International Journal of Spatial Data Infrastructures Research, 2007, Vol. 2, 132-162

Developing Feature Types and Related Catalogues for the Marine Community - Lessons from the MOTIIVE project*.

Keiran Millard (k.millard@hrwallingford.co.uk)[1], Rob Atkinson (rob@socialchange.net.au)[2], Andrew Woolf (a.woolf@rl.ac.uk)[3], Kristin Stock (kstock@socialchange.net.au)[2, 11], Roger Longhorn (ral@alum.mit.edu)[4], Chris Higgins (chris.higgins@ed.ac.uk)[8], Mark Small (m.small@ed.ac.uk)[8] Sander Hulst (hulst@argoss.nl)[5], Torill Hamre (torill.hamre@nersc.no)[6], Maria Ferreira (m.ferreira@eucc.net)[7], Irene Lucius (i.lucius@eucc.net)[7], Peter Breger (peter.breger@cec.eu.int)[9], John Pepper (john.pepper@ukho.gov.uk)[10], Dominic Lowe (d.lowe@rl.ac.uk)[3], Quillon Harphen (q.harphen@hrwallingford.co.uk)[1,] Stephen Wells (s.wells@hrwallingford.co.uk)[1]

[1]HR Wallingford, [2]Social Change Online, [3]Central Laboratory of the Research Councils, [4]IDG Ltd, [5]Argoss, [6]Nansen Environmental Remote Sensing Center, [7]EUCC, [8]EDINA, University of Edinburgh [9]European Commission, [10]UK Hydrographic Office, [11]University of Nottingham

Abstract

MOTIIVE (Marine Overlays on Topography for annex II Valuation and Exploitation) is a project funded as a Specific Support Action (SSA) under the European Commission Framework Programme 6 (FP6) Aeronautics and Space Programme. The project started in September 2005 and finished in October 2007. The objective of MOTIIVE was to examine the methodology and cost benefit of using non-proprietary data standards for data exchange amongst and between different communities as defined by the INSPIRE data theme annexes. Specifically it embraced the harmonisation requirements between the INSPIRE Annex II data theme 'elevation' (terrestrial, bathymetric and coastal) and INSPIRE Annex III marine thematic data for 'sea regions', 'oceanic spatial features' and 'coastal zone management areas'. This was examined in context of the requirements for interoperable information systems as required to realise the objectives of GMES (Global Monitoring for Environmental Security) for 'global services'. The work draws particular conclusions on the realisation of Feature Types (ISO 19109) and Feature Type Catalogues (ISO 19110) in this respect. More information on MOTIIVE can be found at www.motiive.net

Keywords: ISO 19110, ISO 19109, Marine Data, Coastal Data, INSPIRE, cost-benefit, user requirements.

1 INTRODUCTION

The EC's INSPIRE Directive[1] proposes that harmonised 'Annex I' core spatial data will underpin the further integration and harmonisation of 'Annex II' and 'Annex III' thematic data[2]. Implicit in this statement is that common approaches must exist to model data across Annex III themes such that it can be combined in a meaningful way with data from Annex I and Annex II. The existence of Open Geospatial Consortium (OGC) and International Organization for Standardisation (ISO) TC 211 non-proprietary spatial data standards and interoperability tools is expected to help foster data integration at lower cost than previously experienced when using multiple data sources in integrated data projects.

A key objective of MOTIIVE was therefore to examine this expectation for a selected set of core and thematic data, while fully documenting the processes, procedures, barriers and resource requirements involved in creating, then using, non-proprietary geographic information standards and new interoperability tools to aid such data harmonisation. The cost-benefit analysis included as a key component of MOTIIVE explored the implications of harmonisation across core and thematic data areas as well as within thematic data groups.

This paper provides an overview of the outputs of MOTIIVE. The purpose of this paper is to inform the reader of the scope of MOTIIVE and highlight the main conclusions from the project. Although MOTIIVE considers the issues of data harmonisation from the perspective of the marine community; the scope of the paper is of relevance to any community engaged in intra-domain or inter-domain data harmonisation. This paper covers the four main areas of the project's wide ranging remit:

- The requirements of the broad marine user community for data harmonisation and system interoperability. This also includes: Regional and local Coastal managers, Mapping organisations creating integrated land-sea products, GMES Service Element projects (GSE) and the World Meteorological Organisation (WMO).
- The cost-benefit to the marine community in moving towards open standards to develop a spatial data infrastructure.
- The derivation of GML (Geographic Mark-Up Language) Application Schema of the Feature Types needed within the marine community for data exchange using the ISO TC21 suite of standards.
- A reference implementation of a Feature Type Catalogue. For communities to use to publish their Feature Types and expose them for re-use.

[1] http://www.ec-gis.org/inspire/
[2] INSPIRE Drafting Team "Data Specifications" – deliverable D2.3: Definition of Annex Themes and Scope - D2.3_v2.0.doc

Through these activities MOTIIVE has striven to ensure alignment with the Open Geospatial Consortium[3] (OGC) and ISO TC211[4] on a 'sustained' home for the approaches outlined above. This resulted not in a specific *Marine Working Group* as the MOTIIVE issues are broader than just the marine community. A *Domain Modelling* subgroup under the auspices of OGC was therefore proposed for any 'community of interest' that wishes to model data and deploy services in a consistent manner. The procedures for developing GML Application Schema of the Feature Types has subsequently formed part of the INSPIRE methodology for Data Products Specifications[5].

2 COMMUNITY NEEDS FOR HARMONISATION

The marine community is very diverse covering areas such as meteorology, navigation, water quality, biodiversity and coastal defence amongst others. This chapter reviews the harmonisation needs of four domains of the marine community with which there has been dialogue in MOTIIVE. This includes: regional and local coastal managers, mapping organisations creating integrated land-sea products, GMES Service Element projects and the World Meteorological Organisation. The dialogue with these communities was used to derive a general 'use case' which informed the realisation of the MOTIIVE data specifications.

2.1 Coastal Management Community

MOTIIVE reviewed the information needs across a number of completed and ongoing coastal management activities (Ferreira et al, 2007). These activities included specific research and development projects as well as operational 'day to day' activities of public bodies responsible for marine and coastal management as follows:

- EUROSION GIS database
- Mapping European seabed habitats - MESH project
- DeCOVER - procedures to update land cover data in Germany
- A national information system in the Netherlands
- A GIS infrastructure for coastal law enforcement in Portugal
- Regional data & information GIS system in Catalonia and the DEDUCE project experience.

The review consisted of analysis of literature and also face to face interviews. This concluded that the information needs of coastal managers and professionals

[3] http://www.opengeospatial.org/
[4] http://www.isotc211.org/
[5] http://www.ec-gis.org/inspire/reports/ImplementingRules/inspireDataspecD2_6v2.0.pdf

implementing EU and national polices and coastal legislation can be grouped into three driver activities. These are as follows:

- Monitoring, reporting and surveillance (e.g. real-time monitoring such as providing shipping guidance and routine monitoring such as water quality measurements)
- Design and policy development (e.g. survey and plans for new coastal infrastructures such as hard defence systems or a new harbour)
- Emergency operations/disaster response (e.g. storm surge prevention and sea level rise impact responses)

Within each of these sectors there is a growing need for better and harmonised data and information for the integrated management of the coastal and marine environment. Better ecosystem and seabed information is needed to spatially plan marine protected areas, to regulate marine resource exploitation, including extractive and shipping industries. Regarding the stakeholders involved in shipping and management of low-lying areas, forecasts of waves and surface currents is essential. Furthermore, those developing and/or protecting coastlines need tidal patterns, erosion rates and sea-level rise predictions, to name a few.

The common objective to better facilitate sharing of marine and coastal information can be summarised in two main user aspirations:

- 'Seamless discovery' – the user would like to be able to search widely, at different levels and access all that exists. This entails the needs for agreements in terms of data descriptions, common metadata definitions, common protocols, data access and sharing policy.
- 'Seamless use' – the user would like to easily identify the data available and to easily find what fits the purpose of his/her work. This entails the needs for agreements on definitions of data, common models, measurements, and data sharing principles.

2.2 Meteorological Community

MOTIIVE partners contributed to a workshop sponsored by the UK Met Office to understand the use and adoption of ISO TC211 Standards to the meteorological community (UK MetOffice, 2006). Operational Meteorology is a subject domain with an international scope, and data interoperability between practitioners is crucial. However, existing software tooling is generally bespoke and doesn't take advantage of 'off-the-shelf' technologies, and so is difficult to maintain and extend. Accordingly, the UK Met Office is investigating appropriate technologies with the aim of improving the reusability of both software components and data in its information-processing infrastructure.

