source of the publication specified. The views and facts expressed in this document are those of the MarineXML consortium. Its members and affiliates offer no warranty as to the accuracy of this document and accept no liability whatsoever for any loss or damage arising from the interpretation or use of the information or reliance on views contained herein.

Contents

Catalogue **2**
Contents **3**
Executive Summary **4**
1 Introduction **5**
2 Overview of selected standards **6**
2.1 CATEGORISATION OF STANDARDS 6
2.2 INITIATIVES ADDRESSING MARINE DATA AND/OR METADATA STANDARDS 8
2.2.1 ICES-IOC SGXML (Study Group on XML) 8
2.2.2 AODC's MarineXML 8
2.2.3 Keeley Bricks 9
2.2.4 MIML (Maritime Information Markup Language) 10
2.2.5 INSPIRE (Infrastructure for Spatial Information in Europe) 10
2.2.6 NERC DataGrid 11
2.2.7 MBARI (Monterey Bay Aquarium Research Institute) 13
2.3 PROJECTS, PROGRAMMES AND SYSTEMS 14
2.4 SPATIAL DATA STANDARDS 16
2.4.1 Raster and grid data standards 16
2.4.2 Vector data standards 18
2.4.3 Volume data standards 20
2.5 METADATA STANDARDS 20
2.6 PRESENTATION STANDARDS 25
2.7 OTHER STANDARDS 26
3 Outline of inter-standard relationships model **29**
3.1 ONTOLOGY FACETS 29
3.2 ONTOLOGY IMPLEMENTATION 29
4 Conclusions and recommendations **34**
Acknowledgements **35**
MARINEXML MEMBERS 35
MARINEXML AFFILITATES 35
App. A Synopsis of selected standards **36**
A.1 SPATIAL DATA STANDARDS 36
A.2 PRESENTATION STANDARDS 43
A.3 OTHER STANDARDS 43

Executive Summary

MarineXML (A pre-standardisation development for marine data interoperability using XML) is an Accompanying Measures project part-funded by the European Commission. The project aims to demonstrate that eXtensible Mark-Up Language (XML) technology can be used to develop a framework that improves the interoperability of data for the marine community and specifically in support of marine observing systems.

The first phase of the MarineXML project has focussed on identifying relevant standards, both from marine applications as well as from outside the marine domain. All together, about 50 standards have been proposed for inclusion in the study and thereafter analysed to decide whether and how they should be used in the MML development. Of these standards, many were *de jure* standards like S-57, GML (Geographic Mark-Up Language) NetCDF and GeoTIFF, while other standards have smaller user groups but were included because they covered an important marine application domain or general concept that could be useful in multiple domains.

The purpose of this study is to provide a neutral view of this 'standards landscape' to answer questions such as; 'who is doing what with which standards?' – particularly if certain trends are emerging. To map this 'landscape', an ontology for inter-standard relationships has been developed using the Protégé tool developed at Stanford University. This ontology is a separate deliverable, but is outlined in this report.

One of the key trends is the harmonisation and adoption of OGC standards within the ISO TC211 (ISO19000) series of standards. The adoption of GML as ISO 19136 means potentially that MarineXML no longer has to make a choice between the ISO or OGC path for long term standardisation. In addition, the commitment of IHO to release version 4 of S-57 along ISO19000 standards means not only will it become possible to reuse S-57 features for other purposes beyond navigation. A registry containing the IHO S-57 Feature Type Catalogue could also be used to register other, non-navigational, marine features.

It is recognised that the standards landscape is continually evolving and as such this ontology within the resources of this project can never be 100% complete, accurate or finalised. Accordingly methods are being investigated as part of the post-project exploitation to serve this standard's ontology to wider marine community; in particular responsibilities for its update. In its present state the ontology is served in simple HTML format, together with a download of the Protégé files.

1 Introduction

MarineXML (A pre-standardisation development for marine data interoperability using XML) is a 2-year Accompanying Measures project funded by the EESD Programme under the Fifth Framework Programme of the European Commission. The aim of MarineXML is to demonstrate that eXtensible Mark-Up Language (XML) technology can be used to develop a framework that improves the interoperability of data for the marine community and specifically in support of marine observing systems.

MarineXML will develop a working prototype 'Marine Mark-Up Language' (MML) that will be used to facilitate data interoperability in a set of testbeds in different marine application domains that use diverse and heterogeneous types of data. The MML specification will be built around a marine data ontological framework defining what marine data entities exist and their relationships to other entities, as well as how different marine data entities may co-exist from the perspective of their interoperability. Crucially it will consider a governance model that needs to be established to take the standardisation process forward post-project.

MarineXML draws upon the results of previous and ongoing projects and standardisation initiatives addressing marine data and metadata. In particular it is anticipating to ensure the standardisation of the prototype MML specification by liasing with the IOC/IODE SGXML (Study Group on XML). In addition any MML specification will also have to be interoperable with widely used standards from organisations like IHO (International Hydrographic Organisation) and OGC (OpenGIS Consortium), as well as standards developed by e.g. instrument vendors providing measurements in (near) real time for monitoring and forecasting of the marine environment. The ultimate goal of a fully developed MML is to have a framework for seamless access to diverse and distributed marine data, which originally may reside in a variety of formats and on highly different platforms. By means of XML technologies and tools, data and associated metadata will be delivered to online services in their preferred format regardless of the original format of the data.

The first phase of the MarineXML project has focussed on identifying relevant standards, both from marine applications as well as from outside the marine domain. All together, about 50 standards have been proposed for inclusion in the study and thereafter analysed to decide whether and how they should be used in the MML development. Of these standards, many were de facto standards like S-57, GML (Geographic Mark-Up Language) and GeoTIFF, while other standards have smaller user groups but were included because they covered an important marine application domain or general concept that could be useful in multiple domains. The standards being investigated in this study are described in general in the next section, and in more detail in Appendix B.

The purpose of this study is to provide a neutral view of this 'standards landscape' to answer questions such as; 'who is doing what with which standards?' – particularly if certain trends are emerging. To map this 'landscape', an ontology for inter-standard relationships has been developed using the Protégé tool developed at Stanford University. This ontology is a separate deliverable, but is outlined in this report. A figure showing the entities in the data model is shown in Appendix A.

2 Overview of selected standards

The review process has focused on two main groups of metadata and data standards:

1. Standards used in the marine community
2. Standards based on XML

The aim has been to cover a representative set of data and metadata standards used in marine applications and in other domains, when the latter standards also has a potential for representing marine data and metadata. The 'Marine Mark-Up Language' (MML) to be developed should unify these standards by providing mechanisms for data interoperability between the different standards. The main criteria for inclusion of a standard in the review have been that the standard should either cover an important marine data type or application domain, and/or offer a general concept that can be used as a building block when developing the MML specification.

2.1 Categorisation of standards

Some 50 standards have been considered during the review. The majority of these have been included in the analysis and development of an inter-standard relationship model, after being categorised according to the main concepts shown in Figure 2.1. A geographic object is described by a (conceptual) metadata model, which in turn is represented by a metadata content model that is implemented by an implementation model in form of an encoding scheme (usually a file format). Further, the data of a geographic object is modelled by a (conceptual) spatial data model, which is represented by a spatial content model that is implemented by a data encoding scheme. Finally, a geographic object can be depicted by a conceptual presentation data model whose content model is encoded by a symbol encoding. Table 2.1 shows the categorisation of each reviewed standard.

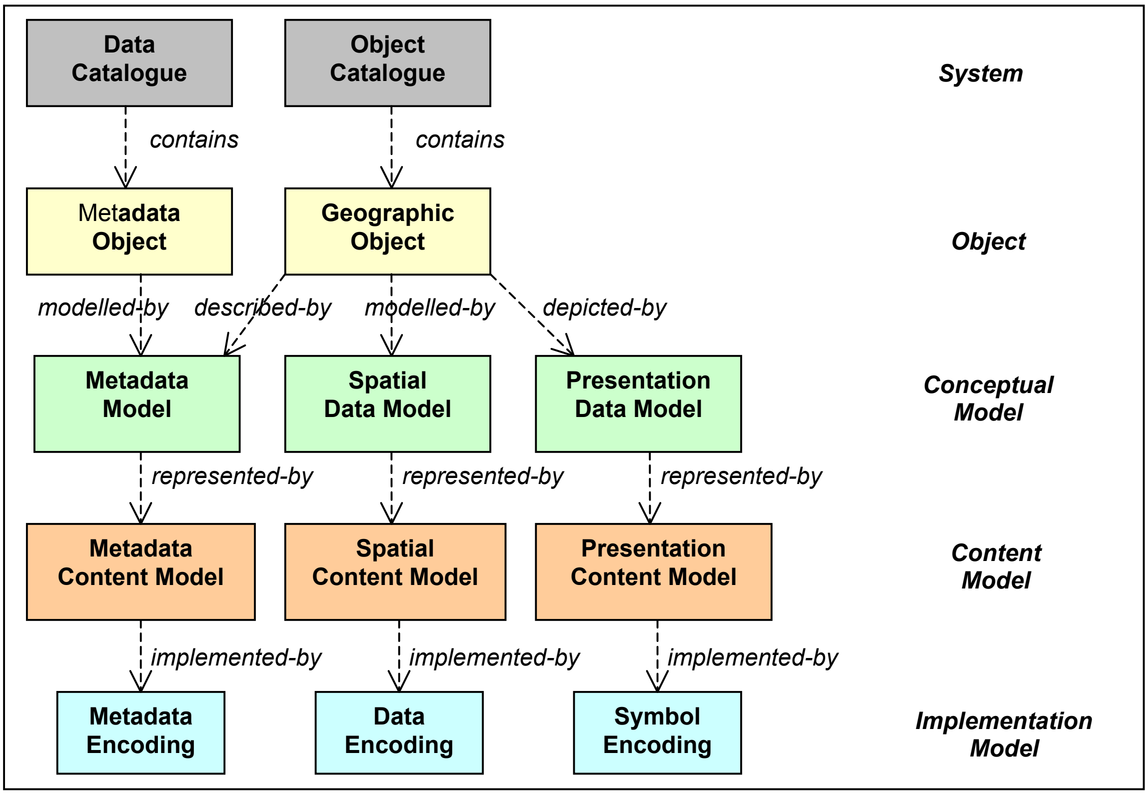

Figure 2.1 Different categories of standards for geographic data and metadata.

Table 2.1 Categorisation of standards.

Conceptual model	Content Model	Encoding Scheme
Raster:		
• Pixel	BSQ (Band SeQuential) BIL (Band Interleaved by Line) BIP (Band Interleaved by Pixel) SDTS Spatial Data Model BUFR Data Model GML Data Model ArcGIS Marine Data Model DIGEST Theoretical Model	TIFF, GeoTIFF. NetCDF, HDF GIF, CEOS (all binary) ISO 8211 Data Encoding of SDTS BUFR (binary) XML (Schema) to be decided Multiple (e.g. ISO 8211 and ISO 8824)
• Grid cell (matrix)	SDTS Spatial Data Model BUFR Data Model GRIB Data Model JMGRIB Data Model MMML Data Model GML Data Model ArcGIS Marine Data Model NetCDF Data Model HDF Data Model XSIL Data Model NDG Data Model DIGEST Theoretical Model	ISO 8211 Data Encoding of SDTS BUFR (binary) GRIB (binary) XML (DTD), data in GRIB for one case XML, data in HDF XML (Schema) to be decided NetCDF (binary) HDF (binary) XML (DTD) XML (Schema), data in GRIB Multiple (e.g. ISO 8211 and ISO 8824)
Vector:		
• Geometry (points, lines, polygons)	Shapefile Data Model BUFR Data Model TBEIC Data Model HydroML Data Model S57 Theoretical Data Model MIML Data Model SDTS Spatial Data Model GML Data Model Observations and Measurements MIML Data Model (for geography + hydrographic markup components) XSIL Data Model CML (Chemical Markup Language) NDG Data Model ChemicalUsageML Data Model WISTML Data Model NMEA 0183 & 2000 Data Model DIGEST Theoretical Model	Shapefile (binary) BUFR (binary) XML (Schema) XML (Schema) S57 (ISO/IEC 8211) XML (DTD or Schema) ISO 8211 Data Encoding of SDTS XML (Schema) XML (Schema) XML (DTD/Schema to be defined) XML (DTD) XML (DTD + Schema) XML (Schema), data in GRIB XML (Schema) XML (Schema) Proprietary format (binary/ASCII) Multiple (e.g. ISO 8211 and ISO 8824)
• Topology (nodes, networks)	S57 Theoretical Data Model MIML Data Model SDTS Spatial Data Model GML Data Model Observations and Measurements DIGEST Theoretical Model	S57 (ISO/IEC 8211) XML (DTD or Schema) ISO 8211 Data Encoding of SDTS XML (Schema) XML (Schema) Multiple (e.g. ISO 8211 and ISO 8824
• ISO feature (object)	ArcGIS Marine Data Model HydroML Data Model S57 Theoretical Data Model MIML Data Model SDTS Spatial Data Model GML Data Model Observations and Measurements	to be decided XML (Schema) S57 (ISO/IEC 8211) XML (DTD or Schema) ISO 8211 Data Encoding of SDTS XML (Schema) XML (Schema)

	TBEIC Data Model MIML Data Model (for geography + hydrographic markup components) XMML Data Model MarineXML (of AODC) NDG Data Model	XML (Schema) XML (DTD/Schema to be defined) XML (Schema) XML (Schema) XML (Schema), data in GRIB
Volume:		
• Voxel		
• 3D cell	MMML Data Model NetCDF Data Model HDF Data Model	XML NetCDF (binary) HDF (binary)
Presentation:		
• Map symbol	S52 Data Model SVG Data Model GraphML Data Model SMIL Data Model	S52 XML (Schema) XML (DTD & Schema) XML (DTD + Schema)
Metadata:		
• Discovery & technical metadata	ISO 19115 Metadata Model OGC Catalogue Specification Dublin Core Metadata Model CSDGM Specification NDG Metadata Model EDIOS Model NcML IWICOS Metadata Specification Directory Interchange Format	XML (DTD/Schema to be defined) XML (DTD) XML (Schema) XML (DTD) XML (Schema) Database schema (SQL) XML (Shema) XML (Schema) DIF (ASCII)

2.2 Initiatives addressing marine data and/or metadata standards

2.2.1 ICES-IOC SGXML (Study Group on XML)

The SGXML (ICES-IOC Study Group on the Development of Marine Data Exchange Systems using XML) is investigating how XML technology can be used to make marine data management and exchange more efficient between the different data centers in IOC/IODE and which can serve as guidelines for the marine community in general. At present, SGCML is focusing on three areas: development of parameter dictionaries, representation of point data, and metadata. For parameter dictionaries, both a DTD and an XML Schema have been defined, and "Keeley Bricks" (see section 2.2.3) have been used for profile data (including taxonomy). For metadata, SGXML will evaluate and compare established standards (such as ISO 19115 and EDMED) and make recommendations for an optimal metadata definition for marine data.

