Using XML Technology for Marine Data Exchange
A Position Paper of the MarineXML Initiative

Keiran Millard[1], Rob Atkinson[2], Andrew Woolf[3], Roy Lowry[4], Pieter Haaring[5], Francisco Hernandez[6], Torill Hamre[7], Dan Pillich[8], Holger Bothien[8], Ward Vanden Berghe[6], Ian Johnson[3], Brian Matthews[3], Hans Dahlin[9], Nic Flemming[9], Vladimir Vladymyrov[10], Bev Meckenzie[11], Alan Edwards[12], Willem van der Hoeven[5], Quillon Harphen[1], Gil Ross[13]

[1]HR Wallingford, [2]Social Change Online, [3]Central Laboratory of the Research Councils, [4]British Oceanographic Data Centre, [5]RIKZ, [6]VLIZ, [7]Nansen Environmental Remote Sensing Center, [8]Seven Cs, [9]EuroGOOS, [10]IOC/IODE, [11]UK Marine Information Council, [12]European Commission, [13]Met Office UK

January 2005

1. Abstract

The eXtensible Mark-Up Language (XML) developed by the World Wide Web Consortium can be used to develop a framework that improves the interoperability of data in support of marine observing systems. There have been several initiatives investigating how this can be achieved in practice, but to date there has been no strong direction that the marine community should take. This paper presents the findings of an international project tasked with providing such direction. In particular it provides 'pre-standardisation' advice and justification to the international marine community on the use of XML standards for data interoperability. Importantly this embraces the very broad extent of the marine community, covering the very many disciplines and domain areas that exist.

Keywords: XML, GML, ISO-TC211, Marine Data Exchange

2. Introduction

The huge diversity of data formats, proprietary data management systems, analysis packages, numerical models and visualization tools complicate the processing, management and accessibility of marine data. This diversity severely limits the multiple re-use of data and reduces our access to the data. Additionally the present complexity of accessing and integrating data makes it a tedious, often labour intensive to generate knowledge of marine processes and risks associated with activities in the marine environment.

Today, the marine community is in a situation where large volumes of data are collected by an increasing number of agencies and scientists for an increasing number of purposes. The desire to build regional or global databases, aggregating data from multiple sources is also increasing. This is evident at the European level through the development of INSPIRE[1] and GMES[2] and at an international level through GEO[3] and GOOS[4]. The need for a common data framework to enable this integration has now become an essential component of building our knowledge of the marine environment. It has been considered that the eXtensible Mark-Up Language (XML) developed by the World Wide Web Consortium can be used to develop such a framework. This would build on the approaches that have been used in other domain areas such as the geospatial (Lake R, 2004), chemistry, genetics and finance (Millard K, 2003).

[1] www.ec-gis.org/inspire/
[2] www.gmes.info
[3] http://earthobservations.org/
[4] ioc.unesco.org/goos/

MarineXML is an initiative under the governance of the IOC/IODE[5] of UNESCO[6] to improve marine data exchange within the marine community. The European Commission has provided a funding contribution to this initiative as part of its 5th Framework Programme to undertake a 'pre-standardisation' task of identifying the approaches the marine community should adopt regarding XML technology to achieve improved data exchange. This project (MarineXML EC) ran from February 2003 to January 2005. This position paper reports on the findings of this project. Other projects have contributed to MarineXML in this timeframe including the Study Group on XML (SGXML) of ICES/IOC[7], the UK NERC Data Grid Project[8] and the UKHO S-57 GML project [9]. All these projects worked closely together to reach consensus on using XML for marine data exchange. The authors accordingly believe that they have been successful in providing a level of pre-standardisation to inform a route-map towards an extensible framework of standardisation for data exchange in the marine community based on the adoption of ISO[10] and OGC[11] standards.

A key enabler of standardisation is effective engagement with the communities that stand to benefit. MarineXML has successfully attracted associates and spawned related activities based on the principles developed. The standardisation effort will need to effectively resource governance and dissemination of an evolving suite of components, and support delegation of responsibility for both development and dissemination of components to sub-communities. The intermediate results of MarineXML have been well received in presentations to OGC and there is little doubt that opportunities have been created for future collaboration.

3. Using XML for information exchange

XML is simply a method for encoding data developed by the World Wide Web Consortium[12]. It is a text-based system meaning that both humans and machines can understand it directly, and is 'self-describing' in so much as each data element can be traced to a definition. XML provides a formal grammar for text-based encoding which means that general parsing and querying tools can be used to process the data. XML also supports improved data integrity because XML editors can check documents satisfy defined constraints on content and structure. It is also extensible, which means that components can be linked together and schemas extended without compromising any existing data standards and hence minimises re-inventing the wheel. . As it has a formal schema language, XML Schema Definition Language[13], that is human readable (that is to a degree self-documenting), and interpretable for computer software.