A number of specific challenges were considered at the workshop:

1. How to create data product and service specifications for Met Office implementations, such as internal processing chains,
2. How to take these, and predict, test and drive standardisation of data products and service specifications for international exchange of data, and
3. How to maintain metadata at all stages in the processing chain.

The initial starting position was that the above challenges were essentially unsolved, but that the ISO (19100 series) standards provided a theoretical and governance basis for development of appropriate data standards, and that exploiting a Service Orientated Architecture (SOA) based on the Open Geospatial Consortium (OGC) conceptual formalisms provided a practical implementation route.

This output from the workshop was summarised in eleven recommendations which can be categorised in two main areas:

- GML Application Schema Development. The standards used to model geospatial data. Specifically the tools and standards that can be used to develop data models and how these can be realised in 'application schema' to enable data models to be incorporated into processing systems. This examined the work of the Climate Science Modelling Language and the OGC Observations and Measurements Recommendations. These two topics are covered in more detail in Chapter 4.
  The recommendations from the workshop stressed the need for the data modelling within a community to be guided by the governance remit of that community. In the context of the workshop this would mean that meteorological community should only seek to develop concepts related to meteorology. The meteorological community does not need to develop generic concepts such as geospatial or temporal referencing, or generic models for an observation or measurement; these should exist outside the remit of any meteorological office.

- Achieving Interoperability. The methods and approaches to making use of application schema in processing systems. This considered 'what can be done' within the existing standards landscape, and what changes are required to improve this for the meteorological community. This examined the use of registries to store Feature Types (Feature Type Catalogues). Feature Type Catalogues are discussed in more detail in Section 5.

The recommendations from the workshop highlighted that having the ability to query a feature type in a standard way and understand what can be done with it is a key consideration for the implementation of service chaining. This is termed as the 'processing affordance' of a Feature Type, but there was no clear approach as to how the 'affordance' of a Feature Type could or should be declared within a Feature Type Catalogue.

2.3 GMES Service Elements

Service chaining is an integral part of the GMES service elements (GSE), binding data providers to value-added services. In this respect the requirements of the GSE are very similar to those of the international Met Offices. The conclusions reached from the consultation with the Met Offices we discussed with the WIN project (www.win-eu.org). WIN is providing the architecture for the GSE MARCOAST (www.gmes-marcoast.com). WIN agreed that service binding at the feature-type level (the 'processing affordance of feature types') was a natural progression towards realising the concept of a service 'mash-up', enabling communities of users to derive and deliver their own services from 'base' services.

2.4 Hydrographic Community

MOTIIVE examined the problems of producing chart products that spanned the land-sea interface in the UK (Longhorn et al, 2007). Specifically it considered the data integration tasked faced in the Integrated Coastal Zone Map (ICZMap) Project and the (subsequent) many processing steps implemented by the UK Hydrographic Office (UKHO) to overcome the barriers to integrating land and sea data at the coastal zone. In retrospect it became clear that these workflow are directly analogous to the steps in the INSPIRE Guidelines for establishing a Data Product Specification[5].

The central problem is that, throughout the world, hydrographic and topographic data has been collected by two different organisations for different purposes (sea navigation and land management / urban planning) and in different data models, including most vexingly, a different vertical datum. Establishing common feature catalogues helps the process as it facilitates agreements that the same 'thing' is being referred to. The land-sea harmonisation efforts in the UK identified that ~20% of all Feature definitions needed for land-sea harmonisation were duplicated between organisations with responsibilities for mapping and charting the coastal zone.

Agreeing common Features is a difficult process and requires above anything else reaching political consensus on the definition of the 'thing' and who maintains this definition. In addition, even with commonly agreed Features, the datasets containing these features may be subject to different quality control. The last point is important as it can result in Features being 'in the wrong' place as a result of human error that may go unchecked. It is identifying and correcting these ‘unknown’ errors that have the greatest impact on the harmonisation process making it labour intensive, even with the aid of mature practices, methods and models.

There is no doubt that had common harmonisation frameworks been in place in the past, the efforts of ICZMap and UKHO would have been easier than today. Now there is a large cost associated with re-engineering datasets to establish an integrated land-sea product. The benefits of such harmonisation will not be realised by organisations tasked with charting the land and sea respectively, but by downstream users who will save time using a single, harmonised dataset, and will use the same harmonised dataset, over time.

2.5 MOTIIVE “Use Case”

From the analysis of the specific needs from the above communities, a generic ‘high-level’ scenario was established that could be used to drive the standardisation process in MOTIIVE (Woolf & Millard, 2007). This high-level scenario is presented in Box 1 and embraces the general needs of the ‘coastal manager’ for ‘integrated information’ and operational service provision across several actors as required by Met Offices and the GSE. The following are key elements of this scenario:

- Data discovery. The user is able to find what they are looking for
- Service discovery (by interface or type). The user is able to discover services that exist, either by what the service does, or by what the service requires as input.
- Querying a feature-type catalogue (for inheritance and behavioural properties). The user is able to determine the Features represented in the dataset; how they relate to other Features and what operations can be performed on them.
- Data processing, binding service to data. Related to the above, the user is able to actually execute data processing against data sets in an automated manner
- Portrayal catalogue. Extension of data processing case, but for data visualisation

Box 1: MOTIIVE high level scenario

Jane is contributing to the development of an integrated coastal management plan and needs access to a range of data concerning conditions in the littoral zone and offshore. Jane's current concern is to investigate sediment transport and coastal erosion in the UK's Thames estuary.

Jane is working in an era of a harmonised spatial data infrastructure and is able to access on-demand a range of data across the internet. She wishes to evaluate a variety of tidal models available to her by validating against historical tide gauge measurements. Helga (NO) and Willem (NL) both run tidal models and wish to make their datasets available for re-use – they employ standards-based interoperable web services for this purpose, and these are amongst the datasets available to Jane for evaluation.

Data from the selected model is used to generate plots of current and other tidal parameters. None of the data suppliers provide dedicated portrayal services for their data, but a third-party portrayal service is available for a range of compatible 'base' feature types. Moreover, another service provides feature-type transformation (defined

3 COST BENEFIT

The use case clearly states that the broad coastal and marine community would benefit from enhanced interoperability between data sets and data processing systems based on open standards. One of the main issues realising this interoperability is, however, the cost in deploying new data systems and migrating from existing data management frameworks. For this reason MOTIIVE considered what approaches could be used to determine if the benefits where in fact greater than the costs.

Many cost-benefit studies have been conducted in relation to geospatial information systems (GIS) projects and technology, beginning as far back as the early-1980s. Far fewer have been conducted looking specifically at quantifiable benefit for implementing spatial data infrastructure (SDI). Table1 shows the range of typical studies in time, geography, nationality, sector and diversity of these 'value of GI' type studies. Many of these studies investigated cost-benefit only for single industry or government sectors or agencies, or types of spatial data technology or applications. Others tried to look at a wide range of sectors and regions, from national to trans-national. Some studies considered only government (public sector) data, while others tried to factor in the impact of private industry on SDI strategies and the impact of those strategies on private industry.

Following a thorough review of these different types of cost-benefit analysis (CBA) methodologies, MOTIIVE found that the Multi-Criteria Approach came closest to matching the decision-making requirements for projects where the benefits could not easily be given a money value. This approach also relies on assumptions made during the analysis process no more or less extensively than in standard financially oriented CBAs. Yet it permits stakeholder involvement from the outset, to the extent possible within the resource limitations (time, experts, modelling expertise, etc.) of those doing the analysis. That is, the proposed methodology can be used for relatively simple assessments or for more complex analyses, depending upon how much effort is put into the initial analysis framework development.

During the development of this methodology, MOTIIVE took part in a special workshop conducted by the SDI Unit of DG JRC in Ispra, in January 2006. This 2-day workshop focused specifically on CBA issues and metrics, i.e. Return on Investment (ROI) and Net Present Value (NPV), for a range of methodologies and practices. Participants included the developers of GeoVMM (Geographic Value Measuring Methodology) and a more complex simulation modelling approach – NB-Sim – developed for US Geological Survey (USGS). Both methodologies had been used in the USA by large government agencies, in the case of GeoVMM to compare the cost-benefit of two large geospatial information projects for NASA, and for an the entire US national SDI "The National Map" programme, using NB-Sim. One of the conclusions of the workshop was that no one methodology is 'best', partly because of the wide range of types of projects and programmes for which a CBA might be conducted, with widely varying goals and expected success metrics.