Ref: http://www.marinexml.net/

2.2.2 AODC's MarineXML

MarineXML is developed by the Australian Oceanographic Data Centre (AODC). It is the internal format used as the basis for all management of marine data at AODC, and represents both metadata (e.g. custodian, quality, source (sensor), etc.) and data (parameters). Incoming data are converted to MarineXML, and stored in a marine database. The MEDI (Environmental Data Information Referral Catalogue) system then extracts the metadata from the MarineXML file and creates a metadata record that complies with NASA's GCMD (Global Change Master Directory) format and most fields of the Australian ANZLIC metadata standard (see section 2.5). The structure of MarineXML is defined in an XML Schema. The main elements of the AODC MarineXML schema are shown in Figure 2.1.

Ref: http://www.aodc.gov.au/products/prod/marinexml.html
http://www.aodc.gov.au/products/prod/documentation/marine_xml_schema.html

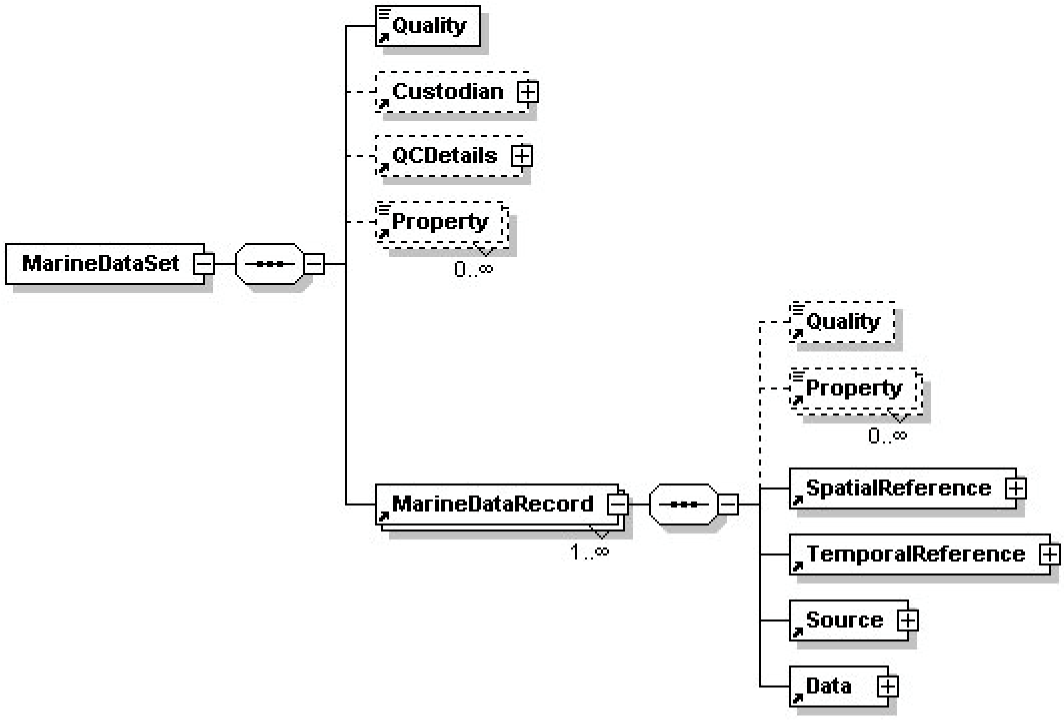

Figure 2.1 Excerpt of the main elements in AODC's MarineXML standard.

2.2.3 Keeley Bricks

Keeley Bricks are being developed as basic building blocks for representation of marine data and metadata in a Marine XML. The concept is that a Marine XML file will be built up by assembling those bricks necessary to represent a particular dataset as well as any pertinent metadata associated with it. Metadata includes, among others, a description of the instruments used to collect the data, the cruise, data quality, and of the availability (e.g. restrictions on use). Data that can be represented is currently limited to profiles. Three Canadian laboratories, which have also participated in the ICES-IOC SGXML, have developed the Keeley Bricks. Figure 2.2 shows some of the main elements of the Keeley Bricks XML schema. More information on Keeley Bricks are found in e.g. Keeley et al., 2003.

Ref: http://ioc.unesco.org/marinexml/contents.php?id=19
http://ioc.unesco.org/marinexml/files.php?action=viewfile&fid=4&fcat_id=3
Keeley, Robert, Anthony Isenor and Joe Linguanti, XML Bricks, 20 Jan 2003.

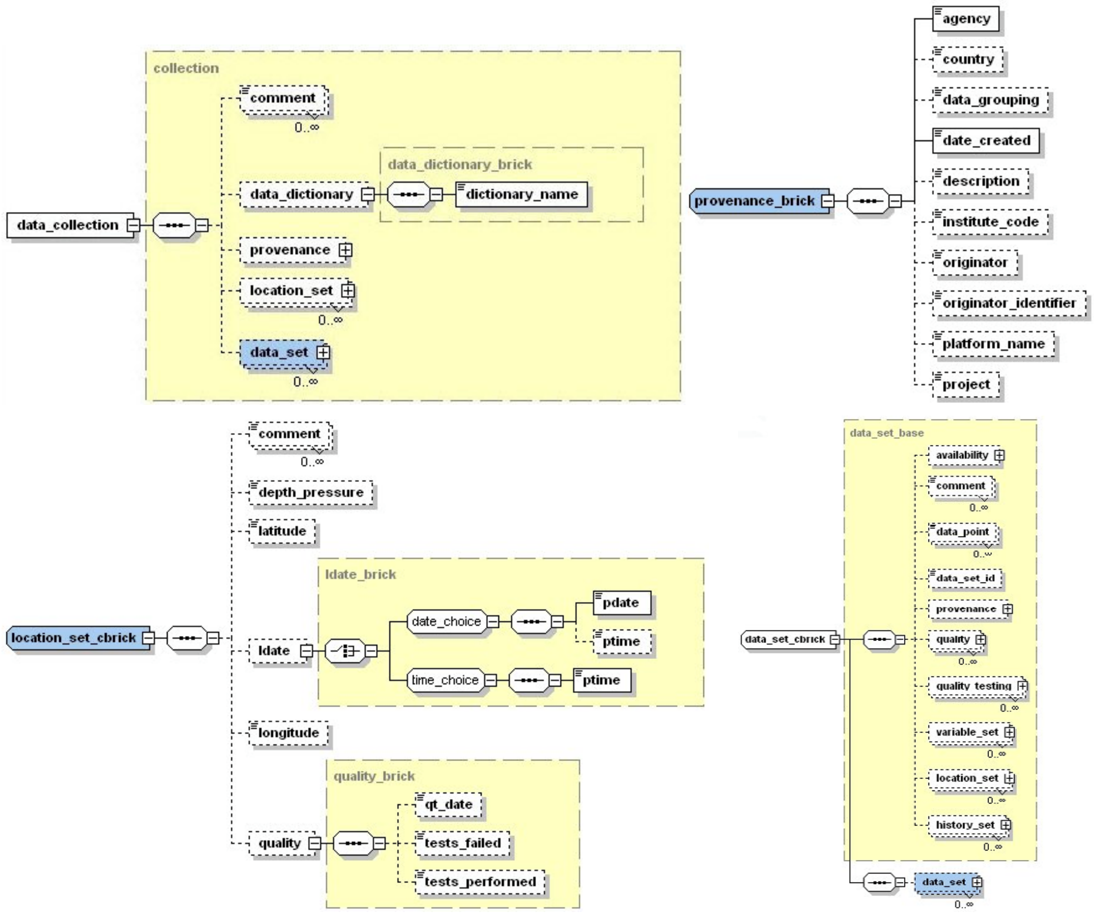

Figure 2.2 Some of the defined Keeley Bricks. (The element data_set in lower right figure is of type data_set_cbrick.)

2.2.4 MIML (Maritime Information Markup Language)

MIML (Maritime Information Mark-Up Language) is an USCoastguard standard for marine information, which is currently under development. MIML is developed for the Waterways Information Network (WIN) of the Coast Guard Research and Development Center (RDC). Its aim if to use XML to describe the format of data transfer mechanisms and the structure of the different types of data available in WIN. MIML models are categorised at three levels. At the first (minimum) level, XML is used to tag the data. A second level model will incorporate the first level (i.e. an XML document) accompanied by a 'reference' document explaining the (XML) structure and providing examples of valid use. The third level model would be a formal representation of the data model (and data transfer model) using e.g. an ER (Entity-Relationship), UML (Unified Modelling Technique) or ontology model (e.g. by means of the Protégé tool).

Ref: "MIML Development Version 0.3 – June 25, 2003" (Available online at http://www.rdc.uscg.gov/iws/pubs/miml.pdf.)
www.eas.asu.edu/~gcss/papers/rmm-ion2002.pdf (paper by R. Malyankar, Arizona State Univ.)

2.2.5 INSPIRE (Infrastructure for Spatial Information in Europe)

INSPIRE (Infrastructure for Spatial Information in Europe) is an initiative aiming at making available relevant, harmonised and quality geographic information to support formulation, implementation, monitoring and evaluation of Community policies with a territorial dimension or impact. Its working group on architecture and standards concluded that "the INSPIRE profile and guidelines for the implementation shall be based on the ISO 19100 series of standards for geographic information, and where necessary and appropriate, results of others standardisation initiatives can be considered (e.g. Dublin Core, OGC)".

Further information and documents can be found on the INSPIRE web site: http://inspire.jrc.it/, including the INSPIRE Architecture and Standards Position Paper, which can be access at http://inspire.jrc.it/reports/position_papers/inspire_ast_pp_v4_2_en.pdf.

2.2.6 NERC DataGrid

The NERC DataGrid project is a joint project between BADC (British Atmospheric Data Centre), BODC (British Oceanographic Data Centre) and CCLRC e-Science Centre. It aims at developing a joint framework for data discovery and data access with format transparency within the environmental community funded by NERC (Natural Environment Research Council, UK). Data and metadata will be searchable and retrievable through the Earth System Grid, an online portal based on standard web and Grid technologies.

To support data interoperability within the ESG, the NERC DataGrid project is developing a metadata and data model, termed the NDG Metadata Model and NDG Data Model, respectively. Both models will be compatible with relevant ISO 191xx standards, and alignment with ISO standards will be an ongoing process as these become available after final ratification. The NGD Metadata Model divides metadata into five categories: (A) use metadata, (B) core metadata, (C) ancillary metadata, (D) discovery metadata, and (E) 'extra' metadata. The metadata structure is illustrated in Figure 2.3 and is being implemented as an XML Schema. The NDG Data Model is developed in alignment with ISO 19101 Geographic Information – Reference model, with the dataset as its core concept. Figure 2.4 shows the main elements of the NDG Data Model. The data model will be implemented using standard data formats, including self-describing XML as well as de facto standards like GRIB and NetCDF (CF convention).

Ref:http://ndg.badc.rl.ac.uk/index.htm http://www.e-science.clrc.ac.uk/web/projects/nercdatagrid

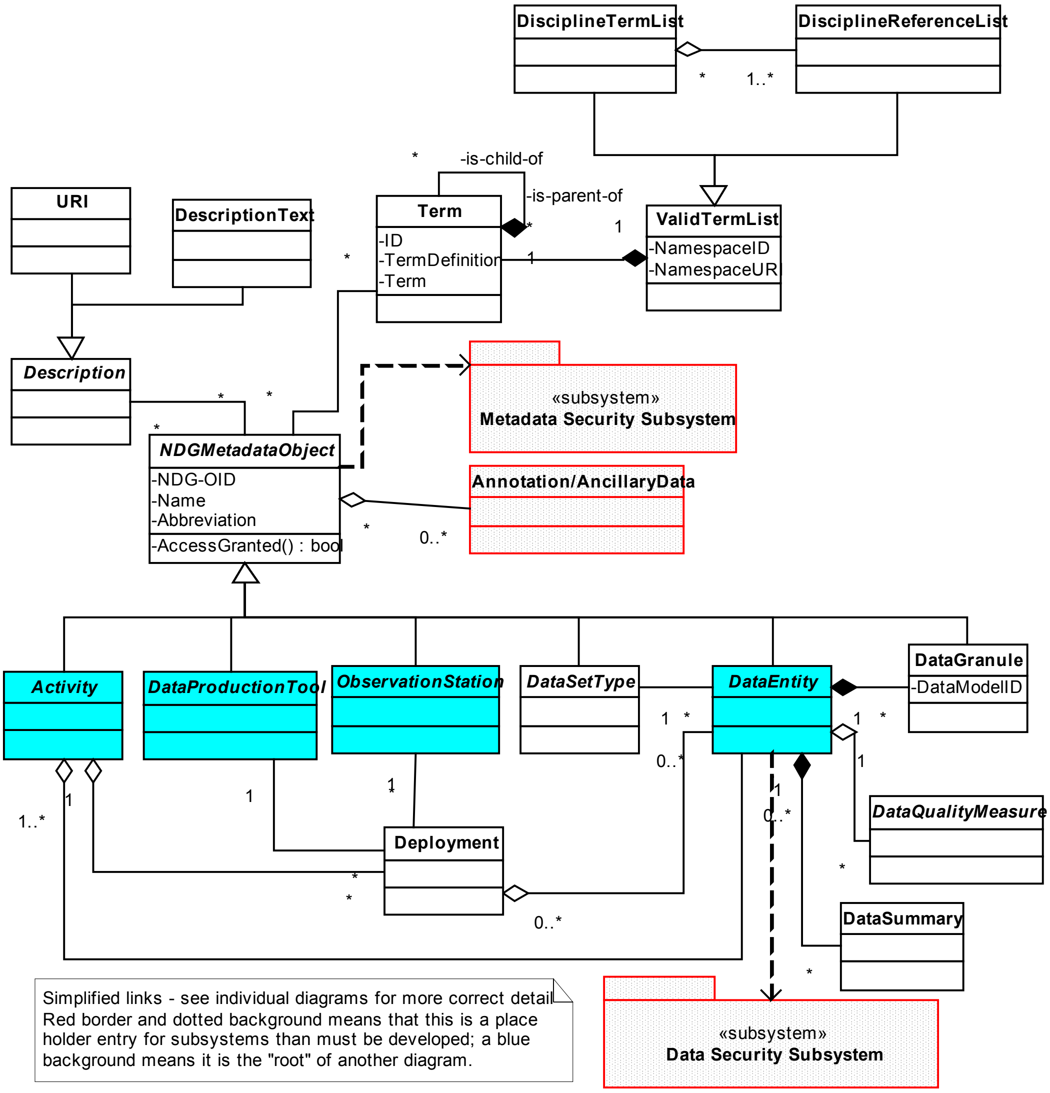

Figure 2.3 NDG Metadata Model.