XML Schema have been widely adopted in many communities to express a vocabulary (i.e. structure and content models) for information exchange using XML. XML Schemas are managed and maintained by the community defining them. For example SVG[14] for information on 2D-graphics, SOAP[15] for transporting messages, WSDL[16] for defining web services, GML[17] for geographic information. Where these communities provide 'stable vocabularies', services have been developed around them. For example transforming geographic data in GML to SVG such that text data can be displayed graphically, e.g. as a map.

---

[5] International Oceanographic Data and Information Exchange Committee of the International Oceanographic Commission of UNESCO
[6] United Nations Education, Scientific and Cultural Organisation.
[7] www.marineXML.net. ICES/IOC SGXML focused on the issue of parameter dictionaries.
[8] http://ndg.nerc.ac.uk
[9] www.ukho.gov.uk/b2b_gml_home.asp
[10] International Organisationfor Standardisation (www.iso.org)
[11] Open Geospatial Consortium (OGC) (www.opengis.org)
[12] Extensible Markup Language (XML) 1.0 (Second Edition) www.w3.org/TR/REC-xml
[13] XML Schema Definition Language (XSD) www.w3.org/TR/xmlschema-0/
[14] Scalable Vector Graphics(SVG) - www.w3.org/TR/SVG/
[15] Simple Open Access Protocol (SOAP) - www.w3.org/TR/soap/
[16] Web Services Description Language (WSDL) - www.w3.org/TR/wsdl
[17] Geography Mark-Up Language (GML) www.opengis.org/gml/

A key advantage of XML is its 'transoformability'. Transformation between XML documents complying to different schemas is commonly undertaken using the XML Stylesheet Transformation Language (XSLT)[18]. This releated standard allows the means to translate and format XML encoded data to be distributed to all the potential users. Binary formats require specialised software components (or detailed technical specifications) to be distributed, leading to a range of policy, security and efficiency issues, as well as chronic "version fatigue" where the cost of maintaining software versions forces activities to stagnate and eventually become obsolete.

4. Analysis of XML for data exchange across the marine community

The marine community embraces a range of disciplines (e.g. water quality, navigation science, dredging, meteorology, fishing and conservation) undertaking a range of activities (e.g. modelling, measuring, forecasting, archiving, dissemination) for a range of phenomena (e.g. benthic flora, waves, tides, salinity). This richness gives rise to a range of sub-communities, often defining their own standards or standards implementation to meet the needs of specific tasks.

MarineXML EC undertook a review and analysis of these data exchange standards in the marine community to understand what standards were being used for data exchange and where they were being applied (Hamre et al 2004). This was not limited to standards that involved XML encoding, but rather any situation where an information community has defined its needs sufficiently to contemplate a standardised data exchange. This provided a useful starting point for understanding the needs of the user community. It was particularly important to understand the governance mechanisms that were in place for the standards (i.e. which standards are likely to be around in the future). A strong governance model is also an indication of the value and importance of the standard to the community. The standards review identified around fifty standards in use in the marine community and these were documented as an ontology using RDF/OWL[19] (Johnson et al 2004). This ontology can be browsed and downloaded from www.marineXML.net.

It was especially useful to review the work of groups that had developed their own XML Schema for different aspects of marine data exchange. These include MMML[20], KeelyBricks (Keely et al, 2003), MBARI[21] , MIML[22], ESML[23], ANZLIC[24], OBIS[25]. None of these are inherently wrong; in fact they represent a good view on certain parts of the marine community. However none offer an extensible route for other communities to make use of their vocabulary. The Standards Review also showed that binary data formats such as NetCDF and GeoTIFF are widely used within narrowly defined communities able to agree on specific data products. Such encoding standards are indeed necessary for large data sets and any XML-based framework must take account and reference of this. Following this analysis related projects have been initiated to reconcile the NetCDF usage practices with scalable, extensible data exchange architecture.

The most important conclusion from the Standards Review was that there is a momentum from organisations such as IHO[26] and WMO[27] to adopt consistent approaches for the vocabulary of their data along the reference implementation of ISO Standards defined by the Open Geographic Consortium. Alignment with OGC has advantages for the marine community in that it provides not only a well-supported XML-encoding for geographic data (GML), it also offers services for accessing and processing geographic data such Web Feature Server, Web Map Server, Web Coverage Server, and Coordinate Transformation.