MOTIIVE proposed a hybrid methodology based on two pre-existing, proven methodologies. One of these is GeoVMM, mentioned above, a form of Multi-Criteria Analysis approved for use by the US government in federal project justification, which combines monetary cost estimates with value-based (non-monetary) benefits scores. Two success metrics result – an ROI based on "saving to invest" derived from the cost figures and a 'value' score that is both risk-adjusted and weighted, in consultation with experts and stakeholders, using a decision framework agreed in the initial stages of the project by stakeholders. In GeoVMM, benefit or value metrics are developed in five groups, including: direct user value/benefit, social value, institutional operational benefits, institutional financial value, and stragetic and policital value. Each main category is further broken down into from two to five sub-categories, yielding 18 measures of value or benefit in all. The full methodology is described in Booz Allen Hamilton (2005).

Secondly, the cost structure from the INSPIRE SDIGER Project ("SDIGER: A cross-border inter-administration Spatial Data Infrastructure to support WFD information access for Adour-Garonne and Ebro River Basins") was used in our hybrid methodology, as this related specifically to using geospatial information in a cross-border context, and included cost elements relating specifically to interoperability of both data and applications, based on real world experiences in the SDIGER project. Cost categories considered included: initial system study, pre-implementation work, hardware and software acquisition and maintenance, development and development support, developing web applications, handling legacy data, setting up the new system(s), on-going data related costs and on-going general system costs (training, support, maintenance, etc.), These were further sub-divided into over 40 sub-catgories, for which original estimates and a assessment of risk factor were recorded, leading to a risk-adjusted cost figure. This hybrid methodology is too complex to be fully described in this paper. The reader is directed to the MOTIIVE cost-benefit report, which contains some worked examples of applying the methodology and template spreadsheets used in the analysis (Longhorn, 2007).

Table 1: Benefit:Cost Ratios from prior studies

Date	Organisation	Country	Type of Study	Benefit:Cost
1990	New South Wales state	Australia	Economic aspects of digital mapping	2:1 to 9:1
1990	Western Australia Dept. of Land Administration	Australia	Land Information Programme	5.9:1
1991	Office of Information Technology of South Australia	Australia	GI in the Public Sector	2.9:1 to 5.8:1
1992	AUSLIG	Australia	Economic & Social Benefits of Public Interest Programme	3.8:1
1992	Dept. of Defence	Australia	Economic Benefits of Hydrographic Programmes	2.7:1
1993	Gov. of Victoria	Australia	Strategic Framework for GIS Development	5.5:1
1994	Australian Bureau of Agricultural & Resource Economics	Australia	Economic Analysis of Remote Sensing for Land Management	[1]
1995	ANZLIC	Australia/NZ	Australian Land and Geographic Data Benefits Study	4:1
1996	Coopers & Lybrand for OS GB	U.K.	Economics of collecting, disseminating and integrating government GI	[N/A]
1998	US Department of Agriculture	USA	ROI for GIS Projects from agency-wide Business Process Re-engineering study	$168M savings/yr
1999	Dept. of Land & Water Conservation, NSW	Australia	Business Case for Community Access to Natural Resources Information (1999-2003)	1.82:1 average
1999	OXERA for OS GB	U.K.	Economic contribution of OS GB	[2]
2000	PIRA International (USA)	EU-wide	Commercial Exploitation of Europe's Public Sector information	[3]

Date	Organisation	Country	Type of Study	Benefit:Cost
2000	Centre for International Economics, Sydney (for GSDI)	Global	Describes preferred methodology for preparing business case for SDI	N/A
2001	Baltimore County (Maryland, USA) Office of Information Technology	USA (local gov)	10-year forecast CBA for savings across local government departments using GIS and geodata	IRR - 64% to 168%
2002	Austrian Federal Ministry of Economics and Labour	Austria & Europe	Economic analysis of CBA for Austrian cadastral GI	23:1 [4]
2003	Environment Agency UK & Univ. of Sheffield, UK	EU-wide	Contribution to the Extended Impact Assessment for INSPIRE	4.4:1 to 8.9:1
2004	European Commission INSPIRE	EU-wide	Extended Impact Assessment for INSPIRE	5.4:1 to 12.4:1
2004	US Geological Survey US Dept. of Interior	USA	Determined net present value (NPV) of USA National Map programme over 30-years	$2 billion benefit
2005	Booz Allen Hamilton (USA)	USA	Geospatial Interoperability ROI Study	RoI 26.2%

[1] Remote sensing returned 'net gain" of AUS$1.5 million and AUS$66 million in monitoring trees and fertilizer use.

[2] OXERA reported estimated value to UK economy of £100 billion (£100 000 million) from GI maintained by Ordnance Survey GB at an annual cost of around £100 million.

[3] Economic potential to society of wider use of PSI, of which GI played a major part (over 50% of total PSI value).

[4] The Austrian analysis includes tax revenues in the benefits to the state, as well as registration fees; this is more a "monetary revenue:cost" ratio than the CBAs reported in other studies.

[N/A] = no specific figures are stated or the studies looked mainly at non-quantifiable, qualitative benefits, so figures are "Not Available", or looked at "benefits of GIS", thus "Not Applicable".

4 FEATURE TYPES FOR THE MARINE COMMUNITY

4.1 Context

This section provides a description as to how standards published by ISO TC211 can be used as a harmonisation mechanism in the coastal and marine community. It builds open earlier work conducted as part of the MarineXML initiative of the International Oceanographic Commission of UNESCO part funded by the European Commission (Millard et al, 2005).

From the MOTIIVE use case, the harmonisation requirements of the coastal and marine community stem from the need to establish data products that are based on integrating datasets from multiple sources. These datasets represent processes that occur in the natural world across a wide spectrum of spatial and temporal scales, and considerable science informs the design of experimental sampling strategies. The geometry and topology of observation sets are a fundamental determinant of the scientific uses to which they may be put. Moreover, the properties of the instruments used to generate data themselves place constraints on their interpretation (for example accuracy, precision, calibration, required post-processing.

These two factors – the scientific utility of a sampling regime and the limitations of an observing process – lead to a natural, scientifically important, classification of data types along these axes. This provides the basis for defining the Features and subsequently realising a set of Application Schemata for the data sets. The Application Schemata developed by this process have been given the name Climate Science Modelling Language as their applicability extends beyond the Marine and Coastal community to any community where datasets inherently represent observations and measurements of the natural world.

4.2 Climate Science Modelling Language

The Climate Science Modelling Language (CSML) provides a semantic model and encoding for representing a range of conceptual information classes of relevance to environmental sciences. These classes may be leveraged to build intelligent services for data subsetting, aggregation, processing. As well, CSML provide a 'wrapper' mechanism to encapsulate legacy file-based data, exposing them instead through the conceptual view. The motivation and context to CSML has been given in (Woolf et al, 2003 & 2006). Some key elements are reviewed below. A full reference manual can be found in *Climate Science Modelling Language – Reference Manual*, (Woolf 2007).

4.2.1 Standards-based data modelling

An emerging series of international standards for geospatial data, metadata and services (Woolf et al, 2005) provides a timely basis for meeting the CSML data integration requirements (Woolf, 2006). The series (comprising some 40 standards) is being developed by Technical Committee 211 (Geographic Information and Geomatics) of the International Organisation for Standardisation (ISO).

The overall scope of these standards is outlined in ISO 19101[6], which introduces the "Domain Reference Model" – an abstract information architecture for geospatial data infrastructures. At the core of the model is a geospatial 'Dataset'. A Dataset contains 'Feature' instances (see section 4.2.2 below) and related objects, and is described by 'Metadata'. 'Geographic information services'

[6] ISO 19101 Geographic information – Reference model

operate on a Dataset, while the logical structure and semantic content of a Dataset is described through an 'Application schema' (see section 4.2.3 below).