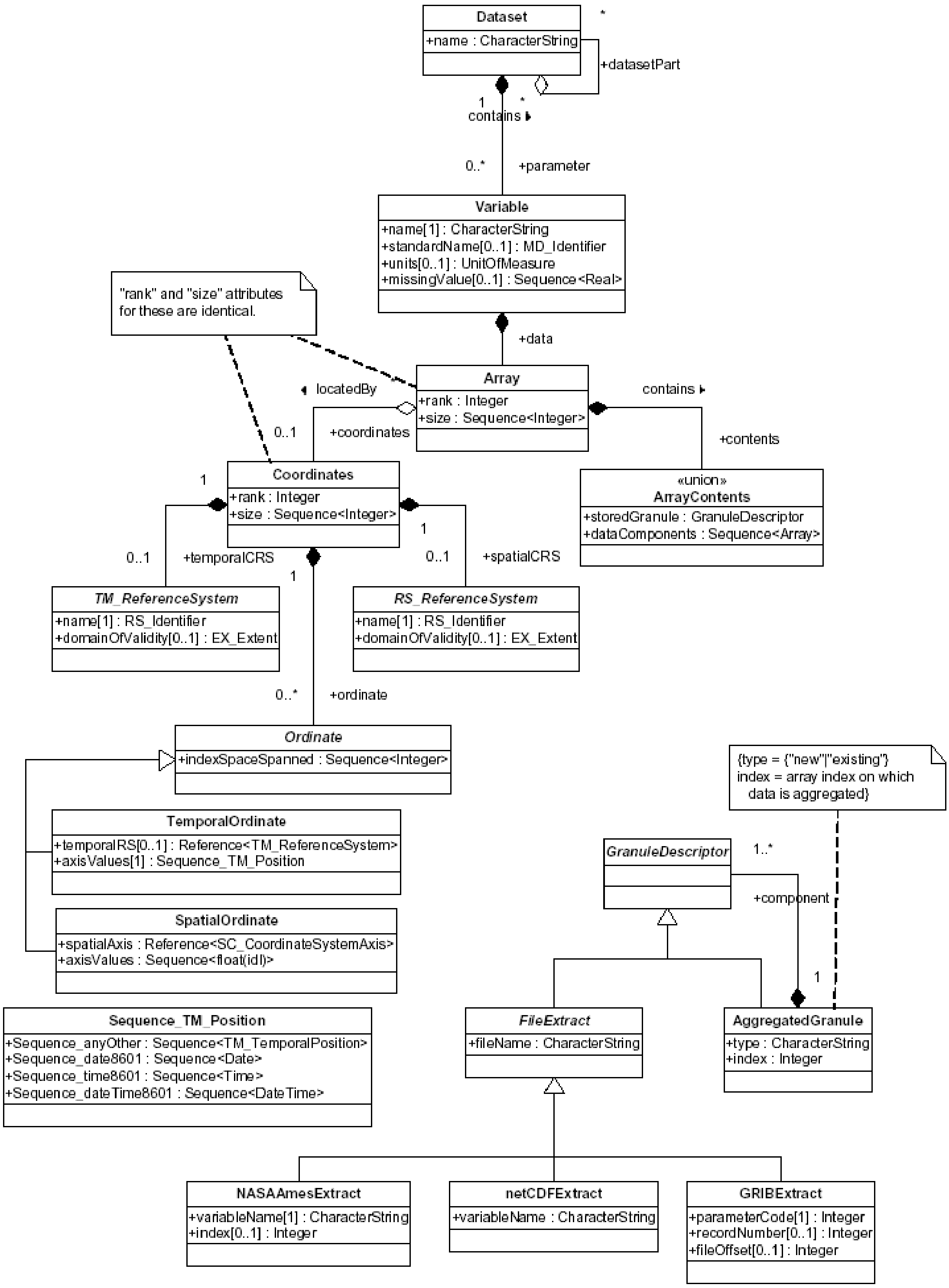

Figure 2.4 NDG Data Model.

2.2.7 *MBARI (Monterey Bay Aquarium Research Institute)*

The Monterey Bay Aquarium Research Institute (MBARI) has collected science data for 15 years from many oceanographic instruments and systems. The Monterey Ocean Observing System, or MOOS, presents new oceanographic data management challenges. To meet the data management requirements, MBARI is developing a flexible, extensible data management solution, the Shore

Side Data System (SSDS). This data management solution addresses the complete data life cycle, including instrument (and metadata) development, data ingest, archival, search and access, and visualisation and analysis. Working with MOOS infrastructure software, the SSDS can easily support new instruments, data streams, and data sets, from all types of instruments and platforms (for example, moorings, AUVs, and ships).

The MOOS uses XML definitions and schema, and provide user tools to simplify metadata entry. The SSDS takes advantage of, and provides support for, open standards where possible. Its developers monitor, and in some cases collaborate with, development of other data management components such as OPeNDAP[1], THREDDS[2], and LAS[3]. The functionality is compatible with an open source framework, including Linux, Apache, and MySQL, although it is not yet fully tested in that environment. In the data standards arena, SSDS developers have been seeking XML schema that address the oceanographic science mission MBARI represents. No schema matching SSDS needs have been identified. Because of its widespread oceanographic use, the netCDF[4] file format is a likely intermediate storage format. MBARI has started collaborating with multiple oceanographic science institutions to tailor its use of this standard.

Ref: http://www.mbari.org/ssds/ReferenceDocuments/MOOSMetadataSchema.xsd.txt

2.3 *Projects, programmes and systems*

The following projects, programmes and systems have been reviewed. (Some are specific to marine data, while others are not.)

MEDI (Marine Environmental Data Information Referral Catalogue) is the metadata management tool of the AODC (Australian Oceanographic Data Committee). The internal data format is AODC's MarineXML (described in section 2.2.2). MEDI is an inventory of information about marine related datasets and is used within the framework of the IOC's International Oceanographic Data and Information Exchange (IODE) system. MEDI uses the Directory Interchange Format (DIF) that has been developed by NASA's Global Change Master Directory (GCMD). The IODE programme distributes the MEDI metadata authoring tool which stores metadata records as DIF-XML files. XML is used to transfer MEDI-DIF and GCMD-DIF records.
Ref: http://ioc.unesco.org/medi
http://www.aodc.gov.au/products/prod/medi.html

ROSCOP provides a low level inventory for tracking oceanographic data collected on Research Vessels etc. Most marine disciplines are represented in ROSCOP, including physical, chemical, and biological oceanography, fisheries, marine contamination/pollution, and marine meteorology.

Ref: http://www.ices.dk/ocean/roscop/
http://www.meteo.ru/nodc/project/inventory/descrip2.htm

EDMED covers a wide range of disciplines and is a high level inventory, describing both Datasets and Data Holding Centres. At present, EDMED already describes more than 2814 Datasets from over 574 Data Holding Centres across Europe.

Ref: http://www.bodc.ac.uk/services/edmed/, http://www.sea-search.net/edmed/welcome.html

---

1) OPeNDAP: Open Source Project for a Network Data Access Protocol, <http://opendap.org/>
2) THREDDS: Thematic Realtime Environmental Data Distributed Services <http://www.unidata.ucar.edu/projects/THREDDS/>
3) LAS: Live Access Server, *NOAA*, <http://ferret.wrc.noaa.gov/Ferret/LAS/ferret_LAS.html>
4) netCDF: network Common Data Form, *Unidata/UCAR*, <http://www.unidata.ucar.edu/packages/netcdf/>

EDMERP is a European directory of research projects relating to the marine environment. It covers a wide range of disciplines including marine meteorology; physical, chemical and biological oceanography; sedimentology; marine biology and fisheries; environmental quality; coastal and estuarine studies; marine geology and geophysics etc.

Ref: http://www.sea-search.net/mrp/content.htm

EDIOS (European Directory of the Initial Ocean-observing System) is a 3-year RTD project funded by EESD under FP5. EDIOS has as main aim to "build a meta-database (computerised Directory) that includes information on all European ocean-observing sites/devices in routine and repeated operation (to be continuously updated) and to use this Directory to define the Initial European Ocean-observing System". EDIOS has defined a database structure that includes, among others, technical specifications of the data collection methods used, geo-referencing, specifications of the measurements with their spatio-temporal characteristics, and accuracy.

Ref: http://www.edios-project.de/
http://www.edios-project.de/Project%20Information/InformationIndex.html

MarLIN (The Marine Life Information Network for Britain & Ireland): The Marine Life Information Network (*MarLIN*) programme was established in 1998 by the Marine Biological Association with the aim of providing information for marine environmental management, protection and education. The programme was developed in collaboration with the major environmental protection agencies in the UK together with academic institutions and was designed to make information freely and rapidly available through the Internet. Researchers and other data collectors are encouraged to submit their data, which are quality controlled and inserted into a database that can be searched via Internet. Database can be searched by species or by locations (on a map). Data displayed shows species counts, abundance or presence absence as well as any physical data (e.g. depth or substratum where provided).

Ref: http://www.marlin.ac.uk/

OBIS (Ocean Biogeographic Information System) is a web-based provider of global geo-referenced information on marine species. The OBIS portal can be searched by species, by region, or by taxonomic group, plot distributions on maps, view data, and/or export data to on-line modelling packages or to your own applications. OBIS also provides links to data contributors and to many other sources of information about individual marine species, including a variety of educational and technical resources.

Metadata about marine species datasets are stored in the OBIS schema (an extension of Darwin Core, version 2), which is a list of data fields with names, descriptions, and format notes. When the OBIS portal sends queries out to its distributed data contributors, the portal will request data using these fields and needs to have data returned using these fields. OBIS uses DiGIR software (see below) to communicate with data providers, who must map their databases to the OBIS schema to enable searches via the portal.

Ref: http://www.iobis.org/

DiGIR (Distributed Generic Information Retrieval) is a protocol for retrieving structured data from multiple, heterogeneous databases across the Internet. The DiGIR protocols defines request and response message formats for communication between provider, portal engine, and user interfaces, specifically: (1) metadata requests, (2) search requests, and (3) inventory requests.

Design goals for the DiGIR project is to use open protocols and standards, such as HTTP and XML, decouple the protocol, software and semantics, make new data provider installations as easy as possible, and to have open source development and GNU General Public Licensing. Implementation is ongoing, and available from http://digir.sourceforge.net/

DoD XML Registry provides METOC XML components through the WWW. The DoD XML Registry constitutes guidance in the generation and use of XML among DoD (US Department of Defense) communities of interest and is the authoritative source for registered XML data and metadata components.

Ref: http://diides.ncr.disa.mil/xmlreg/user/index.cfm

The **DEAL Data Registry** facilitates access to data and information about offshore Oil & Gas Exploration & Production for the U.K. The DEAL data registry catalogues link to sources of the underlying data, such as well logs, seismic data and reports. Users can for instance send requests for copies of UK Department of Trade and Industry (DTI) 2D seismic data or requests to view DTI well cores. The system uses XML for data exchange, and uses XML Schemas for well header and well events information. Schemas are being developed in cooperation between BGS (British Geological Survey), CDA (Common Data Access, Ltd.), DTI, and POSC.

Ref: http://www.ukdeal.co.uk/

BIOMARE (Implementation and networking of large scale, long term MArine BIOdiversity research in Europe) is an FP5 EESD Concerted Action led by NIOO-CEMO Netherlands Institute of Ecology, with 20 other partners (and 7 more associated partners) from more than 20 European countries. The objectives of BIOMARE are to achieve a European consensus on the selection and implementation of: (1) a network of Reference Sites as the basis for long-term and large-scale marine biodiversity research in Europe, (2) internationally agreed standardised and normalised measures and indicators for (the degree of) biodiversity, and (3) facilities for capacity building, dissemination and networking of marine biodiversity research.

Ref: http://www.biomareweb.org/

CORES is a project funded by IST in FP5. The central objective of the CORES project is to encourage the sharing of metadata semantics. CORES will address the need to reach consensus on a data model for declaring semantics of metadata terms in a machine-readable way.

Ref: http://www.cores-eu.net/

2.4 Spatial data standards

The following spatial data standards have been reviewed. (Some are specific to marine data, while others are not.)

2.4.1 Raster and grid data standards

SDTS (Spatial Data Transfer Standard) is a standard for transfer of spatial digital data developed by the USGC (U.S. Geological Survey). Federal agencies in the U.S. must comply with SDTS, which is also widely used in other governmental organisations. Large commercial companies like ESRI, ERDAS and MapInfo have included STDS handling in their software. SDTS can hold both vector and raster data, with geometry and topology, as well as associated metadata (e.g. data quality). STDS is stored according to the ISO 8211 Data Encoding standard.

DIGEST Edition 2.1 is developed by the Digital Geographic Information Working Group (DGIWG) to support efficient exchange of digital geographic information between national agencies, data producers and end users. DIGEST (Digital Geographic Information Exchange Standard) is a comprehensive "family of standards" that can represent raster, matrix, and vector data (and associated text), and supports entire range of levels of topological structures. DIGEST can be encoded in multiple ways, e.g. using ISO 8211 and ISO 8824 standards.