---

[18] XSL Transformations (XSLT) www.w3.org/TR/xslt
[19] Resource Description Framework / Ontology Web Language
[20] Model and Monitoring Data Mark-up Language (Bultman G et al, 2002)
[21] Monterey Bay Aquarium Research Institute- http://www.mbari.org/ssds/ReferenceDocuments/MOOSMetadataSchema.xsd
[22] Marine Information Mark-up Language – http://www.rdc.uscg.gov/iws/pubs/miml.pdf
[23] Earth Science Mark-Up Language - http://esml.itsc.uah.edu/index.jsp
[24] www.anzlic.org.au/infrastructure_metadata.html
[25] Ocean Biogeographic Information System - iobis.org//obis/obis.xsd
[26] International Hydrographic Organisation
[27] World Meteorological Organisation

5. OGC, Geography Markup Language and Feature Types

GML, initially a specification of OGC, is now being standardised as ISO TC211 project 19136. It provides reference implementation XML schemas for a range of conceptual models from other ISO projects[28], as well as other general-purpose building blocks for application schema complying with ISO 19103 ('Conceptual schema language'), ISO 19109 ('Rules for application schema') and ISO 19118 ('Encoding'). A GML application schema describes the logical structure and semantic content of a dataset. A key concept is that of the 'feature' – an 'abstraction of a real world phenomenon'. Typically, a data set contains one or more feature instances, with related information. Application schemas describe feature types, which may also be catalogued in a 'feature type catalogue'.

A feature may be classified as different feature in different domains. Even when you establish a common vocabulary within a domain, different people may disagree about the classification, e.g. the definition of 'swell waves'. Features can be classified in a hierarchy. The same feature name can have a different meaning depending on the domain. Conversely different feature names may have the same meaning in different schemas. The same physical object can have different names in different natural languages. Hence accuracy of a GML application is subjective to the community defining it. This reflects the nature of reality and the relationship of different communities to it.

6. IHO-S-57 Feature Types and Registers

The IHO has maintained a catalogue of hydrographic objects (features) and attributes as part of the IHO S-57 standard for many years. This is a very mature catalogue; however, it is only suitable for use within the context of the IHO S-57 standard edition 3.1 because it is closely bound to that standard. IHO has adopted a policy of building edition 4 of S-57 based on the ISO TC211 suite of standards (IHO 2003a, IHO 2003b). One benefit of structuring the S-57 object catalogue in conformance with the ISO standards is that it allows for re-use of the objects defined by IHO in a broader context, beyond the Electronic Navigational Chart. Another benefit is that it provides a mechanism to import additional features in future S-57 versions that have been defined by related communities

A register is simply a managed list and the ISO register standards provide a schema for the structure of the list and a process for the management of its contents. This increases consistency, reduces errors, and makes the list easily machine searchable. ISO has developed two separate standards for registration of geographic information. The first standard, ISO 19135, addresses procedures for the registration of items of geographic information. The ISO 19126 standard is built upon the procedures defined in 19135 and the data model described in 19110 and addresses specifically items in feature data dictionaries and feature catalogues.

The IHO registry is a compound registry that supports four independent but related registers that share a common code space but are managed by separate organisations[29]. Other similar registers of feature types could be added to this compound registry.

7. WMO and Feature Types

A WMO (World Meteorological Organisation) expert team with contributors from five European countries has been developing a meteorological community profile of the metadata standard ISO 19115. The team is also working on a feature catalogue (based on the ISO 19110 abstract model) which describes as feature types and coverages, the content of WMO bulletins and forecasts which are distributed in real time around the world. This work is part of WMO's FWIS development (Framework for a WMO Information System).

---

[28] E.g. ISO 19107 'Spatial schema', ISO 19108 'Temporal schema', ISO 19111 'Spatial referencing by coordinates', ISO 19123 'Coverages'

[29] S-57, OpenECDIS, AML and ICE. More information can be found at Sevencs.com

8. Principles for Developing an XML framework for Marine Data Exchange

From the research and analysis several principles for an XML-based framework for marine data exchange were established and these are presented in this section.

- Principle 1 – Marine Data exchange based on ISO19136 (GML Feature Types)
  The alignment between ISO and OGC on ISO19136 makes GML the clear (only) choice for developing an XML-based framework for marine data exchange. The adoption of this approach by significant organisations in the marine community such as IHO and WMO also reinforce this.

- Principle 2 – There is no single 'Marine Feature Type'
  From the standards review it is clear that it is not possible to have a single 'Marine Feature Type'. Given the diversity of the marine community, what is needed is a mechanism to represent what needs to be exchanged (Atkinson et al 2004). This was a challenge to early views held by the project on how an XML-based exchange framework could work. Whilst it was accepted that some degree of modularity was required, it was perhaps not conceived how broad this modularity had to be to represent the whole of the marine community.

- Principle 3 – Different sub-communities take responsibility for their feature types.
  The breadth of the marine community means that it becomes wholly impractical for any single organisation, such as IOC, to manage and maintain all the Feature Types that the marine community could require. Different parts and operations of the marine community need to subscribe to their own data standards as these are integral to their operations. These standards are often highly specialised to meet the requirements of particular services. For example, the highly specialised feature types for universal exchange of meteorological observations (SYNOPs, METARs etc.). The definition of such specialised feature types is rightly the domain of international organisations such as the WMO.