4.2.2 "Features"

A cornerstone of CSML is to use the "Feature" model as its key to integration (Woolf et al, 2006). Defined broadly as an "abstraction of a real world phenomena", a feature may represent any important aspect of a universe of discourse; and may be characterised in terms of its attributes, associations with other Features, constraints and operations that may be performed (the so-called "General Feature Model"[7],). In essence, Features provide an object view of data, and may occur as both types and instances. Feature type definitions may be stored for re-use in catalogues (ISO 19110[8]). Since Features encapsulate important data semantics within communities of practice, such Feature Type Catalogues may be regarded as 'semantics repositories' within an overall information architecture. Feature instances are the primary constituents of geospatial Datasets.

4.2.3 Application schema

While feature types define individual information classes, a "Dataset" is described with an "Application schema"[9], [10]. It defines the logical structure and semantic content of a dataset, and specifies the allowable feature types that may be contained. CSML is an application schema of the Geography Markup Language (GML,[11], see section 4.2.5 below).

4.2.4 Governance issues

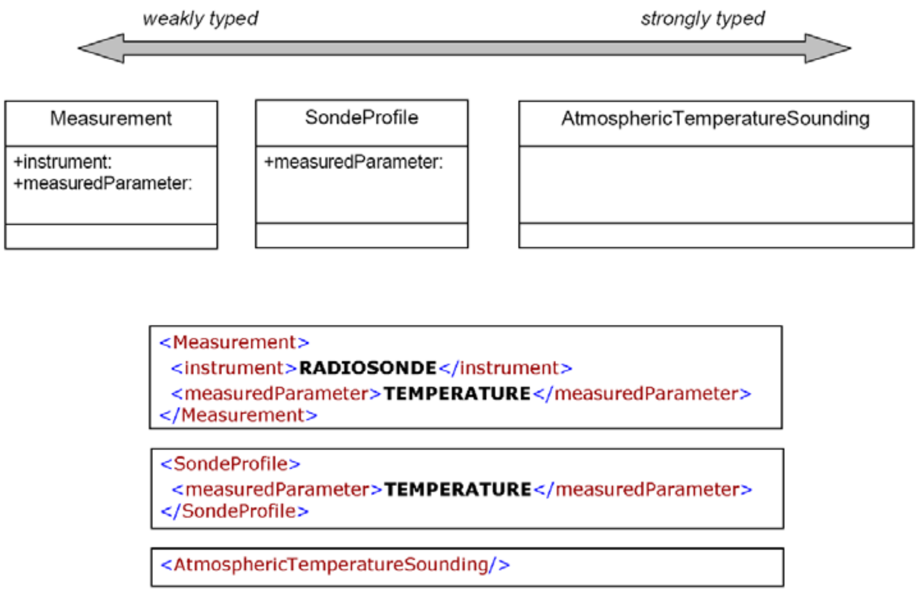
The integration framework outlined above aims to capture semantics of important community information types. To be successful, any attempt to apply this framework must engage with the community concerned. Various points of agreement must be established: what are (1) the information objects that should be modelled as Features; (2) the precise definition of feature types (i.e. their attributes, associations, constraints and behaviours); (3) common dictionaries (e.g. for physical units, coordinate references systems, physical "phenomena" definitions, etc.) and (4) maintenance procedures for agreed definitions.

Governance also is an important control for typing granularity of Features (Atkinson, 2004). *A-priori*, it is not always obvious to a designer of feature types how strongly typed they should be. In general, the more specialised the feature types, the greater will be the number required to capture the spectrum of information types used by the community. Figure 1 illustrates this. The marine

[7] ISO 19109 Geographic information – Rules for application schema
[8] ISO 19110 Geographic information – Methodology for feature cataloguing
[9] ISO 19101 Geographic information – Reference model
[10] ISO 19109 Geographic information – Rules for application schema
[11] ISO 19136, Geographic information – Geography Markup Language

community may agree to have one very weakly typed Feature covering all 'Measurements' or a large number of very strongly type Features for each measurement type (of which AtmosphericTemperatureSounding is one). Within the marine science community, bodies that might be expected to hold a mandate for maintaining detailed Feature Type Catalogues include the UN agencies World Meteorological Organisation (WMO) and International Oceanographic Commission (IOC).

Figure 1: Typing granularity (from Woolf, 2007)



4.2.5 Geography Markup Language, GML

GML provides reference XML schemas for a range of conceptual 'building-blocks' from other ISO standards, e.g. for geometry and topology [12], spatial [13] and temporal[14] reference, 'coverage' and gridded data[15], etc.

While GML is a large specification, its purpose is often misunderstood. GML is "an XML grammar written in XML Schema for the description of application schemas as well as the transport and storage of geographic information"[16]. The specification notes that "*… few applications will require the full range of capabilities described by the GML Schema*". The intention with GML is that its

[12] ISO 19107, Geographic information – Spatial schema
[13] ISO 19111, Geographic information – Spatial referencing by coordinates
[14] ISO 19108, Geographic information – Temporal schema
[15] ISO 19123, Geographic information – Schema for coverage geometry and functions
[16] ISO 19136, Geographic information – Geography Markup Language

schema constructs should be used as building blocks to provide canonical encodings for feature-types designed within a particular application domain. While GML elements *may* be used in arbitrary XML documents, the main purpose is to encode datasets according to an application schema. A valid GML instance document may contain feature instances or collections, dictionaries (of coordinate reference systems or units of measure etc.), or collections of topological objects.

4.2.6 The 'model-driven approach'

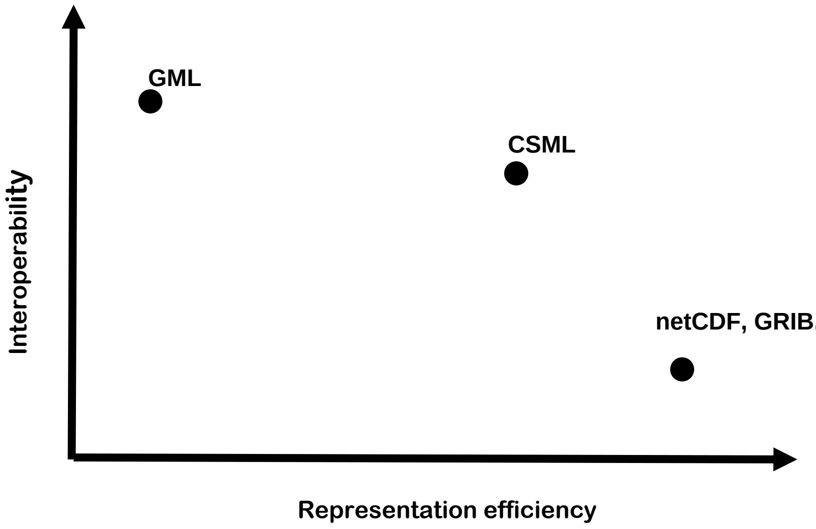
The ISO framework implements a 'model-driven approach' to data encoding. Feature types and datasets are first modelled at a conceptual level using a conceptual schema language (UML,[17]. The conceptual UML model then provides an interoperable view of data independent of actual representation. For canonical encoding and exchange, an automated mapping to XML is defined (see[18], and Annexes E and F of GML[19]). Indeed, apart from some historical quirks, GML itself is little more than the programmatic XML realisation of UML conceptual models defined in other ISO standards. CSML provides an additional mechanism for mapping file-based data onto a GML instance, as described next.

4.2.7 An interoperability model for legacy data

The ambition of CSML to provide standards-based interoperable information models for the climate sciences is in tension with the reality of data management praxis in this domain – very large volume (often terabyte-scale), file-based data in proprietary or ad-hoc but efficient formats, with a considerable legacy of existing – often elaborate – operational infrastructure built around them. It is infeasible on numerous grounds to imagine warehousing existing data into a new 'interoperable' XML representation. CSML attempts to strike a middle-ground and is proposed as a best practice approach for similar problem domains. Figure 2 attempts to show schematically the positioning of CSML with GML and commonly used binary encodings for exchange of oceanographic and meteorological data such as netCDF[20] and GRIB[21]

[17] ISO 19123, Geographic information – Schema for coverage geometry and functions
[18] ISO 19118, Geographic information – Encoding
[19] ISO 19110, Geographic information – Methodology for feature cataloguing
[20] http://www.unidata.ucar.edu/software/netcdf/
[21] http://www.grib.us/

Figure 2: Interoperability *vs* efficiency in the oceanographic sciences (from Stock et al, 2007)