NetCDF (network Common Data Form) is a machine-independent format for representing scientific data, which was developed by Unidata Program Center in Boulder, Colorado. NetCDF is array-oriented and is used for many different types of data, e.g. for satellite images and results of model predictions. The format is widely used in the scientific community around the world, and a number of software packages are available for display and manipulation. These are primarily public domain tools, but also some COTS tools exist (e.g. MATLAB and IDL). Metadata can be stored as NetCDF attributes, which are linked to NetCDF variables of a given dimension. NetCDF data are stored in binary files.

Different communities have defined their own conventions for data storage in NetCDF files. For instance, the NetCDF Climate and Forecast (CF) metadata convention is used by the climate research community. CF was designed to handle climate and forecast data, atmosphere, surface and ocean, model-generated data and comparable observational datasets. CF generalizes and extends the COARDS (Cooperative Ocean/Atmosphere Research Data Service) convetions, which were developed for global atmospheric and oceanographic research data sets.

HDF (Hierarchical Data Format) is a standard format for scientific data exchange used in a number of application domains, e.g. for transfer of satellite images. HDF is developed by NCSA (National Center for Supercomputer Applications), and like NetCDF it is a popular format in the scientific community world-wide. Version 5.1.6.2 is the most current version, while older versions 4.x are also widely used (and supported by NCSA). Metadata is stored as attributes in the defined objects (classes). HDF files are binary.

GeoTIFF represents an effort by over 160 different remote sensing, GIS, cartographic, and surveying related companies and organizations to establish a TIFF based interchange format for geo-referenced raster imagery. GeoTIFF 1.0 is the official release version. GeoTIFF implements the geographic metadata formally, using compliant TIFF tags and structures. GeoTIFF data are stored as TIFF (i.e. as binary files).

General-purpose image format include, among others, **TIFF** (Tagged Image File Format), **GIF** (Graphics Interchange Format), **JPEG** (Joint Photographic Experts Group) and **PNG** (Portable Network Graphics). These are also used for marine data, but primarily for presentation purposes.
Ref: http://home.earthlink.net/~ritter/tiff/ http://www.w3.org/Graphics/GIF/spec-gif87.txt http://256.com/gray/docs/gifspecs/ http://www.faqs.org/faqs/graphics/fileformats-faq/part3/section-57.html http://www.faqs.org/faqs/jpeg-faq/ http://www.libpng.org/pub/png/

CEOS (Committee on Earth Observation Satellites) is a standard format for satellite data from e.g. ESA PAFs (Processing Archive Facilities). The CEOS format is also used by other satellite data providers, and can be ingested in many image processing packages. Metadata often includes information on numerous satellite parameters, and is stored in the 'header' file of a CEOS data set, which is typically composed of multiple files. CEOS files are binary.

GRIB (GRidded In Binary, WMO FM92) is a WMO format for storage of weather product information from national weather forecasting centres and other operational organisations. Metadata such as geo-referencing information and units for parameters are included in the format. GRIB files are binary.

BUFR (WMO FM94) (Binary Universal Form for the Representation of Meteorological Data) is a WMO standard for encoding meteorological and other (primarily) observational data as a binary 'message', which can be efficiently transmitted and stored. Metadata (geo-referencing, units, compression, etc.) is stored as part of the BUFR message. BUFR files are binary, and all codes used for encoding observations must be predefined to enable correct decoding at the receiver.

ArcGIS Marine Data Model defines an object-oriented data model (constrained to, and implemented as, a geo-relational database) for a wide ranges of marine data and (physical) objects. For instance, point data, raster and gridded data, volume data and time series (of points, etc.). The model also includes information about the instruments that generated the observations, cruises, etc. Metadata includes, among others, geo-location and time stamping. The data model is developed in UML (Unified Modelling Language), which is transformed to an ArcGIS 8 database schema

GML 3 is described under Section 2.4.2.

2.4.2 Vector data standards

S57 3.1 is an IHO standard for the exchange of digital hydrographic data. Version 3.1 is the current version, which is endorsed by all major commercial ECDIS (Electronic Charting and Display System) developers. S57 3.1 is used almost exclusively for encoding Electronic Navigational Charts (ENCs), but can include hydrographic data that also can be used in other application areas, e.g. tidal data. Metadata for S57 3.1 can include, among others, coordinate system, projection, horizontal and vertical datum used, source scale and the units of height and depth measurements, information about data origins, and a description of the accuracy of the locational data. S57 3.1 data can be stored in ASCII or binary files.

S57 4.0 is under development by IHO's Transfer Standard Maintenance and Application Development (**TSMAD**) WorkingGroup. In addition to extending the S57 data model, version 4.0 will facilitate XML encoding of the geographic objects in the data model (in addition to ASCII and binary representation). S57 4.0 will also support new object types such as raster and gridded data, temporal and 3-dimensional data. The IHO standards process

GML 3 (Geography Markup Language) is developed by OGC (OpenGIS Consortium) as a standard for "modelling, transport and storage of geographic information". OGC is an international industry consortium of more than 250 companies, government agencies and universities participating in a consensus process to develop publicly available geo- processing specifications. GML provides an XML encoding for exchange of geographic information , and incorporates both spatial and non-spatial properties. The current version, GML 3.1 (ISO 19136), can be used to represent a number of geographic objects and concepts, including features, coordinate reference systems, geometry, topology, time, units of measure and generalised values, based on the abstract models defined in the ISO19000 series. Features can be simple features[5] (like points and lines), feature collections (sets of simple features), coverages (e.g. images) and observations (i.e. values captured with some form of instrument). GML 3.1 is an specification for XML technology implementation of several of OGC's Abstract Specifications and ISO standards for geographic information, among others the recently approved standard ISO 19115 Geographic Information – Metadata..

SDTS is described under Section 2.4.1.

---

[5] Defined by OGC as "features whose geometric properties are restricted to 'simple' geometries for which coordinates are defined in two dimensions and the delineation of a curve is subject to linear interpolation".

DIGEST is described under Section 2.4.1.

BUFR is described under Section 2.4.1.

Shapefile is a simple, non-topological format for storing the geometric location and attribute information of geographic features. I.e. the metadata is encoded within the geographic objects. The Shapefile format is developed by ESRI, and is widely used for transfer, display and manipulation of vector data. Shapefiles can be handled by a number of software packages, both commercial and free. A Shapefile consists of a main file, an index file and a dBASE table. All files are in binary format.

ArcGIS Marine Data Model is described under Section 2.4.1.

HydroML is an extension of the eXtensible Markup Language (XML) providing the Hydrologic Scientific Community with a standard definition of XML tags and concepts of structure to allow the definition of hydrologic information. HydroML can represent the following information: site information, computation instructions, corrections, ratings, shifts, time-series data including unit values and daily value statistics, peak flows, and site visit measurements. HydroML is currently under development. Data are stored and validated against an XML Schema.
Ref: http://water.usgs.gov./nwis_activities/XML/nwis_hml.htm

The **TBEIC** (Tokyo Bay Environmental Information Center) data model is based on ISO/TC211 recommendations and implemented by means of GML 3.0. The TBEIC data model is under development, and currently handles water quality (point) data.
Ref: http://www.tbeic.go.jp/

ChemicalUsageML is developed by POSC (Petrotechnical Open Software Corporation) for describing the nature, usage and potential hazards of chemicals to be used in the exploration and production of petroleum hydrocarbons. ChemicalUsageML is maintained by POSC, and can represent objects like points (well sites) and profiles (sections). Data are stored in XML format, using an XML Schema.
Ref: http://www.posc.org/, http://www.posc.org/technical/chemicalusage/index.html

WISTML v.1.2 (Wellsite Information Transfer Standard Markup Language) is developed by POSC (Petrotechnical Open Software Corporation) for well data. The standard is maintained by a SIG (Special Interest Group) where major oil companies (e.g. Statoil and BP) and contractors (e.g. Sclhumberger) participate. It is also supported by the UK Department of Trade and Industry (UK DTI). WISTML can represent objects like points (well sites) and profiles (drilling data). Data are stored in XML format, using an XML Schema.
Ref: http://www.witsml.org/

LandXML is a standard for developed by an Industry Consortium lead by Autodesk, consisting of some 40 land development and transportation companies and governmental organisations. LandXML is designed to facilitate the exchange of data created during the Land Planning, Civil Engineering and Land Survey process. The current version 1.0 was ratified by the LandXML.org consortium in July 2002. . The next version of LandXML is intended to be a GML application schema.

Ref: http://www.landxml.org/

XMML (The Exploration and Mining Markup Language) is a language for data needed in mining exploration and operation, including ore-bodies, boreholes, geophysics and samples. XMML is

based on GML and will therefore support all the (3-D) geometries described in ISO 19107 (Geographic information – spatial schema).
Ref: http://xmml.arrc.csiro.au/ https://www.seegrid.csiro.au/twiki/bin/view/Xmml/WebHome

NMEA 0183 & 2000 is two standards defined by NMEA (National Marine Electronics Association), which are widely accepted by manufacturers of marine electronics. These standards describe how instruments like GPS, echosounders, gyro repeaters etc. should communicate with each other, and are also used by CTD systems, ADCP software etc. for collecting navigational data.

The following description of NMEA 0183 is taken from the NMEA FAQ (Version 6.4):

> "Under the NMEA-0183 standard, all characters used are printable ASCII text (plus carriage return and line feed). NMEA-0183 data is sent at 4800 baud.
> The data is transmitted in the form of "sentences". Each sentence starts with a "$", a two letter "talker ID", a three letter "sentence ID", followed by a number of data fields separated by commas, and terminated by an optional checksum, and a carriage return/line feed. A sentence may contain up to 82 characters including the "$" and CR/LF."

In addition to position (latitude, longitude, elevation/depth), parameter measurements, e.g. water temperature, wind direction and speed, can also be sent in an NMEA "sentence".
Ref: http://vancouver-webpages.com/peter/nmeafaq.txt

2.4.3 Volume data standards

NetCDF is described under Section 2.4.1.

HDF is described under Section 2.4.1.

ArcGIS Marine Data Model is described under Section 2.4.1.

MMML (Model & Monitoring Markup Language) is a standard for data from hydrodynamic models, with metadata from the EDIOS model (encoded in XML). MMML is developed by RIKZ (National Institute for Coastal and Marine Management, the Netherlands).
Ref: http://hosting.xi-alles.nl/RWS/MMML/

2.5 Metadata standards

The following metadata standards have been reviewed. (Some are specific to marine data, while others are not.)

ISO 19115 is a geo-spatial metadata standard developed by ISO/TC 211. ISO 19115 defines a comprehensive metadata model for geographic objects, and was approved in late March 2003. Recognising that all data sets need not have such as complex metadata specification, ISO/TC 211 also defined a smaller set of *core* metadata elements (Table 2.2). This core contains the minimum elements that satisfy the requirements of an ISO conformant metadata record. The ISO 19115 standard does not specify storage format, but XML schemas are under development for an XML encoding of it (in full or for specialised profiles).
Ref: http://www.isotc211.org/publications.htm
http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=26020&ICS1=35&ICS2=240&ICS3=70
http://www.ned.dem.csiro.au/research/visualisation/metadata/geospatial/
Cox, Simon. Summary of some Geospatial Metadata Standards. Available [online] http://www.ned.dem.csiro.au/research/visualisation/metadata/geospatial/.

Table 2.2 ISO 19115 Core Metadata element set (from Cox, 2001).

Element	**Obligation**
Dataset title	M
Dataset reference date	M
Dataset responsible party	O
Geographic location of the dataset (by four coordinates or by geographic identifier)	C
Dataset language	M
Dataset character set	C
Dataset topic category	M
Scale of the dataset	O
Abstract describing the dataset	M
Dataset format name	O
Dataset format version	O
Additional extent information for the dataset (vertical and temporal)	O
Spatial representation type	O
Reference system	O
Lineage statement	O
On-line resource	O
Metadata file identifier	O
Metadata standard name	O
Metadata standard version	O
Metadata language	C
Metadata character set	C
Metadata point of contact	M
Metadata time stamp	M

Obligation: M = Mandatory, C = Conditional, O = Optional.

CSDGM (Content Standard for Digital Geospatial Metadata) is a standard for metadata for geographic objects developed by FGDC (Federal Geographic Data Committee). The CSDGM metadata model is also comprehensive, but many of the elements are optional (Figure 2.5). It also enables development of profiles, i.e. customisation of the standard to suit the needs of a particular application domain (while staying within the framework of the standard). CSDGM is widely used within the U.S. and to some lesser degree in other countries. An XML encoding (DTD) is provided by FGDC.
Ref: http://www.fgdc.gov/metadata/contstan.html

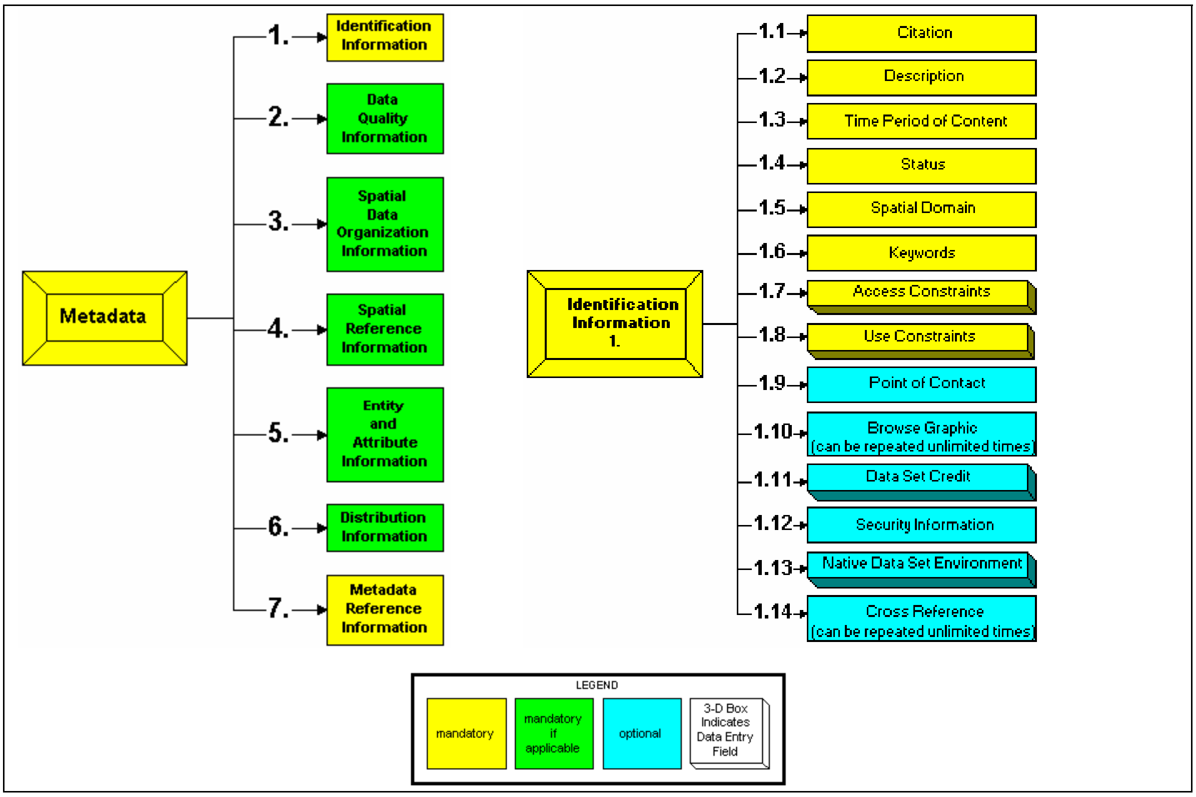

Figure 2.5 CSDGM meta-data model. Left: main components. Right: Elements for identification.