- Principle 4 – MarineXML Feature Type responsibilities
  Through its MarineXML initiative, IOC/IODE should act as the authority (registry owner) for the specialist feature types that are central to the marine community and the general purpose feature types for exchange within the marine community (e.g. to enable 'cruise' and 'met observations' to be effectively combined). These general-purpose feature types should be developed in liaison with key organisations in the marine community such as IHO and WMO, not least to combine resources for standards maintenance and update.

- Principle 5 – Wrapper for Legacy Data
  The XML based framework should not only encode text-based data as XML, it is also required to provide a wrapper to data that exists or is best delivered by binary encoded files.

9. General Purpose Feature Types for the Marine Community

Taking account of the above principles, MarineXML looked to develop and test some general purpose Feature Types for data exchange within the marine community. To undertake this activity MarineXML collaborated with the NERC Data Grid Project. This project was investigating the Feature Types necessary for data exchange between the oceanographic and meteorological communities.

This timely collaboration enabled a very rigorous development and documentation of GML Application Schema to encode these Feature Types. The resultant schema at the time of the testbed was referred to as the 'NDG data model', although it is now released as Climate Science Mark-Up Language (CSML) (Woolf et al, 2004a). In addition to modelling various climate science data types, CSML also provides a wrapper mechanism to encapsulate the representation of those data objects in file-based storage artefacts such as NetCDF and GRIBB.

Feature Types in CSML are defined primarily on the basis of geometric and topologic structure, and not the semantics of the observable or measurand according to the following principles (Woolf, 2004b)).

- Offloading semantics onto Coordinate Reference System
  If two features would 'look the same' structurally, apart from the underlying coordinate reference system, then they are modelled as the same feature. e.g. a vertical sounding radar is a sequence of profiles in time with a series of vertical heights as the spatial domain; a scanning radar is a sequence of sloping profiles (at fixed elevation) in time and in azimuth with a series of ranges as the spatial domain. These are modelled as the same feature type, just with different underlying coordinate reference systems.

- Offloading semantics onto physical parameter type
  If two features would 'look the same' structurally, apart from the physical parameter type, then they're modelled as the same feature. e.g. a vertical wind-profiler is plotted with wind barbs, while a temperature sounding is plotted as a line– these are regarded as the same feature type, distinguished only by the physical parameter type.

- Sensible plotting as discriminant
  A principle to suggest a workable minimum granularity of feature type definition is to use the requirement for 'sensible plotting' as a discriminant. That is, there should be sufficient detail in, and sufficient difference between, feature types to enable unsupervised 'sensible plotting' by an appropriate piece of software.

The Feature Types described by the CSML Model are presented in Table 1 below and shown conceptually in Figure 1.

CSML feature type	*Description*	*Examples*
TrajectoryFeature	Discrete path in time and space of a platform or instrument.	ship's cruise track, aircraft's flight path
PointFeature	Single point measurement	raingauge measurement
ProfileFeature	Single 'profile' of some parameter along a directed line in space.	wind sounding, XBT, CTD, radiosonde
GridFeature	Single time-snapshot of a gridded field	gridded analysis field
PointSeriesFeature	Series of single datum measurements	tidegauge, rainfall timeseries
ProfileSeriesFeature	Series of profile-type measurements	vertical or scanning radar, shipborne ADCP, thermistor chain timeseries
GridSeriesFeature	Timeseries of gridded parameter fields.	numerical weather prediction model, ocean general circulation model

Table 1 CSML Feature Types

The feature types do not carry any explicit topologic descriptions. For instance, both a scanning radar and a vertical sounding radar are modelled with the same feature type (CSMLProfileSeries). The spatial geometry of the first may be modelled topologically as a series of curves connected at an origin node, while the second is a single curve. Similarly, the topological relationship between a series of marine CTD casts and the associated ship track is left implicit in the CSMLProfileSeries feature type. GML provides rich schemas for describing topology – a requirement may arise in the future to add this information to the CSML Data Model.