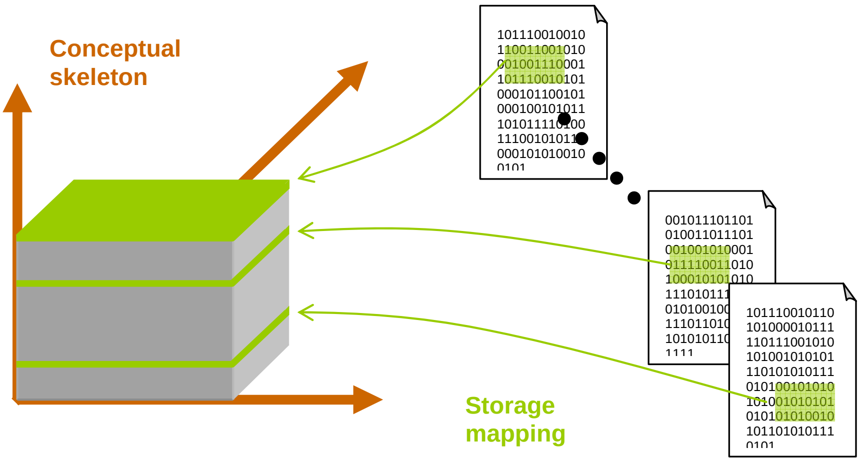


The goal of interoperability suggests the standards-based approach described above – conceptual modelling of important community feature types, with model-driven serialisation into standardised XML representations. On the other hand, compatibility with existing infrastructure and efficiency of storage demands the continuing use of community data formats.

In fact, both are needed. The CSML solution formalises the working habits of practitioners, who don't plan their activities in terms of low-level file formats, but rather at a higher conceptual level. For instance a researcher may wish to "calculate the annual average surface temperature over continental Europe". The conceptual model held by the researcher is of a field of temperature data extending over the Earth's surface, and vertically, and varying in time over some number of years. Regardless of the actual representation of the data (be it in one or more netCDF files, or GRIB 2-d layers, or in pages of an Excel spreadsheet), it is clear that this representation can always be *logically mapped onto the conceptual view* (otherwise the researcher's task would not be possible). The mapping might not be straightforward (e.g. it might involve extracting and aggregating arrays of data from many hundreds of files), but it is logically possible. Thus we arrive at the CSML model for interoperability. It consists of two components (Figure 3):

1. a conceptual model of information capturing essential feature types independent of storage details – an interoperable 'skeleton'; with
2. a mechanism for mapping stored (file-based) data into feature instances to provide the content.

Figure 3: The CSML interoperability model for legacy data (from Stock et al, 2007)



This two-component model for interoperability is represented in CSML through two schemas – a GML application schema for feature instances, and a 'storage descriptor' schema for describing the logical extraction and aggregation of data from storage, with the mapping between them being achieved through the use of xlink.

4.3 Feature-type principles for the climate sciences

We examine here some of the guiding principles behind the choice of feature-types in CSML.

4.3.1 Scientific data types

Two factors – the scientific utility of a sampling regime, and the limitations of an observing process – lead to a natural, scientifically important, classification of data types along these axes. Quite often the two are highly correlated (certain instruments generate certain samplings), and so scientific communities of practice adopt more abstract conceptual information classes that nevertheless reflect artefacts of sampling or instrument-type. This is particularly evident in the environmental sciences, where broad information classes based on measurement-set geometry and topology have almost universal acceptance. The following examples are illustrative.

- The US National Oceanographic and Atmospheric Administration (NOAA) is developing a plan for a Global Earth Observing Integrated Data Environment (GEO-IDE) to integrate measurements, data and products and create interoperability across data management systems. In the GEO-IDE Concept of Operations[22], the following 'structural data types' are defined: *Grids*, *Moving-sensor multidimensional fields*, *Time series*, *Profiles*, *Trajectories*, *Geospatial Framework Data*, *Point Data*, *Metadata*.

- The ESRI 'ArcMarine' Data Model for marine data includes classes like *Instant*, *Location Series*, *Time Series*, *Profile Line*, *Track*, *Sounding*, *Survey*, *Regularly Interpolated Surfaces*, *Mesh Volume*, etc.
- File formats such as netCDF and NASA Ames utilise data models that reflect these structures (e.g. netCDF four-dimensional gridded lat-lon-height-time variables, or NASA Ames time-series at a point).

Based on this observation, CSML feature types are classified primarily around geometric and topologic structure, and not the semantics of the observable or measured. The CSML feature types do not carry any explicit topologic descriptions at present. However GML provides rich schemas for describing topology – a requirement may arise in the future to add this information to CSML.

4.3.2 'Soft-typing' on phenomenon

If two Features are structurally identical, except for the physical 'phenomenon' of interest (temperature, salinity, wind vector, humidity, etc) then they are modelled in CSML as the same feature type. For example, both a vertical wind profile and an atmospheric temperature sounding have similar characteristics (in terms of attributes and data handling operations that may be performed), differing primarily in the distinguishing physical parameter (vector wind *vs* temperature). Their representations are collapsed, in CSML, into the same feature type. This principle ('soft-typing' on phenomenon) has been identified as appropriate for GML modelling of coverage-type data in operational oceanography and meteorology (UK MetOffice, 2006).

4.3.3 Conventional portrayal

Most climate science data has a conventional portrayal used by practitioners (e.g. model output is typically displayed as shaded 2-d slices, an atmospheric sounding as a line graph against height in the vertical, a marine temperature section as a 2-d contoured field against depth and ship track distance, etc.). A workable minimum granularity for CSML feature types is determined by applying a discriminant of "sensible plotting": there should be sufficient detail within a feature type – and sufficient difference between feature types – to enable in-principle unsupervised rendering, in the conventional manner. This criterion is somewhat loose, and it remains to be seen in practice the extent to which it is satisfied with the feature types chosen. In that sense, the principle may play a more important role in evaluation than in design.

[22] https://www.nosc.noaa.gov/dmc/docs/NOAA_GEO-IDE_CONOPS-v3-3.pdf

4.3.4 The 'Observations and Measurements' pattern

The Open Geospatial Consortium's Observations and Measurements best practice paper (Cox, 2003) provides a conceptual model and encoding for observations and measurements. Under this model, an *observation* is an event whose *result* is an estimate of the value of some *property* of a *feature-of-interest* obtained using a specified *procedure* (Figure 4). Each of the major classes in the model may be specialised for a particular application domain. The observed property of an observation may be modelled using the O&M Phenomenon taxonomy (Figure 5).

Figure 4: The Observations and Measurements model (from Cox, 2003)