Dublin Core Metadata is a general metadata standard for describing resources available on the WWW. It is created by the Dublin Core Metadata Initiative, which is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. Dublin Core (DC) is widely used to document the availability of papers, book, and similar resources that can be accessed via the Internet.

Table 2.3 gives an overview of the elements in the Dublin Core metadata element set. In addition, there are a number of other elements and element refinements that can be used to provide more information about an online resource. As can be seen from Table 2.3, the DC enables a very general description of online resources, and can therefore be used for a number of different resource types. Dublin Core is especially popular for describing various types of documents (articles, books, etc.).

However, Dublin Core can also be used to describe the location and time of geographic data, for instance, through the element labelled *Coverage*, which can include spatial location, temporal period or jurisdiction (e.g. administrative entity). The *Coverage* element can be refined using the "type" qualifiers *Spatial* and *Temporal*, which also give the opportunity to describe the spatial and temporal characteristics of a resource. Hence, DC can also be used for description of location and timing of geographic data sets.

Ref: http://dublincore.org/ and http://dublincore.org/documents/

Table 2.3 Dublin Core metadata element set.

Label	**Definition**
Title	A name given to the resource.
Creator	An entity primarily responsible for making the content of the resource.
Subject	The topic of the content of the resource.
Description	An account of the content of the resource.
Publisher	An entity responsible for making the resource available.
Contributor	An entity responsible for making contributions to the content of the resource.
Date	A date associated with an event in the life cycle of the resource.
Type	The nature or genre of the content of the resource.
Format	The physical or digital manifestation of the resource.
Identifier	An unambiguous reference to the resource within a given context.
Source	A reference to a resource from which the present resource is derived.
Language	A language of the intellectual content of the resource.
Relation	A reference to a related resource.
Coverage	The extent or scope of the content of the resource.
Rights	Information about rights held in and over the resource.

DIF (Directory Interchange Format) is a general format for exchange of information about scientific data sets. DIF was first proposed at a workshop on catalogue interoperability held in February 1987. After several demonstrations, workshops, and feedback from the scientific community, the Directory Interchange Format (DIF) was formally approved and adopted by a CI science advisory group at a CI workshop in 1988. Since then it has been widely used in online catalogue systems, among others, in the Global Change Master Directory (GCMD) for Earth sciences applications (formerly called NASA's Master Directory (NMD)). DIF is a standard for creating directory entries describing data, and contains a collection of fields detailing specific information about the data. Six of these fields are mandatory (directory entry identifier, directory entry title, parameters, data center, summary and document author); others are optional and used to expand upon and clarify the information provided. A skinny DIF is a DIF that consists of only the required DIF fields. The Directory Interchange Format is compatible with the U.S. federally mandated Federal Geographic Data Committee's Content Standard on Digital Geospatial Metadata (CSDGM). DIF data are stored in ASCII files.
Ref: http://gcmd.gsfc.nasa.gov/User/difguide/difman.html Directory Interchange Format (DIF) Writer's Guide, Version 8. 2003. Global Change Master Directory. National Aeronautics and Space Administration. http://gcmd.nasa.gov/User/difguide/.

ANZLIC – the Spatial Information Council is a joint initiative of the Australian Government, the New Zealand Government and the governments of the States and Territories of Australia. The **ANZLIC metadata standard** defined a set of core metadata elements (Figure 2.6), which must be filled in for each geographic dataset. The ANZLIC core elements are defined in the ANZMETA Document Type Definition (DTD).

Ref: http://www.anzlic.org.au/infrastructure_metadata.html http://www.ga.gov.au/anzmeta/

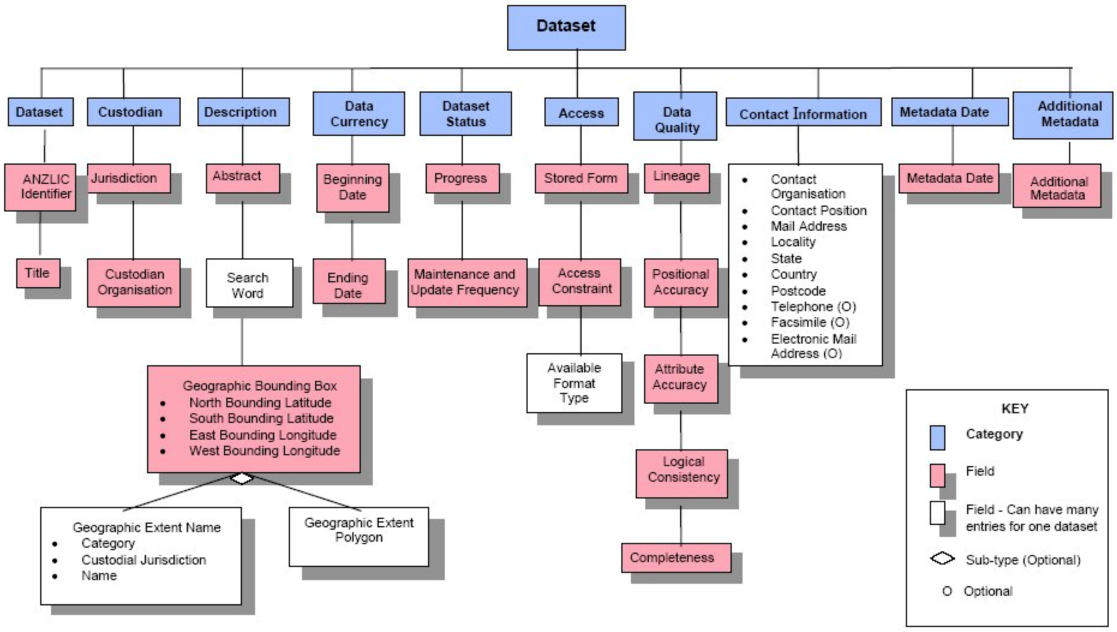

Figure 2.6 ANZLIC Core Metadata Elements (from ANZLIC Metadata Guidelines: Core metadata elements for geographic data in Australia and New Zealand – Version 2.0 (February 2001)).

NcML (NetCDF Markup Language) defines metadata for generic NetCDF data. NetCDF is a common format used in global oceanographic programmes such as Argo. The Netcdf Metadata Object Model (NMOM) was developed for generic data in NetCDF format. Thus, NcML only includes features that are common to all data in NetCDF format, and therefore was developed to be as simple and generic as possible. Figure 2.7 shows the main objects in the NcML data model. It is implemented in an XML Schema.

Ref: http://www.vets.ucar.edu/luca/netcdf/

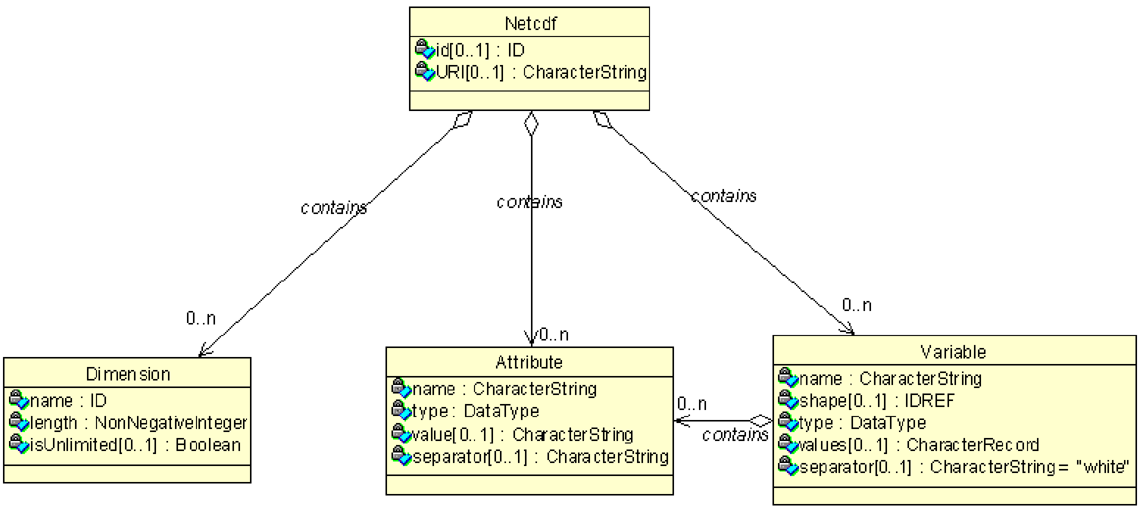

Figure 2.7 NcML Data Model.

JMGRIB 1.5 is an application of XML to describe a set of gridded data, which are packed in GRIB records. JMGRID defines a collection of elements that can be represented in four different

formats. First, a "raw" JMGRIB file containing metadata (title, location, time, projection, etc.) and has a reference to an external GRIB file (which holds the data). Secondly, an "encoded" JMGRIB file, which as the same metadata, but the GRIB file is included –as ASCII characters using BASE64 encoding. Both of these formats will require a GRIB decoder to extract the data. The third format, "expanded" JMGRIB includes metadata (same as the previous two) and then the data values encoded (i.e. expanded) as ASCII text. The fourth format, a grid-point JMGRIB file is suited for small datasets because it includes location parameters for each data point. The fourth format is in XML too. A DTD is defined for all four types of JMGRIB files.
Ref: http://zowie.metnet.navy.mil/~spawar/JMV-TNG/XML/JMGRIB.html

IWICOS (Integrated Weather, Sea Ice and Ocean Service System) is a metadata specification developed during an IST project running from 2000 to 2002. The IWICOS metadata specification is similar in structure to the CSDGM standard, but has some additions that were needed to represent metadata for the various types of met-ice-ocean products produced in the IWICOS project. The IWICOS specification for metadata is defined using XML Schema. The data values were kept in separate files, in BSQ, GRIB, Shapefile or XML format.
Ref: http://web.dmi.dk/pub/IWICOS/metadata/newdef/

2.6 Presentation standards

The following presentation standards have been reviewed. (Some are specific to marine data, while others are not.)

S52 is a specification for Chart Content and Display Aspects of ECDIS, and is developed by IHO. S52 is a "Presentation Model" in the sense that it contains information about how the encoded geographic objects should be displayed on the screen, e.g. what colour, line thickness and type should be used for a line object.

SVG (Scalable Vector Graphics) is a W3C standard for describing two-dimensional graphics in XML. SVG also supports animations (by adopting SMIL (see below)). The current version is 1.1 (W3C Recommendation), while SVG 1.2 is under development. A number of stand-alone SVG viewers exist, and a list is provided on http://www.w3.org/Graphics/SVG/SVG-Implementations.htm8. Several browser plug-ins also exist, enabling SVG to be displayed in a common desktop browser.
Ref: http://www.w3.org/Graphics/SVG/

SMIL 2.0 (Synchronized Multimedia Integration Language) is defined by W3C for simple authoring of interactive multimedia presentations. SMIL is typically used for "rich media"/multimedia presentations which integrate streaming audio and video with images, text or any other media type. W3C has defined the SMIL language in both a DTD and an XML Schema.
Ref: http://www.w3.org/TR/smil20/

GraphML is a comprehensive and easy-to-use file format for graphs. Its development started by a group from academia and industry with the aim to define a new XML-based file format that is to eventually form a standard for the graph drawing community and its co-operators. The version GraphML 1.0rc (release candidate) was released in mid March 2003.
Ref: http://graphml.graphdrawing.org/

General-purpose presentation formats include, among others, **HTML** (HyperText Markup Language), **DHTML** (Dynamic HTML), **XHTML** (eXtensible HyperText Markup Language) and **PDF** (Portable Document Format). These are typically used for displaying marine data and metadata in a web browser or (for PDF) for generating hardcopies of information sheets etc.

2.7 Other standards

The following other standards have been reviewed:

JMCDM (METOC conceptual data model) is a logical data model that integrates the geophysical data requirements of all DoD (Department of Defense) components. JMCDM and its supporting encyclopedia will serve as input for the expansion of the DoD Data Model. JMCDM is being developed by the Meteorology and Oceanography (METOC) Data Administration.

Epicentre v3.0 is a logical data model for E&P (Exploration & Production) information. It is used by several or POSC (Petrotechnical Open Software Corporation) data models (e.g. ChemicalUsageML). POSC is the standards body for technical information relating to the Oil and Gas exploration and production industry. A not-for-profit organization, POSC has members in 24 countries and is strongly supported by all sectors of the industry, including oil companies, software companies, other standards organizations within the oil and gas sector and governments. http://www.posc.org/Specifications/Epicentre_V30/index.html

EDCS (Environmental Data Coding Specification) is a general data coding of environmental objects, developed by ISO (as specification ISO/IEC 18025). EDCS defined nine dictionaries of environmental concepts and a functional interface. These directories contain the following information:

1. classifications (type of environmental objects)
2. attributes (the objects' state)
3. attribute value characteristics (information on the values of attributes, e.g. data type)
4. attribute enumerants (allowed values for an enumerated attribute)
5. units (quantitative measures of the state of an environmental object)
6. unit scales (allow a wide range of numerical values to be stated)
7. unit equivalence classes (sets of units that are mutually comparable)
8. organizational schemas (for locating classifications and attributes sharing a common context), and
9. groups (for collecting concepts sharing a common context).