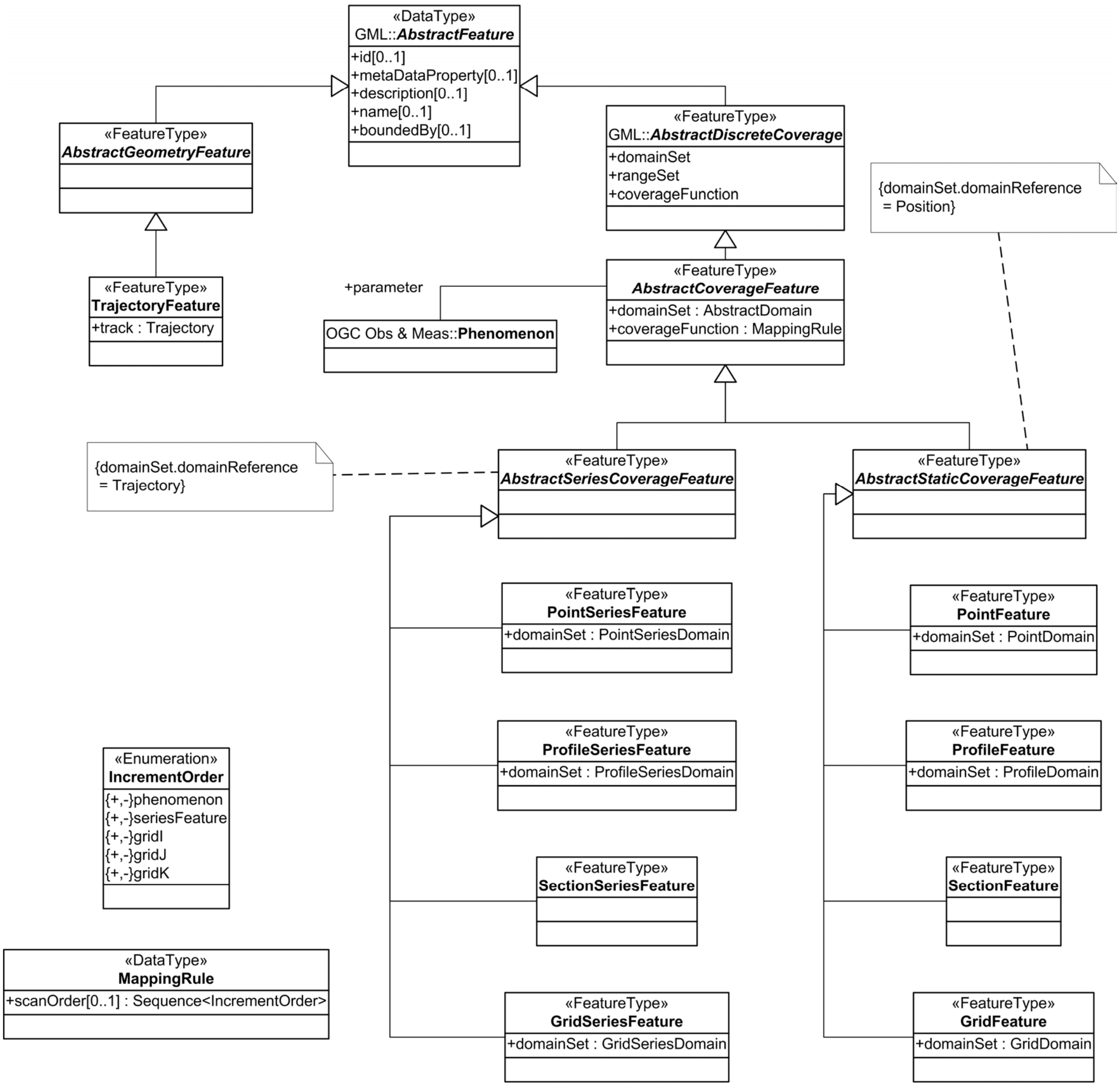

Figure 1 Conceptual Model of CSML Feature Types

10. Metadata and Feature Types

Appropriate structure and content models for metadata vary dependent of the service requirements of the user, e.g. discovery, browse, use (O'Neill et al, 2004). CSML is a schema for 'use' metadata. ISO19115/19139 profiles do the same job for the more familiar 'discovery' metadata. There are two approaches to developing a discovery metadata profile. One is to populate 'use' data with additional information needed for 'discovery'. The other is to use separate schemas for 'use' and 'discovery' metadata with a link between them. MarineXML adopted this 'separation' approach.

GML application schema provides the mechanism to embed or link to an XML encoding of an ISO19115 metadata profile. It should be noted there is some information that is required both for discovery and use (such as spatio-temporal coverages), so there will inevitably be some overlap between CSML and 19115/19139 records. Consequently, lightweight discovery metadata could be harvested directly from CSML, but the result would not meet the requirements of 19115/19139. Discovery records according to the standard required can be produced using XML transformation technology on the superset of 'discovery' metadata.

11. Test Bed Implementation

The testbed was established to appraise at a practical level how the CSML Feature Types functioned to provide data exchange in the marine community. The test bed was based around the tools and standards used by the navigation community. This includes the S-57 standard for exchange of hydorgraphic data and a viewer for these electronic navigation charts conforming to this standard called SeeMyDENC[30]. Specifically, the aim was to demonstrate how this display tool for navigation data could also display other data that did not conform to the S-57 standard Pillich et al 2004).