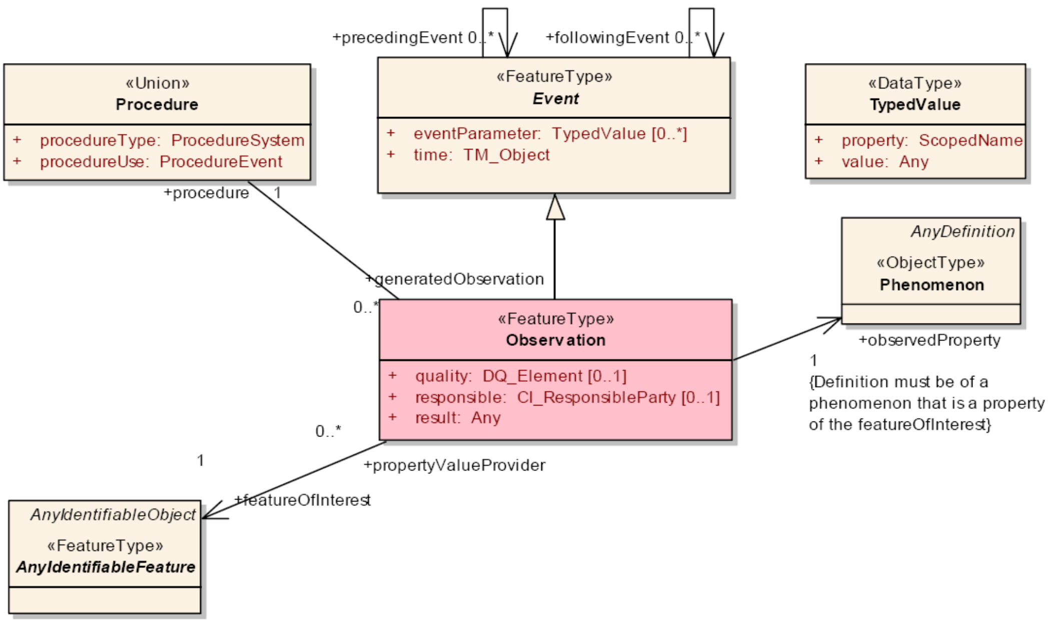
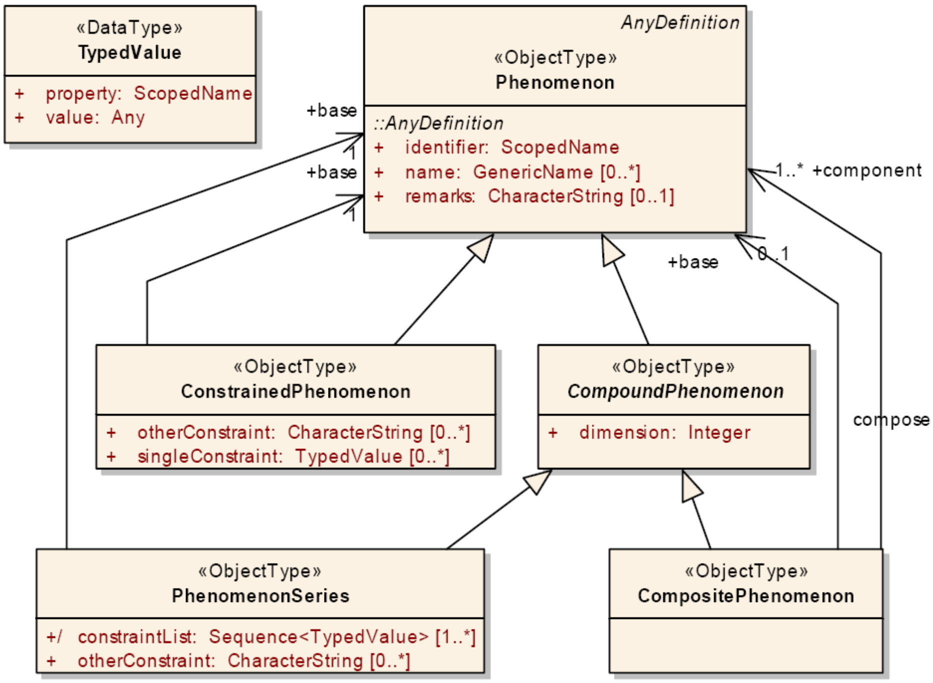


Figure 5: The O&M Phenomenon model (from Cox, 2003)

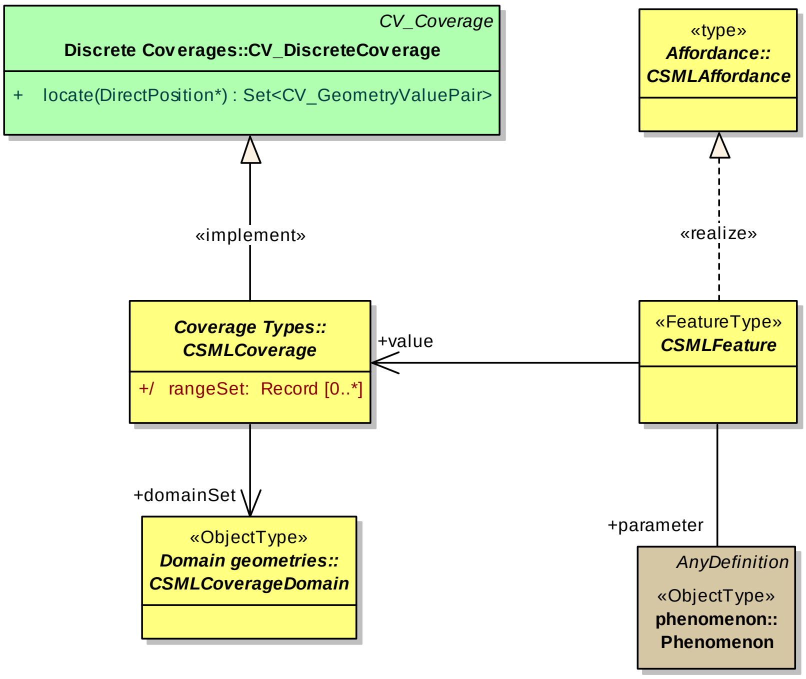


4.4 CSML feature types

The general model for CSML feature types was is given below in Figure 6. Important elements to note are:

- A CSML Feature Type *realises a «type»* (representing a 'processing affordance'). Thus, the operations of the «type» are inherited, and in the CSML application of this pattern attributes of the «type» are copied down (not strictly required by UML)
- A CSML Feature Type represents some observed or simulated *physical parameter* (modelled through the OGC O&M Phenomenon class taxonomy).
- The *value* of this parameter is provided by a CSML 'coverage'. A coverage is, conceptually, a mapping from some spatiotemporal domain to a value range.
- CSML feature types are distinguished primarily on the basis of the *geometry of the coverage domain*. There is a strong relationship between these 'sampling geometries' and the Feature Type operations that are afforded.

Figure 6: General Model for CSML (from Woolf, 2007)



The set of 13 CSML Feature Types are listed in Table 2 below.

Table 2: CSML Feature Types (from Woolf, 2007)

Feature type	Description	Example
PointFeature	Single point measurement.	Raingauge measurement
PointSeriesFeature	Time-series of single datum measurements at a fixed location in space.	tidegauge, rainfall timeseries
TrajectoryFeature	Measurement along a discrete path in time and space.	surface salinity along a ship's cruise track; atmospheric aerosols along an aircraft's flight path
PointCollectionFeature	Collection of distributed single datum measurements at a particular time	2m temperatures measured at weather stations across the UK at 0600z.
ProfileFeature	Single 'profile' of some parameter along a vertical	wind sounding, XBT, CTD, radiosonde

	line in space.	
ProfileSeriesFeature	Time-series of profiles on fixed vertical levels at a fixed location	vertical radar timeseries, thermistor chain timeseries
RaggedProfileSeriesFeature	Time-series of unequal-length profiles, but on fixed vertical levels, at a fixed location	repeat daily balloon soundings of atmospheric temperature from the same location
SectionFeature	Series of profiles from positions along a trajectory in time and space.	shipborne ADCP
RaggedSectionFeature	Series of profiles of unequal length along a trajectory in time and space	marine CTD measurements along a ship's cruise track
ScanningRadarFeature	Backscatter profiles along a look direction at fixed elevation but rotating in azimuth	weather radar
GridFeature	Single time-snapshot of a gridded field.	gridded analysis field
GridSeriesFeature	Time-series of gridded parameter fields	numerical weather prediction model, ocean general circulation model
SwathFeature	Two-dimensional grid of data along a satellite ground-path	AVHRR satellite imagery

5 FEATURE TYPE CATALOGUES

5.1 MOTIIVE Approach to Feature Type Catalogues

This section discusses the work undertaken by MOTIIVE to deploy a web-enabled implementation of a Feature Type Catalogue (FTC) (Stock et al, 2007). The specification for this FTC has subsequently been submitted to OGC as a 'best practice' paper.

In order to develop and demonstrate the concept of cross-discipline, cross-border integration of information about land and sea, MOTIIVE included the creation of a small scale spatial data infrastructure. This infrastructure allowed participants to serve their data using OGC compliant web services to a broader community. The web services explored for the purposes of the project include those conforming to the OGC Web Feature Service and Web Coverage Service specifications.

The successful discovery and execution of these services within a spatial data infrastructure (SDI) requires a registry containing information about the resources (both services and data) that are available within the SDI. The heterogeneous nature of these resources suggested that a registry containing a rich feature type catalogue is most appropriate to support discovery and integration of data sets from a wide range of different sources. In this way, users will be able to identify semantic linkages between resources and thus identify all information that is of interest.