Each directory has a defined set of attributes that are used to describe the corresponding concept. For instance, a classification is described by attributes Label, Definition, Code, Groups, Reference Type, and References.
Ref: http://www.gscassociates.com/wg8/edcs/

SensorML (Sensor Model Language) is an OGC initiative that aims to provide an XML schema for defining the geometric, dynamic, and observational characteristics of a sensor. Sensors are devices for the measurement of physical quantities. There are a great variety of sensor types from simple visual thermometers to complex electron microscopes and earth observing satellites. The current status of SensorML is a discussion paper.
Ref: http://opengis.org/techno/discussions/02-026r4.pdf

Observations and Measurements is another OCG initiative that aims to develop general models and XML encodings for sensor observations and measurements. This was originally an OpenGIS® Interoperability Program Report, but is now a module within GML 3
Ref: http://www.opengis.org/techno/discussions/02-027.pdf

Both SensorML and Observations and Measurements are part of a larger OGC initiative named OpenGIS SensorWeb™, which aims at developing interoperability interfaces and meta-data encodings for real time integration of heterogeneous sensors into the Web.
Ref: http://ip.opengis.org/ows2/

Coastal XML / CML (Coastal Mark-Up Language) is being developed in an RTD project lead by Prof. D. Wright, Oregon State University, US. At present, no information is available about the status of Coastal XML / CML.
Ref: http://www.diggov.org/archive/projects/046.jsp [worked 5 March 2004]

CML 2.1.1 (Chemical Markup Language) is an XML based encoding of molecular information, ranging from macromolecular sequences to inorganic molecules and quantum chemistry. CML can be converted to SVG for display in e.g. a web browser. There are also several software tools, e.g. JUMBO, that can read and display a CML file. Both DTD and XML Schema are defined for CML.
Ref: http://www.xml-cml.org/, http://xml.coverpages.org/materials.html http://www.xml-cml.org/information/position.html

XSIL (Extensible Scientific Interchange Language) is a "flexible, hierarchical, extensible, transport language for scientific data objects". XSIL was developed by a team at CACR (Center for Advanced Computer Research), California Institute of Technology, and has been used in multiple projects in various domains, e.g. LIGO (Laser Interferometer Gravitational Wave Observatory) and Digital Puglia (a digital archive of remote sensing data).
Ref: http://www.cacr.caltech.edu/SDA/xsil/

BSML 3.1 (The Bioinformatic Sequence Markup Language) is an XML Data Standard for Genomics. BSML encodes biological sequence information and includes graphical representations of biologically meaningful objects such as sequences, genes, electrophoresis gels, and multiple alignments. Encoding is specified as a DTD.
Ref: http://www.bsml.org/

IML 1.0 (Image Markup Language) is metadata structure for storing textual annotations to GIF or JPEG images, and is used in medical education software and clinical information systems. IML is developed by the Structural Informatics Group, Department of Biological Structure, School of Medicine, University of Washington. The implementation of the structure is specified as a DTD.
Ref: http://faculty.washington.edu/lober/iml/

ESML 3.0 (Earth Science Markup Language) is developed to facilitate easier data exchange in the Earth science community, by defining metadata in a platform independent language (XML) and referencing the associated data file(s). ESML is developed by the Information Technology and Systems Center, University of Alabama in Huntsville. The ESML metadata structure is defined in an XML Schema, which is illustrated in Figure 2.8.

Ref: http://esml.itsc.uah.edu/index.jsp

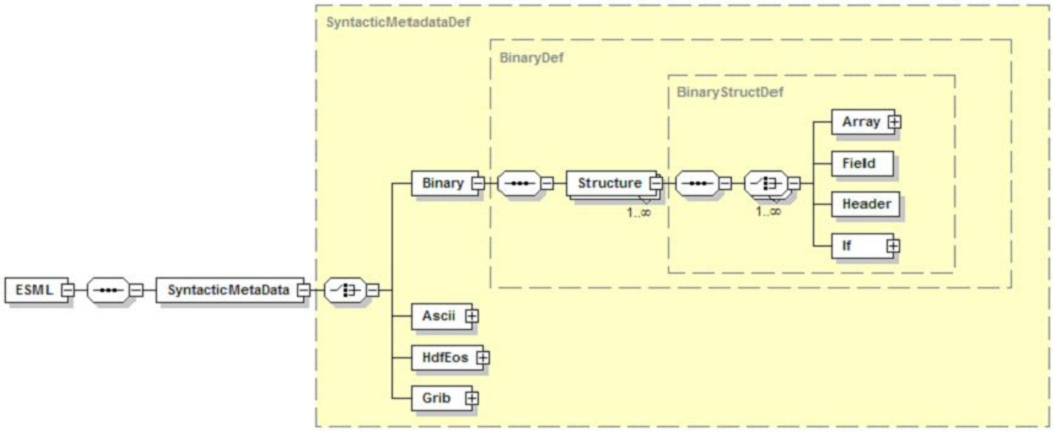

Figure 2.8 Excerpts of ESML schema.

EPSG 6.5 (European Petroleum Survey Group) has developed a database of geodesy parameters, which are available online in MS Access format. This database is used by e.g. the DEAL Data Registry.
Ref: http://www.epsg.org/

3 Outline of inter-standard relationships model

3.1 Ontology Facets

During the review phase, an inter-standard relationships model has been developed that can be understood by computers, i.e. an ontology. This ontology is termed MISTR (marine inter-standard relationships model),

The main part of ontology development is the definition of classes, slots, facets and relationships between sets of classes or instances. The CERIF[6] (Common European Research Information Format) model was used as a base model for the cataloguing of standards and their relationships. CERIF is an EC standard developed to harmonise databases on research projects. CERIF builds on the CRIS (Current Research Information System) data model, which contains:

- Principal entities: Person, Project, OrgUnit
- Secondary entities: Results, Classification, Contact, Event
- Level 3 (translation): Project-title, Project-abstract, Project-keywords, etc.
- "Look-up" tables: Project-status, Person-honorific-title, etc.
- Level 5 ("many-to-many"): e.g. Project_Person and Person_Person.

As the CERIF data model was designed for another purpose than MISTR, the approach in MarineXML has been to adapt the entities of CERIF to our needs. For instance, principal entities like Person was kept, but project specific attributes like Project-Result_Publication was left out. The classes defined for MISTR are the following:

Project
Person
Organisation
Contact
Project_Person
Organisation_Person
Organisation_Project
(adapted from CERIF)

Standard
Standard_Project
Standard_Organisation
ContentModel
DataStructure
Encoding
Presentation
Standard_Person
Standard_Relationship
Topic
(new classes)

Another important part of the ontology development was to define the relationships between classes or instances. Figures 3.1-3.5 illustrate some of the relationships in the MISTR ontology.

3.2 Ontology Implementation

Protégé[7], an open source ontology editor written in Java, was used to generate the ontology and to populate it with information obtained during the review. Protégé is freely available as source code or pre-compiled for many platforms (e.g. Windows, Linux, Solaris, Mac OS X). In addition to providing tools for defining the ontology, Protégé also automatically generates input forms, and enables the user to fill in data into these forms to populate the ontology. Protégé can save the ontology, including its data, as text (CLIPS format), JDBC and RDF files. The tools also has an "Save as HTML" option, which allows an browsable version of the ontology to be generated.

Finally, the ontology was populated with information on standards etc. from the review process by means of the Protégé tool. MISTR can be downloaded from www.marineXML.net.

[6] Current version is CERIF-2002, available from http://www.cordis.lu/cerif/
[7] Current version is 1.8, available from http://protege.stanford.edu

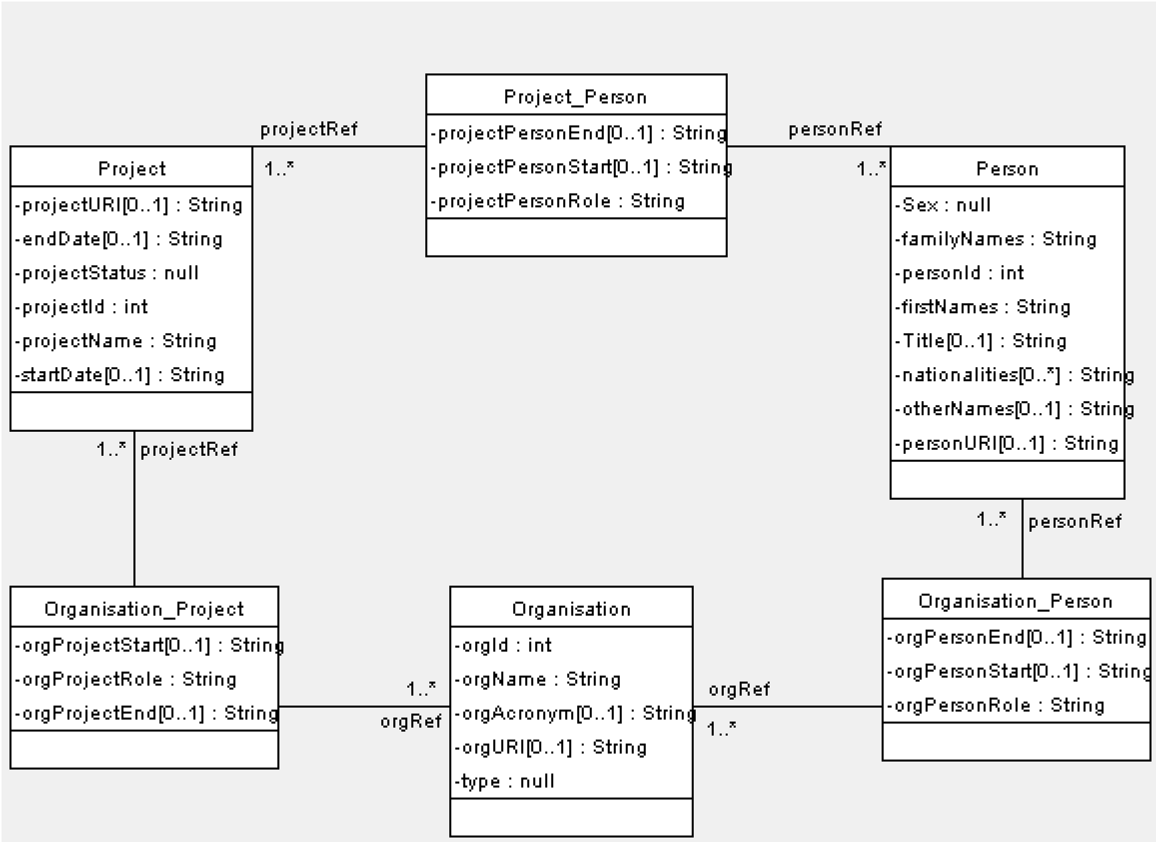

Figure 3.1 Relationships between classes in the MISTR ontology (part 1).

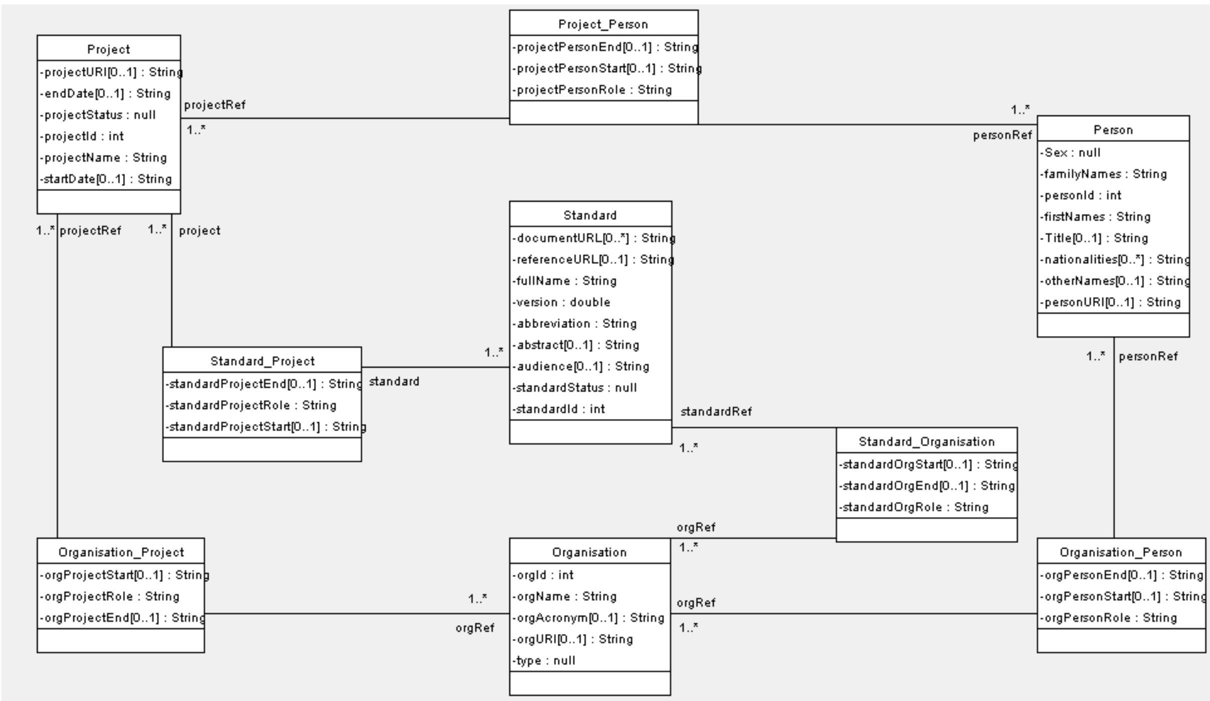

Figure 3.2 Relationships between classes in the MISTR ontology (part 2).