Source data was provided for domains covering marine biology[31], marine remote sensing, measured Hydrodynamics and modelled hydrodynamics. These communities have their own, often *ad-hoc* vocabularies for expressing the content of their datasets. This data was translated to XML conforming to general-purpose Feature Types described by CSML. The software in SeeMyDENC was extended to parse these CSML Feature Types and display them accordingly. This is shown conceptually in Figure 2 below.

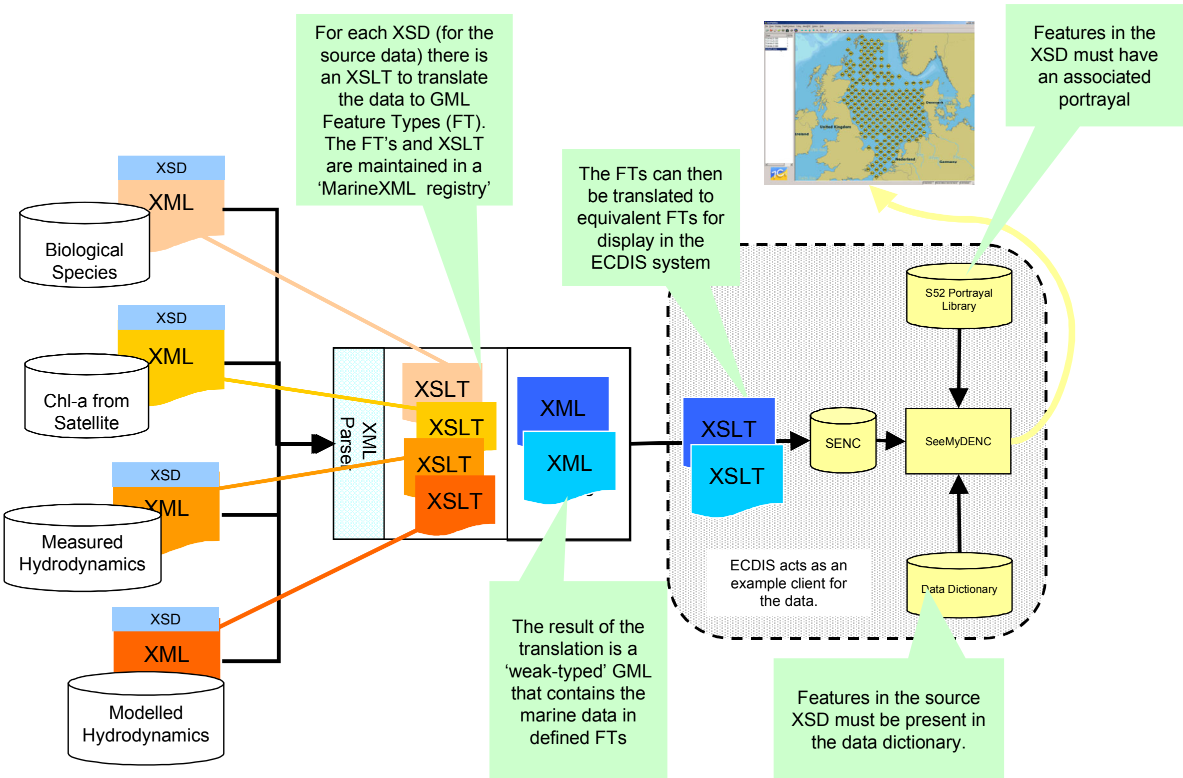

Figure 2Test Bed Conceptual Diagram

CSML proved a robust recipient for the data from each community. It produced economical files with few redundant elements, striking about the right balance between weak and strong typing. As CSML is based on geometric and topologic structures, the mappings worked less efficiently for data structured around groupings of phenomena. Such mappings were possible but resulted in more unwieldy results. Most discussion was not on the Feature Types themselves, but on ensuring equivalent naming of the equivalent phenomena. For the test bed a translation of the CF Standard Names Table[32] into GML dictionary format[33] was used for physical phenomena. For

[30] See www.sevencs.com for more information on SeeMyDENC
[31] Biological data collected by the North Sea Benthos Project of ICES. The data has several abiotic parameters that can not by described in the Darwin schema used by Obis
[32] hwww.cgd.ucar.edu/cms/eaton/cf-metadata/standard_name.html
[33] hndg.nerc.ac.uk/csml/CFStandardNames.xml

biological data, phenomena were taken from OBIS[34] with a webservice established to the OBIS taxonomic database to provide definitions.

Standard dictionaries of both phenomena (what was observed) and units in which the phenomena have been measured are required if automated parameter interoperability is to be realised. Whilst the test-bed proved the concept of using a GML encoding of CF standard names, this is not an extensible option. The final report of SGXML (Isenor and Lowry, 2005) recommended that the oceanographic phenomena dictionary developed by BODC[35] be adopted as the marine oceanographic community standard.