5.2 Registry Standards used for MOTIIVE

The MOTIIVE SDI depends on a registry that contains all of the necessary information to allow resources to be discovered and executed by clients. Relevant standards for registry design include:

- ebRIM (OASIS/ebXML Registry Technical Committee, 2002), which describes a registry information model;
- the OGC Catalogue Services specification (Open Geospatial Consortium, 2005), which describes a grammar for a minimal query language and a set of elements that should be returned by compliant services;
- CSW, the Catalogue Services for the Web specification, which is a binding of the OGC Catalogue Services specification for HTTP and
- WRS, the Web Registry Service[23], which is the ebRIM application profile of CSW (Open Geospatial Consortium, 2005a) and provides:
  - a set of schemas for the basic package and for discovery, capabilities and publications services;
  - some exception codes to be used during request processing;
  - a mapping from ebRIM elements to the specified set of brief and summary elements from CSW;
  - a mapping from ebRIM elements to the CSW Record elements that are returned by some operations;
  - additional components in the GetCapabilities responses and requests:
    - ♦ permissible section names within the request;
    - ♦ new elements added to the response;
    - ♦ specific service Features and properties that may be provided by a service and specified in the response;
    - ♦ an additional element that provides a link to the WSDL description for a service;
  - additional components and restrictions on the GetRecords requests and responses:
    - ♦ restrictions on the use of the request;

[23] 'WRS' is used to refer to the OGC Catalogue Services ebRIM application profile for CSW

  - ♦ a method for allowing a catalogue record to be returned as ebRIM or CSW representations;
  - ♦ additional constraints on the query syntax to allow the complex ebRIM information model to be queried and to avoid ambiguity when navigating multiple associations between registry objects;
  - ♦ support for queries specified using XPath and XQuery syntax;
  - ♦ a method for invoking stored queries;
  - ♦ a method for defining stored queries;
  - ♦ the specification of temporal operators from GML and
- additional minor components and restrictions on the requests and responses made by the other CSW operations.

5.3 MOTIIVE Feature Type Catalogue Requirements

The WRS specification includes modifications to CSW to handle ebRIM in a generic way, but does not explicitly incorporate a rich feature type catalogue. MOTIIVE extended the WRS specification by the creation of a new extension package that incorporates an ISO 19110 compliant Feature Type Catalogue. Specifically, this involves the addition of elements to the ebRIM model to handle the following:

1. Feature types.
2. Feature attributes
3. Feature operations.

   The inclusion of operations in the FTC is particularly useful in the context of a registry because:

   a) The operations of a feature type can be linked to the web services that implement those operations. In most practical cases, an operation will only be specified if it is implemented as a service. However, the capability will also be included to specify an operation for a Feature Type that is not implemented as a service in order to allow for future development and to provide a more complete semantic representation of a feature type. Such semantic representations may assist with possible future development of semantic querying and discovery tools, semantic similarity assessment and reasoning for automated service chaining.

   b) The operations can also be linked to the properties that are related to their use. Such properties may be intimately linked with the operations over and above the link between an operation and the feature type itself. This provides additional semantic information that may be useful in interpretation and querying of the operations and discovery of web services that implement them.

4. Feature relationships.
   A range of different association types between feature types are included in the ebRIM FTC. Traditional feature type catalogue focus on the hierarchical (is-a) relationship between Features, but a richer feature type catalogue requires that links of other types may be expressed, including synonyms, part-of and derived from relationships. These are implemented in a flexible way so that further relationship types can be added and extended in the future by other projects.
5. Inheritance of properties, operations and relationships with other feature types.
   This allows a feature type to automatically include the properties, operations and relationships of the parent feature type in addition to those explicitly expressed for the child.

The extension package created for the WRS to implement these additional characteristics of a rich feature type catalogue includes the following:

1. Extensions to the ebRIM information model as used in the WRS specification to include modelling of a rich feature type catalogue, including all of those requirements described in the previous paragraph. This work is compatible with the FTC proposed in ISO 19110 to provide a bridge between the work done under that standard and that done under ebRIM.
2. Specification of predefined queries that use the query grammar specified in CSW to allow the richer aspects of the feature type catalogue to be interrogated.

5.4 Extensions to the WRS Information Model Required for MOTIIVE

ISO 19109 specifies rules for application schemas, and includes a general feature model that describes the main components that would be used in a feature type catalogue at a meta-level. The specification includes elements like feature types, attributes, operations and association types. ISO 19110 is a model for a feature type catalogue that realises the general feature model in ISO 19109. Figure 7 is a UML diagram for ISO 19110 as a realisation of ISO 19110 (including the most recently proposed revisions N 2053, not yet in the formal specification). The ISO 19110 specification contains details of the derivation of ISO 19110 from ISO 19109 general feature model. MOTIIVE involved mapping the ISO 19110 feature type catalogue into the ebRIM information model. Figure 8 provides this mapping, indicating how the ebRIM objects will be used to represent the ISO 19110 feature type catalogue model. The ISO 19110 objects classes are shown as specialisations of various ebRIM objects. It should be noted that this diagram represents the current consideration for this mapping and testing and further consideration is likely to result in changes to this model.

Figure 7: UML Model of ISO 19110

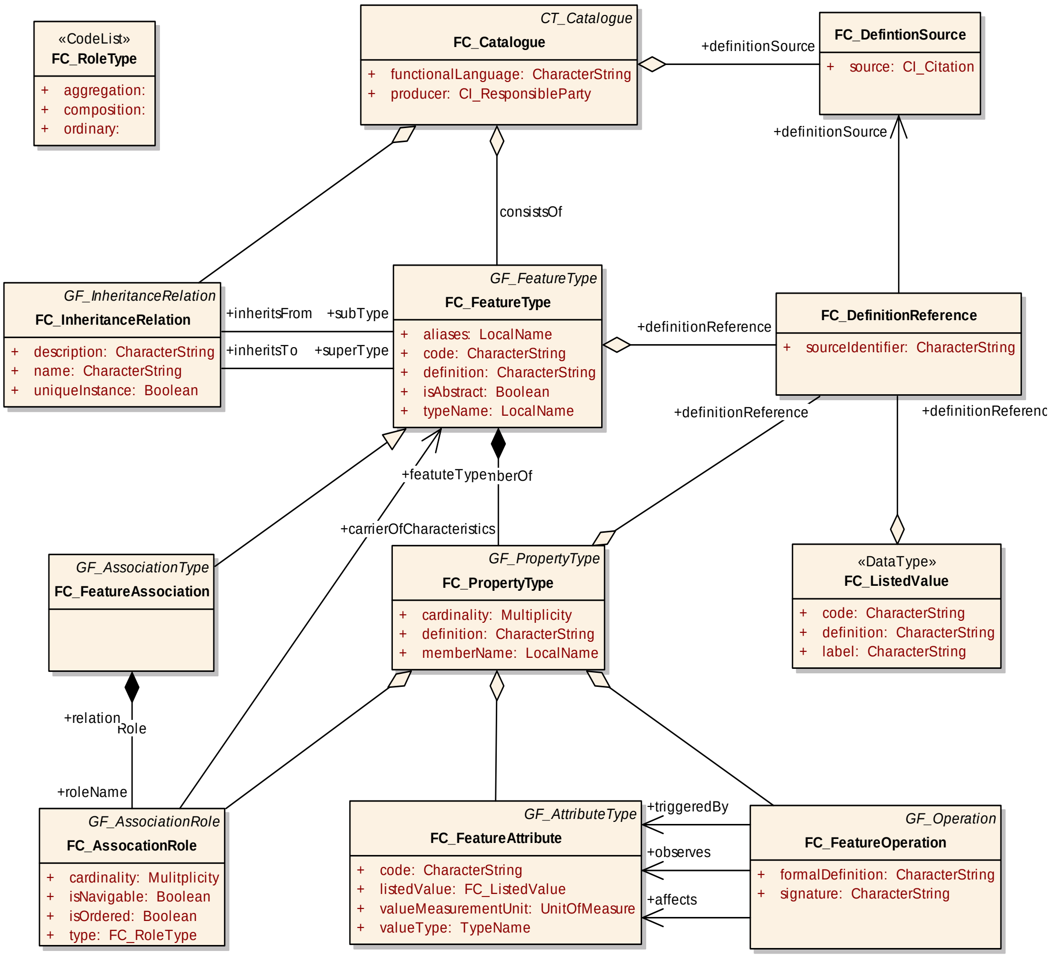
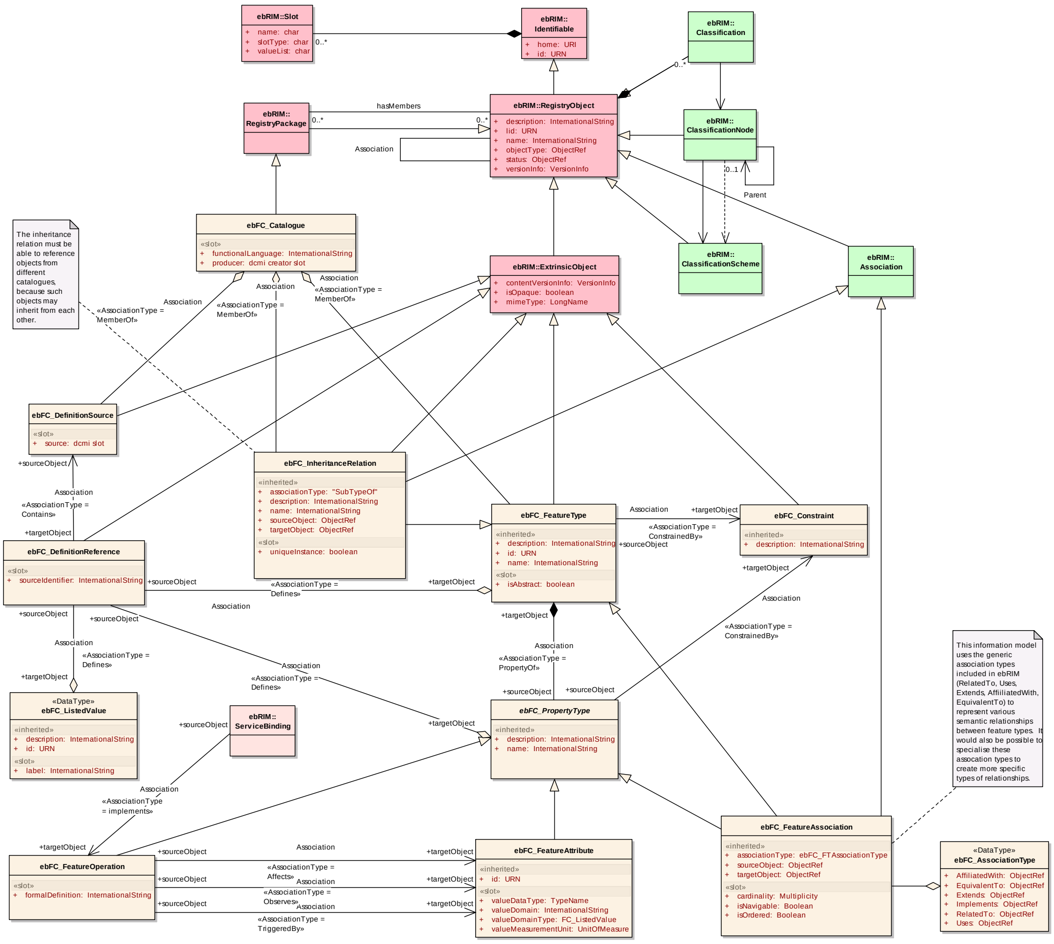