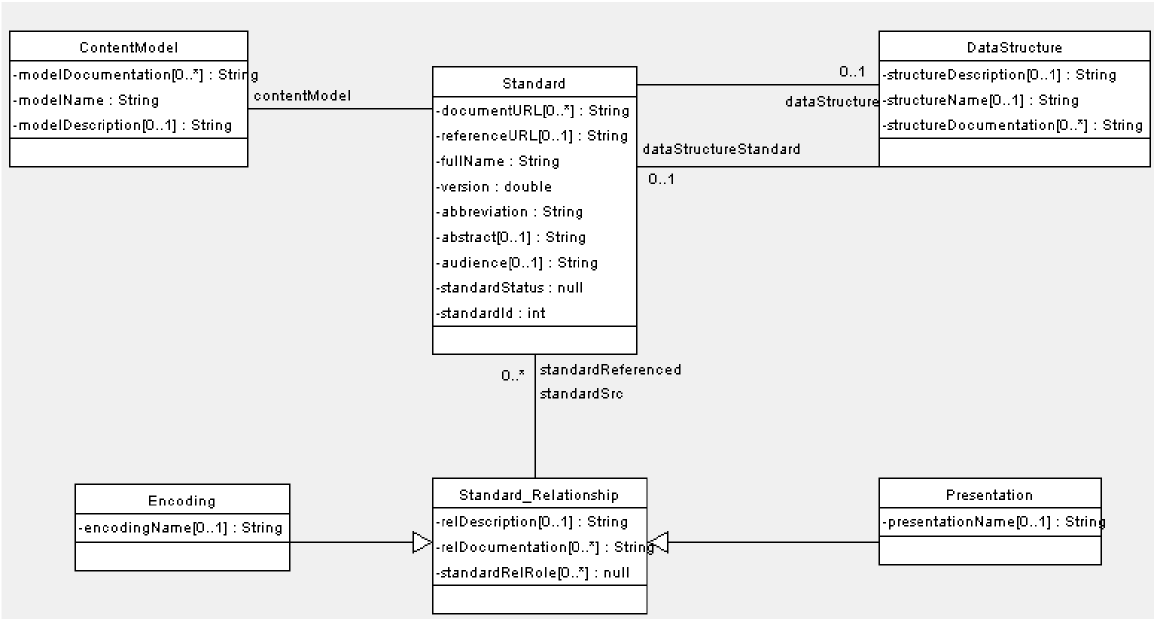

Figure 3.3 Relationships between classes in the MISTR ontology (part 3).

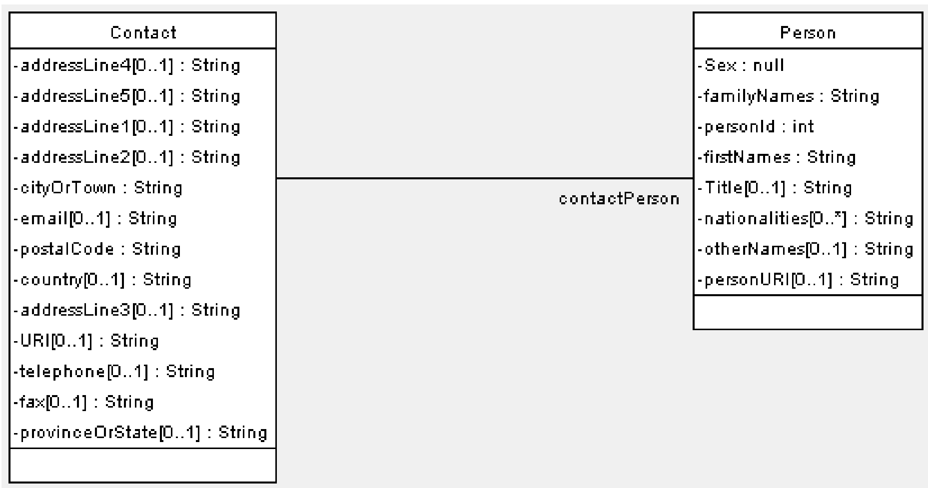

Figure 3.4 Relationships between classes in the MISTR ontology (part 4).

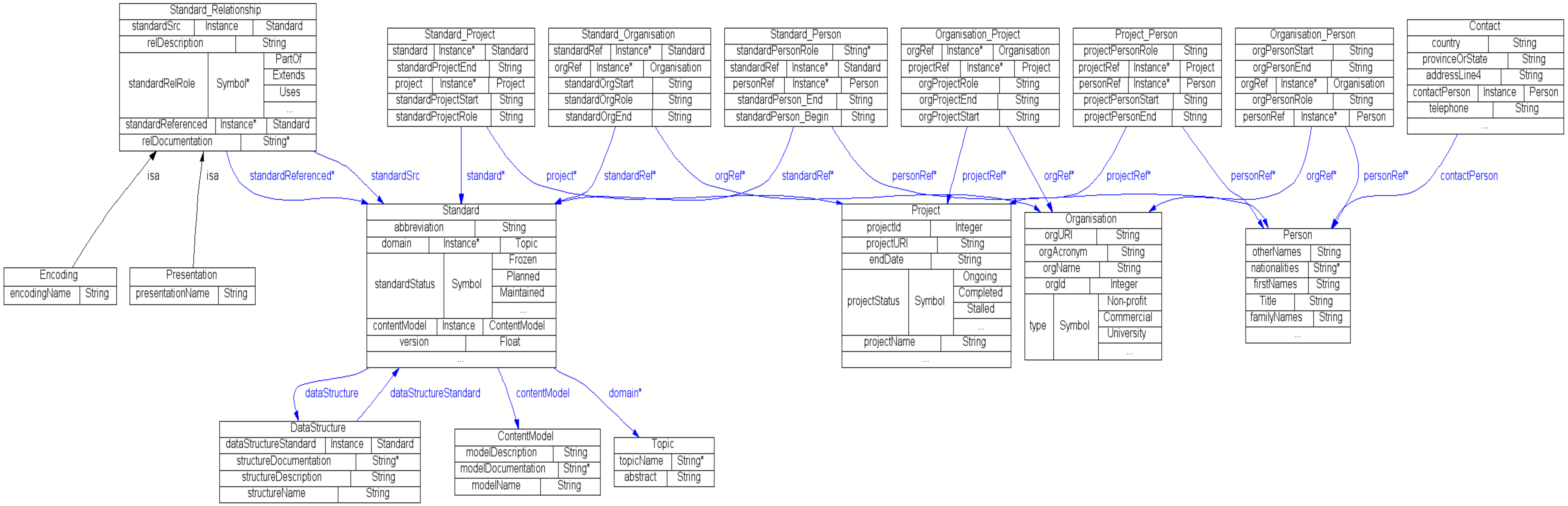

Figure 3.5 Classes, slots and relationships in the marine inter-standard relationships model.

4 Conclusions and recommendations

The review phase has led to the documentation of a number of marine and other data (and metadata) standards, programmes and projects. These have been investigated and evaluated with respect to identifying a candidate set of standards that the MML specification should unify (and demonstrate in a testbed) and the MML specification can use as building blocks

The criteria for selecting the standards to be incorporated in MML is a combination of :

- The importance of the standard (i.e. how widely it is used within the marine community)
- The governance model and future evolution of the standard (i.e. is the lifecycle of the standard considered)
- The desire to incorporate standards in MML to test the strength of the interoperability framework (i.e. MML should rather include one standard for satellite images and one *in situ* instrument specific standard, than two standards for the same type of data)
- What is practical and desirable to achieve within the current project as part of the 'bigger picture'.

A data model (ontology) for inter-standard relationships has also been developed in parallel with the review of standards, programmes and projects. An outline of the process of defining this model has been given in this report. The data model itself is a separate deliverable.

From this initial review process, one of the key trends is the harmonisation and adoption of OGC standards within the ISO TC211 (ISO19000) series of standards. These standards include:

19135 – Procedures for registration of geographic items
19110 – Feature Type Cataloguing methodology
19126 – UML/XML implementation of 19110,19135
19136 – GML
19139 – GML/19115 metadata implementation

The adoption of GML as ISO 19136 means potentially that MarineXML no longer has to make a choice between the ISO or OGC path for long term standardisation. This makes GML the (only) real choice as the basis for developing a marine mark-up language. In addition, the commitment of IHO to release version 4 of S-57 along ISO19000 standards means not only will it become possible to reuse S-57 features for other purposes beyond navigation, but any IHO registry for this standard could be used to register other, non-navigational, marine features. IOC support this idea and IHO and IOC have agreed in principle to investigate how such a joint registry could work.

Given this convergence, the new challenge is to create the tools to support the interoperability of data models across overlapping domains. This includes the concepts and implementation strategies for the implementation of GML application schemas in overlapping domains. In particular, the use of a Feature Type Catalog as a Web Services enabled registry, able to support emerging libraries of component schema building blocks as well as higher-order semantics of relationships between Feature Types.

It is recognised that the standards landscape is continually evolving and as such this ontology within the resources of this project can never be 100% completed, accurate or finalised. Accordingly methods are being investigated as part of the post-project exploitation to serve this standard's ontology to wider marine community; in particular responsibilities for its update. In its present state the ontology is served in simple HTML format, together with a download of the Protégé files. The ontology to be updated at least once prior to the end of the project to make reference the standards used as part of the MarineXML test beds.

Acknowledgements

The European Commission under the Fifth Framework Programme has made MarineXML possible through a funding contribution that the project partners would like to acknowledge. The contract (number EVK2-CT-2002-00177) ran from February 2003 to January 2005.

The authors of this report would like to acknowledge all members of the MarineXML team for their inputs to this document. The team is :

MarineXML Members

Keiran Millard	HR Wallingford
Jane Smallman	HR Wallingford
Jim Clipson	United Kingdom Marine Information Council
Dan Pillich	SevenCs
Holger Bothien	SevenCs
Torill Hamre	Nansen Environmental and Remote Sensing Center
Stein Sandven	Nansen Environmental and Remote Sensing Center
Juan Bicarregui	Central Laboratory of the Research Councils
Brian Matthews	Central Laboratory of the Research Councils
Pieter Harring	National Institute for Coastal and Marine Management
Edward Vanden Berghe	Flemish Marine Institute
Francisco Hernandez	Flemish Marine Institute
Rob Atkinson	Social Change Online
Simon Cox	CSIRO
Hans Dahlin	EuroGOOS
Nic Flemming	EuroGOOS
Peter Pissierssens	International Oceanographic Data Exchange Committee
Greg Reed	International Oceanographic Data Exchange Committee

MarineXML Affilitates

Kathy Shea	US Coast Guard Service
Roy Lowry	British Oceanographic Data Centre
Ray Cramer	British Oceanographic Data Centre
Kevin O’Neill	Central Laboratory of the Research Councils
Andrew Woolf	Central Laboratory of the Research Councils

App. A Synopsis of selected standards

A.1 Spatial data standards

Name:	S57 3.1
Abbreviation:	S57 3.1
Purpose:	S57 3.1 is an IHO standard for the exchange of digital hydrographic data. Version 3.1 is the current version, which is endorsed by all major commercial ECDIS (Electronic Charting and Display System) developers. S57 3.1 is used almost exclusively for encoding Electronic Navigational Charts (ENCs), but can hydrographic data that also can be used in other application areas, e.g. tidal data. Metadata for S57 3.1 can include, among others, coordinate system, projection, horizontal and vertical datum used, source scale and the units of height and depth measurements, information about data origins, and a description of the accuracy of the locational data.
Version:	3.1 (November 2000)
Owner:	IHO (The International Hydrographic Organization)
Maintenance:	S57 3.1 will be maintained and supported IHO also after version 4.0 is finished.
Relationship to others:	Uses S52 for display.
Encoding:	The S57 3.1 data structure can be stored in ASCII or binary files, according to the product specification for the data set to be transferred.
References:	http://www.ohi.shom.fr IHO TRANSFER STANDARD for DIGITAL HYDROGRAPHIC DATA, Edition 3.1 - November 2000, Special Publication No. 57. Published by the International Hydrographic Bureau, MONACO

Name:	S57 4.0
Abbreviation:	S57 4.0
Purpose:	S57 4.0 is under development by IHO. In addition to extending the S57 data model. Version 4.0 will facilitate XML encoding of the geographic objects in the data model. Encodings in ASCII or binary format will still be supported. In addition to the object types supported in version 3.1, version 4.0 will also be support raster and gridded data, temporal and 3-dimensional data. S57 4.0 will be based to the greatest extent possible on geospatial data standards from ISO.
Version:	4.0 (forthcoming)
Owner:	IHO (The International Hydrographic Organization)
Maintenance:	IHO will maintain and support S57 4.0 as the new de facto standards for ECDIS.
Relationship to others:	Extends S57 3.1. Will use S52 for display.
Encoding:	XML, ASCII and binary.
References:	http://www.iho.shom.fr/general/ecdis/COMMITTEE/TSMAD/Welcome_to_IHO/S57_Edition_4_Dev_Page.htm

Name:	Spatial Data Transfer Standard
Abbreviation:	SDTS
Purpose:	**SDTS** (Spatial Data Transfer Standard) is a standard for transfer of spatial digital data developed by the USGC (U.S. Geological Survey). Federal agencies in the U.S. must comply with SDTS, which is also widely used in other governmental organisations. Large commercial companies like ESRI, ERDAS and MapInfo have included STDS handling in their software. SDTS can hold both vector and raster data, with geometry and topology, as well as associated metadata (e.g. data quality). STDS is stored according to the ISO 8211 Data Encoding standard.
Version:	ANSI NCITS 320-1998
Owner:	American National Standards Institute, Inc (ANSI)
Maintenance:	Maintained and supported by ANSI.
Relationship to others:	
Encoding:	ISO 8211
References:	http://mcmcweb.er.usgs.gov/sdts/ Spatial Data Transfer Standard (SDTS) - Part 1, Logical Specifications. Draft. DRAFT for Review, November 20, 1997. American National Standards Institute, Inc.

Name:	Digital Geographic Information Exchange Standard
Abbreviation:	DIGEST
Purpose:	DIGEST is developed by the Digital Geographic Information Working Group (DGIWG) to support efficient exchange of digital geographic information between national agencies, data producers and end users. DIGEST is a comprehensive "family of standards" that can represent raster, matrix, and vector data (and associated text), and supports entire range of levels of topological structures. DIGEST can be encoded in multiple ways, e.g. using ISO 8211 and ISO 8824 standards.
Version:	2.1 (September 2000)
Owner:	Digital Geographic Information Working Group (DGIWG)
Maintenance:	DIGEST is maintained by the DGIWG.
Relationship to others:	Uses several standards for encoding, e.g. ISO 8211 and ISO 8824.
Encoding:	Multiple: ISO 8211 (Information Processing), ISO 8824 (Information Processing Systems – Open Systems Interconnection – Specifications of Abstract Syntax Notation One), and ISO 8825 (Information Processing Systems – Open Systems Interconnection – Specifications of Basic Encoding Rules for Abstract Syntax Notation One). VRF (Vector Relational Format). IIF (Image Interchange Format).
References:	http://www.digest.org/ The Digital Geographic Information Exchange Standard (DIGEST) Part 1 GENERAL DESCRIPTION Edition 2.1. September 2000. Produced and issued by the Digital Geographic Information Working Group (DGIWG).