The test bed worked from a functional perspective, but it should be remembered that it was established to prove the concept that data providers and data services can make use of one another by subscribing to common data vocabularies. These data vocabularies can be expressed as a large number of strongly typed features or a small number of weakly typed features. What the test bed has shown is that the seven feature types represented by CSML provide a good basis for the exchange of marine data.

12. Next steps and End Game

The completion of the work of MarineXML EC and the final report of SGXML provides an important breakpoint to consider the next steps for improving marine data exchange. The 'end game' is for MarineXML to provide a definitive set of resources such as schemas, manuals and software through its website MarineXML.net. This would include both directly hosted resources and links to resources hosted and maintained elsewhere. The links between these resources would be managed as a ontology to show what is available, how it is being used, what organisation are engaged. To achieve this end game a number of next steps can be identified.

- <u>MarineXML Governance Model</u>
  MarineXML is an initiative under the governance of IODE. Contribution to MarineXML has to date largely been on an informal basis and goodwill links between projects. This structure needs to be adapted to make provision for future projects, howsoever funded to contribute to MarineXML. Recommendations have been made to IOC for a strengthened role for the IODE project office in this capacity.

- <u>Standards Registry Deployment.</u>
  It is essential for the uptake of any standard that a robust governance structure is in place. The IOC and IHO have proposed a collaboration to deploy an ISO TC211 compliant standards register. Feature types defined by CSML could, amongst others, be lodged in this register. The IHO have produced a prototype register[36] that covers their four main catalogues, but establishing a formal collaboration with IOC is essential if IHO wish to widen the uptake of their data beyond the navigation community. As the IHO register is a compound register serving four feature type catalogues, the mechanisms adopted for managing the relationships between feature types and their attributes and enumerates is a key technical challenge to be understood.

- <u>Rollout of CSML</u>
  This project established that Climate Science Mark-Up Language (CSML) provided a set of generic Feature Types that work as a good 'general purpose' exchange protocol for marine data. CSML was only released this year and MarineXML has been the first project to test these schemas. The marine community would like to be engaged with NERC in the wider uptake and deployment of this mark-up language (Matthews and Woolf, 2005).

---

[34] http://www.iobis.org
[35] www.bodc.ac.uk/data/codes_and_formats/parameter_code
[36] http://www.ukhoftp.gov.uk/ provides a prototpye version

- Increasing Uptake, Adoption and Understanding of Standards to deliver services
  Building on this pre-standardisation work, education and awareness is required to enable the marine community to benefit from improved data exchange. This needs to provide the marine community with a toolkit for developing and using GML Application Schema, in addition to other OGC specifications - and the costs and benefits of doing this. At a European level, such knowledge is essential for the development of GMES and INSPIRE and for this reason the European Commission has provided funding to a project called MOTIIVE[37] that involves several organisations presently engaged with MarineXML.

- Informing the Development of GML
  As GML is used more in the marine domain, it will become important for the marine community to directly inform the development of GML, particularly where its specification does not meet the marine communities needs. The work of MarineXML EC has been presented at OGC Technical Committee through the Natural Resources Working Group meetings and has been well received. The Motiive project mentioned above provides specific liaison activities with OGC, including the option to establish a working group or special interest group. Liaison with the SEEGrid[38] project exploring aspects of GML semantics and implementation architectures is now formalised. . This provides direct access to members of the ISO GML editing team as well as a community exploring international data standards in the geosciences area.

- Integration of emerging GRID data processing frameworks and legacy binary data exchange systems.
  The NDG data model is an output of a larger ambition to support seamless data access as a foundation for computing GRIDS able to run processes that integrate a wide variety of data. The GRID computing community is undergoing a similar standards convergence experience and the proposed approach should allow harmonisation with future computing infrastructures. At the same time, exploiting the MarineXML approach to add value to existing legacy data transfer arrangements (OpenDAP and NetCDF typically) will provide a integration path. This is being pursued as an activity under AUKEGGS[39] collaboration and will provide an input to the MOTIIVE project.

- Dictionary Development.
  One of the more powerful features of the BODC Parameter Dictionary is that the phenomena descriptions are built by automated concatenation of the elements of a semantic model. This is currently encoded in a relational database, which incorporates significant knowledge as implicit relationships. This resource is currently inaccessible, but could be liberated by encoding in an RDF-based schema such as SKOS[40] There is also a requirement for a standard units dictionary more suited to the needs of the oceanographic community (particularly the chemical and biological disciplines) than UDUNITS[41].

13. Conclusions

MarineXML is an initiative of the IOC/IODE of UNESCO to improve data exchange within the marine community. The European Commission has provided a funding contribution to this initiative to undertake a 'pre-standardisation' project on the approaches the marine community should adopt regarding XML technology. We believe that the results from this project are sufficient to claim that this pre-standarisation has been achieved.

---

[37] MOTIIVE will commence work in 2005 and results will accessible via the MarineXML.net website
[38] www.seegrid.csiro.au
[39] A collaborative mechanism between Australian and UK scientists
[40] http://www.w3.org/2004/02/skos/
[41] http://my.unidata.ucar.edu/content/software/udunits/index.html

Firstly, the alignments between OGC (Open Geographic Consortium) and ISO on the development of a reference standard for feature type encoding using XML means that many communities in the broad marine domain are now considering the deployment of data standards based around Feature Type catalogues. This includes key communities such as the IHO and the WMO. Accordingly any deployment in the marine domain needs to be based on this model. Secondly this project has demonstrated how weakly-typed (i.e. generic) feature types represented as GML application schema can provide an effective interchange between the data formats used across the marine community. This includes the communities engaged in physical and biological oceanography, numerical modelling and remote sensing.

Taking this platform as point of departure, MarineXML has provided a number of next steps to be taken to ensure the wider adoption of open exchange standards for marine data and has identified resources to undertake them. Our continued aim is for the deployment of marine services that are not limited by the short-term technical considerations of users and providers.

14. References[42]

Atkinson, R., et. al., 2004: Next steps to interoperability – Mechanisms for semantic interoperability, EOGEO Workshop, University College London, 23-25 June 2004 [Online:http://y.eogeo.org/files/eogeo2004/eogeo-2004-fullatkinson.ppt, 8 November, 2004].

Bultman G, Tennison J, Heijmans J, 2002, Model and Monitoring Data Mark-Up Language, Ministerie Van Verker en Waterstaat, Rijkswaterstaat. Produced by Xi Advies bv, www.xi-advies.nl

Hamre T, Millard K, Matthews B. , 2004. *Analysis of marine and XML data standards* Deliverable D6 of the MarineXML project, www.marineXML.net, MarineXML_D6_v5.pdf, 17/08/2004.

IHO 2003a, TSMAD Sub WG on S-57 Extensions Draft Report (Version dated 14 October 2003) 29 September – 1 October 2003, Wollongong Australia

IHO, 2003b, Register Structure and Registration Process for an IHO Object Register. WI 2.1 Resolution Sept 0c

Isenor, A.W. and Lowry, R.K. (2005). Final Report of the ICES/IOC Study Group on the Development of Marine Data Exchange using XML. Defence R&D Canada – Atlantic Technical Report TR2005-005.

Johnson I,Matthews B, Atkinson, R, Millard K , 2004. *Ontological model of data standards used by the marine community*. Deliverable D7 of the MarineXML project, www.marineXML.net, MarineXML_D7_vX.pdf, Date.

Keeley, R, Isenor A, Linguanti, J 2003, XML Bricks, 20th January 2003 http://ioc.unesco.org/marinexml/contents.php?id=19

Lake, R, 2004. Geography Mark-Up Language – Foundation for the Geo-Web, Wiley, ISBN 0-470-87154-7

Matthews B and Woolf A, 2005 – *GML Application Schema for the interchange of marine data – An application of CSML*. Deliverable D11 of the MarineXML project, www.marineXML.net, MarineXML_D11_vX.pdf, Date

Millard et al, 2003, MarineXML – using XML technology for marine data interoperability *Proceedings of The Colour of Ocean Data*, IOC Workshop Report No 188, VLIZ Special Publication No 16 Brussels, Belgium 25th-27th November 2002 pp163-175

[42] The final references for deliverables from MarineXML EC is subject to minor changes whilst they are being finalised.

O'Neill K, Cramer R, Gutierrez M, Kleese van Dam K, Kondapalli S, Latham S, Lawrence B, Lowry R, Woolf A. 2004, Googling secure data *Proceedings of the U.K. e-science All Hands Meeting, 2004*. S.J.Cox(Ed) ISBN 1-904425-21-6. http://ndg.nerc.ac.uk/public_docs/AHM-2004-BNL-FullVersion-Submitted.pdf

Pillich D, Bolthien, H, Haaring P, Harphen Q, 2005. *Testbed demonstration of marine data exchange through GML application schema.* Deliverable D10 of the MarineXML project, www.marineXML.net, MarineXML_D10_vX.pdf, Date.

Woolf et al, 2004a Climate Science Modelling Language: Standards-Based Markup For Metocean Data, Proceedings of the American Met Society 85th annual meeting in San Diego 10-13 Jan 2004

Woolf, A, 2004b, Climate Science Mark-Up Language, Users Manual, v1.0, http://ndg.nerc.ac.uk/csml