Name:	network Common Data Form
Abbreviation:	NetCDF
Purpose:	NetCDF is a machine-independent format for representing scientific data, which was developed by Unidata Program Center in Boulder, Colorado. NetCDF is array-oriented and is used for many different types of data, e.g. for satellite images and results of model predictions. The format is widely used in the scientific community around the world, and a number of software packages are available for display and manipulation. These are primarily public domain tools, but also some COTS tools exist (e.g. MATLAB and IDL). Metadata can be stored as NetCDF attributes, which are linked to NetCDF variables of a given dimension. NetCDF data are stored in binary files.
Version:	Version 3 (June 1997)
Owner:	Unidata Program Center in Boulder, Colorado
Maintenance:	NetCDF is maintained and supported.
Relationship to others:	
Encoding:	Binary.
References:	http://www.unidata.ucar.edu/packages/netcdf/ NetCDF User's Guide for C. An Access Interface for Self-Describing, Portable Data. Version 3, June 1997. Russ Rew, Glenn Davis, Steve Emmerson, and Harvey Davies. Unidata Program Center http://www.cgd.ucar.edu/cms/eaton/cf-metadata/index.html http://ferret.wrc.noaa.gov/noaa_coop/coop_cdf_profile.html

Name:	Hierarchical Data Format
Abbreviation:	HDF
Purpose:	HDF is a standard format for scientific data exchange used in a number of application domains, e.g. for transfer of satellite images. HDF is developed by NCSA (National Center for Supercomputer Applications), and like NetCDF it is a popular format in the scientific community world-wide. Version 5.1.6.2 is the most current version, while older versions 4.x are also widely used (and supported by NCSA). Metadata is stored as attributes in the defined objects (classes). HDF files are also binary.
Version:	Version 5.1.6.2 (February 2004)
Owner:	National Center for Supercomputer Applications (NCSA)
Maintenance:	HDF 5 is maintained and supported by NCSA. HDF 4 is supported.
Relationship to others:	
Encoding:	Binary.
References:	http://hdf.ncsa.uiuc.edu/, http://hdf.ncsa.uiuc.edu/HDF5/doc/ Spatial Data Transfer Standard (SDTS) - Part 1, Logical Specifications. Draft. DRAFT for Review, November 20, 1997. American National Standards Institute, Inc.

Name:	GeoTIFF
Abbreviation:	GeoTIFF
Purpose:	GeoTIFF represents an effort by over 160 different remote sensing, GIS, cartographic, and surveying related companies and organizations to establish a TIFF based interchange format for geo-referenced raster imagery. GeoTIFF 1.0 is the official release version. GeoTIFF implements the geographic metadata formally, using compliant TIFF tags and structures. GeoTIFF data are stored as TIFF (i.e. as binary files).
Version:	1.0 (December 2000)
Owner:	GeoTIFF was originally developed by Dr. Niles Ritter, while at NASA-JPL (Jet Propulsion Laboratory). It is in the public domain.
Maintenance:	GeoTIFF is maintained and support by volunteers from the remote sensing community; many large commercial companies are involved.
Relationship to others:	Uses TIFF
Encoding:	TIFF (binary)
References:	http://www.remotesensing.org/geotiff/geotiff.html GeoTIFF Format Specification. GeoTIFF Revision 1.0. Niles Ritter and Mike Ruth. http://www.remotesensing.org/geotiff/spec/geotiffhome.html

Name:	Committee on Earth Observation Satellites
Abbreviation:	CEOS
Purpose:	CEOS is a standard format for satellite data from e.g. ESA PAFs (Processing Archive Facilities). The CEOS format is also used by other satellite data providers, and can be ingested in many image processing packages. Metadata often includes information on numerous satellite parameters, and is stored in the 'header' file of a CEOS data set, which is typically composed of multiple files. CEOS files are binary.
Version:	3.0
Owner:	European Space Agency (ESA)
Maintenance:	Maintained and supported by ESA.
Relationship to others:	
Encoding:	Binary
References:	http://www.ceos.org/ Annex D ERS SAR.PRI CCT and EXABYTE FORMAT SPECIFICATIONS. Document No: ER-IS-EPO-GS-5902.4. Issue: 3.0, August 1, 1998. Ola Grabak, ESA. http://earth.esa.int/rootcollection/sysutil/sarpri.html

Name:	GRidded In Binary (WMO FM92 GRIB)

Abbreviation:	GRIB
Purpose:	GRIB is a WMO format for storage of weather product information from national weather forecasting centres and other operational organisations. Metadata such as geo-referencing information and units for parameters are included in the format. GRIB files are binary.
Version:	Edition 2
Owner:	World Meteorological Organisation (WMO)
Maintenance:	GRID is maintained and supported by WMO.
Relationship to others:	
Encoding:	Binary
References:	WMO GRIB: ftp://ncardata.ucar.edu/libraries/grib/ http://www.wmo.ch/web/www/WDM/Guides/Guide-binary-2.html http://www.wmo.ch/web/www/DPS/grib-2.html ECMWF GRIB: ftp://ncardata.ucar.edu/datasets/ds111.2/format ftp://ncardata.ucar.edu/datasets/ds111.2/software http://www.faqs.org/faqs/graphics/fileformats-faq/part3/index.html

Name:	Binary Universal Form for the Representation of Meteorological Data (WMO FM94)
Abbreviation:	BUFR
Purpose:	BUFR is a WMO standard for encoding meteorological and other (primarily) observational data as a binary 'message', which can be efficiently transmitted and stored. Metadata (geo-referencing, units, compression, etc.) is stored as part of the BUFR message. BUFR files are binary, and all codes used for encoding observations must be predefined to enable correct decoding at the receiver.
Version:	Edition 2 (1991)
Owner:	World Meteorological Organisation (WMO)
Maintenance:	BUFR is maintained and supported by WMO.
Relationship to others:	
Encoding:	Binary
References:	http://www.wmo.ch/web/www/WDM/Guides/Guide-binary-1A.html http://www.wmo.ch

Name:	Shapefile

Abbreviation:	Shapefile
Purpose:	Shapefile is a simple, non-topological format for storing the geometric location and attribute information of geographic features. I.e. the metadata is encoded within the geographic objects. The Shapefile format is binary and is developed by ESRI. It is widely used for transfer, display and manipulation of vector data, and can be handled by a number of software packages, both commercial and free.
Version:	
Owner:	Environmental Systems Research Institute (ESRI)
Maintenance:	ESRI maintains and supports Shapefile
Relationship to others:	
Encoding:	A Shapefile consists of a main file, an index file and a dBASE table. All files are in binary format.
References:	http://www.esri.com/ ESRI Shapefile Technical Description. An ESRI White Paper – July 2998. http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf

Name:	ArcGIS Marine Data Model
Abbreviation:	ArcGIS Marine Data Model
Purpose:	The ArcGIS Marine Data Model defines an object-oriented data model (geodatabase) for a wide ranges of marine data and (physical) objects. For instance, point data, raster and gridded data, volume data and time series (of points, etc.). The model also includes information about the instruments that generated the observations, cruises, etc. Metadata includes, among others, geo-location and time stamping. The data model is developed in UML (Unified Modelling Language), which is transformed to an ArcGIS 8 database schema. The ArcGIS Marine Data Model represents “a new approach to spatial modeling via improved integration of many important features of the ocean realm, both natural and manmade. The goal is to provide more accurate representations of location and spatial extent, along with a means for conducting more complex spatial analyses of marine and coastal data by capturing the behavior of real-world objects in a geodatabase. The model also considers how marine and coastal data might be more effectively integrated in 3-D space and time. Although currently limited to 2.5-D, the model includes "placeholders" meant to represent the fluidity of ocean data and processes.” A working group consisting of ESRI personnel and a number of other GIS experts from academica and industry is currently working on developing the ArcGIS Marine Data Model. Storage format has not been defined at present.
Version:	Beta
Owner:	ArcGIS Marine Data Model Working Group
Maintenance:	The ArcGIS Marine Data Model Working Group will maintain
Relationship to others:	
Encoding:	ArcGIS 8 Geodatabase
References:	http://mcmcweb.er.usgs.gov/sdts/

	Spatial Data Transfer Standard (SDTS) - Part 1, Logical Specifications. Draft. DRAFT for Review, November 20, 1997. American National Standards Institute, Inc.

Name:	HydroML
Abbreviation:	HydroML
Purpose:	HydroML is an extension of the eXtensible Markup Language (XML) providing the Hydrologic Scientific Community with a standard definition of XML tags and concepts of structure to allow the definition of hydrologic information. The goal of HYDROML is to: • enable hydrologic data to be exchanged between persons and organizations, • enable hydrologic data to be exchanged between data collection devices and data bases, and enable hydrologic data to be served, received, and processed on the Web. HydroML can represent the following information: site information, computation instructions, corrections, ratings, shifts, time-series data including unit values and daily value statistics, peak flows, and site visit measurements.
Version:	Draft
Owner:	U.S. Geological Survey (USGS)
Maintenance:	USGS currently supports further development.
Relationship to others:	Relies on XML.
Encoding:	Encoded as XML (using a Schema).
References:	http://water.usgs.gov/nwis_activities/XML/nwis_hml.htm

Name:	Geography Markup Language
Abbreviation:	GML
Purpose:	GML 3 (Geography Markup Language) is developed by OGC (OpenGIS Consortium) as a standard for "modelling, transport and storage of geographic information". OGC is an international industry consortium of more than 250 companies, government agencies and universities participating in a consensus process to develop publicly available geo- processing specifications. GML provides an XML encoding for storage and exchange of geographic information, and incorporates both spatial and non-spatial properties. The current version, GML 3.0, can be used to represent a number of geographic objects and concepts, including features, coordinate reference systems, geometry, topology, time, units of measure and generalised values. Features can be simple features[8] (like points and lines), feature collections (sets of simple features), coverages (e.g. images) and observations (i.e. values captured with some form of instrument). GML 3.0 is conformant with several of OGC's Abstract Specifications and ISO standards for geographic information, among others the recently approved standard ISO 19115 Geographic Information – Metadata. Different metadata structures can be incorporated in GML 3.0, which specifies an abstract meta-data entity (*MetaData*) and a generic concrete meta-data entity

[8] Defined by OGC as "features whose geometric properties are restricted to 'simple' geometries for which coordinates are defined in two dimensions and the delineation of a curve is subject to linear interpolation".

	(*GenericMetaData*). This leaves application developers with two choices: (1) to define a new meta-data entity that extends the *MetaData* entity, or (2) to insert their metadata within a *GenericMetaData* tag. Both solutions provide a high degree of flexibility in terms of what metadata can be associated with the geographic feature or concept, ranging from a free-format text (encsapsulated in the generic meta-data tag) to a well-formed XML element that can be validated against an external schema.
Version:	3
Owner:	OpenGIS Consortium (OGC)
Maintenance:	CML is maintained and supported by OGC.
Relationship to others:	
Encoding:	XML
References:	http://www.opengis.org/techno/documents/02-023r4.pdf

A.2 Presentation standards

Name:	S52 - Specification for Chart Content and Display Aspects of ECDIS
Abbreviation:	S52
Purpose:	S52 is a specification for Chart Content and Display Aspects of ECDIS, and is also developed by IHO. S52 is a “Presentation Model” in the sense that it contains information about how the encoded geographic objects should be displayed on the screen, e.g. what colour, line thickness and type should be used for a line object.
Version:	Edition 5
Owner:	IHO (The International Hydrographic Organization)
Maintenance:	IHO will maintain and support S52 and the display standard for ECDIS.
Relationship to others:	Used by S57 3.1 (and the forthcoming S57 4.0)
Encoding:	
References:	http://www.ohi.shom.fr/ IHO Special Publication S52, edition 5 Dec 1996. Published by the International Hydrographic Bureau, MONACO.

A.3 Other standards

Name:	NMEA 0183 & 2000
Abbreviation:	NMEA 0183 & 2000
Purpose:	NMEA 0183 & 2000 is two standards defined by NMEA (National Marine Electronics Association), which are widely accepted by manufacturers of marine electronics. These standards describe how instruments like GPS, echosounders, gyro repeaters etc. should communicate with each other, and are also used by CTD systems, ADCP software etc. for collecting navigational data. Quoting from http://www.nmea.org/pub/0183/index.html: “The NMEA 0183 Interface Standard defines electrical signal requirements, data transmission protocol and time, and specific sentence formats for a 4800-baud serial data bus. Each bus may have only one talker but many listeners.”

	Quoting from http://www.nmea.org/pub/2000/index.html: "The standard contains the requirements of a serial data communications network to inter-connect marine electronic equipment on vessels. It is multi-master and self configuring, and there is no central network controller. Equipment designed to this standard will have the ability to share data, including commands and status with other compatible equipment over a single channel."
Version:	
Owner:	National Marine Electronics Association (NMEA)
Maintenance:	NMEA is maintaining and supporting these two standards.
Relationship to others:	
Encoding:	
References:	http://www.nmea.org/ http://www.nmea.org/pub/index.html

Name:	METOC Conceptual Data Model
Abbreviation:	JMCDM
Purpose:	JMCDM is a logical data model that integrates the geophysical data requirements of all DoD (Department of Defense) components. JMCDM and its supporting encyclopedia will serve as input for the expansion of the DoD Data Model. JMCDM is being developed by the Meteorology and Oceanography (METOC) Data Administration.
Version:	
Owner:	Meteorology and Oceanography (METOC) Data Administration , U.S: Department of Defense
Maintenance:	Data model is maintained and supported by METOC.
Relationship to others:	
Encoding:	
References:	http://pao.cnmoc.navy.mil/ http://pao.cnmoc.navy.mil/scripts/public_JMCDM/home_pwd.pl