Figure 8: Mapping of ISO 19110 to ebRIM



5.5 Demonstration of the Feature Type Catalogue

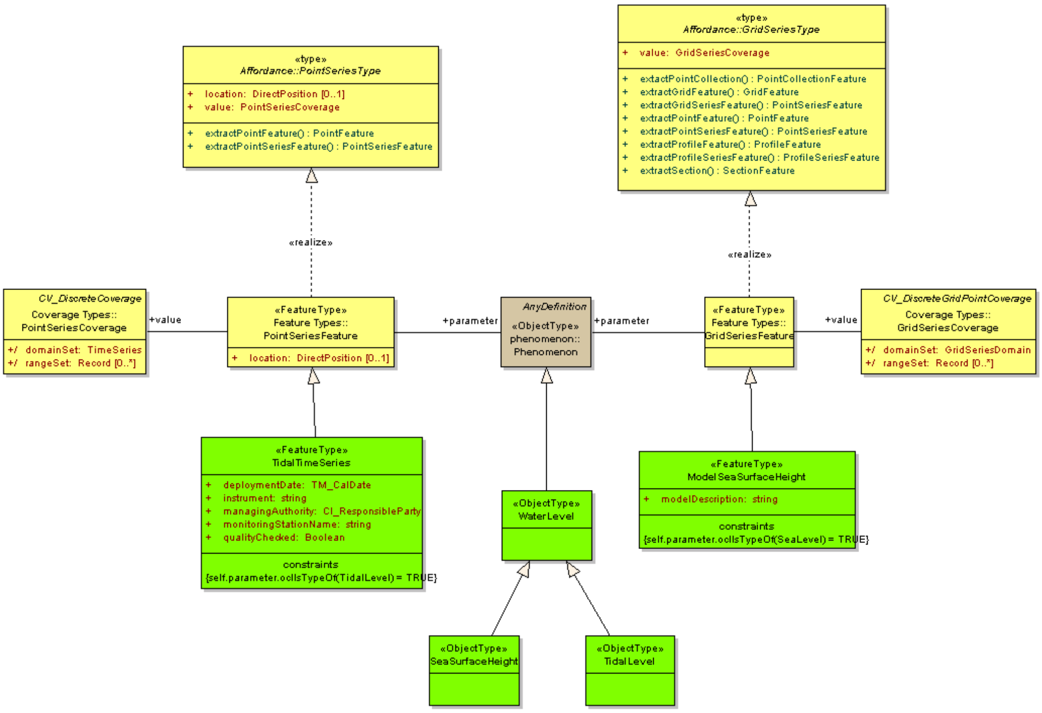
The Feature Type Catalogue (FTC) was demonstrated through test loading of Feature Types established as part of MOTIIVE. A browse client was also established for the FTC to provide a human readable view onto the catalogue and enable navigation through its contents. The MOTIIVE Feature Types were established as direct output from the use case in Box 1 and are described below:

- TidalTimeSeries: An example strongly typed FT governed by the community of users who require precise definition of measurements of water level through time a given point. This was implemented as specialisation of a CSML PointSeries

- ModelSeaSurfaceHeight. An example strongly typed FT governed by the community of users who require precise definitions of water level provided by numerical models over a given areas. This was implemented as a specialisation of a CSML GridSeries

These two Feature Types are shown using UML notation in Figure 7. The colour green is used to indentify the specialisation of the CSML packages (shown in yellow). This also highlights the governance boundaries related to data packing in that the two communities requiring TidalTimeSereis and ModelSeaSurfaceHeight do not need to define data models related to the geometries of the observations as this is provided by CSML. Data modelling in this way is not only more efficient, but also has the advantage that these two specialisations can take advantage of services afforded by the more weakly typed CSML. Such services include generic portrayal and data sub-setting.

Figure 7 Test Feature Types Loaded into the MOTIIVE FTC



6 CONCLUSIONS

This paper has looked at the issues with interoperability across and between the data themes described in the three Annexes in the INSPIRE Directive. It has considered how and why this interoperability is required by users and whether this interoperability can be quantified and justified in cost-benefit terms. The paper then presents how open standards can be used to achieve this interoperability in support of the INSPIRE Directive

In undertaking this work MOTIIVE has focussed on the example of the marine community. The marine community is very diverse covering areas such as meteorology, navigation, water quality, biodiversity and coastal defence amongst others. Although the purpose of information varies between its members (producing a weather forecast, identifying pollution sources, reporting on species abundance or designing a flood defence), there are common users needs for efficient discovery of data and then to be able to process this data in a seamless manner. In many cases this processing is simply enabling 'like for like' comparisons between data, for example comparing water height from a surge forecast against height on a topographic base map.

Our work suggests that these above aspirations are achievable through the adoption of common data types, or Features, to represent the data being exchanged and combined. These 'Features' relate to representations of the natural world obtained from measurement or observation programmes and in the first instance may seem unfamiliar to those who think of Features in terms of constructed objects 'roads' or 'buildings'. Embracing such Features is however important to the realisation of INSPIRE as most data on the environment is inherently based on these feature types. It would be mistake to regard these feature types as unique to the marine community – they are required wherever monitoring or sampling of the environment is needed. This paper has presented the findings of MOTIIVE as to how these Features can be derived as GML application schema. These GML application schema have collectively been referred to as the Climate Science Modelling Language (CSML). CSML encodes a collection of weakly typed Features for the results of common observations and measurements patterns. These can be specialised into more elaborate types as required by a particular community.

Any community deriving a feature type however needs a way of publishing their definitions for re-use. This mechanism ideally needs to be machine readable to such that it can be queried and used within a service-orientated architecture. For this reason MOTIIVE implemented and tested a web-based implementation of a feature type catalogue compliant with ISO 9110. This specification has subsequently been submitted to OGC as a best practice paper.

Finally, the work of MOTIIVE has identified that having clear and unambiguous definitions of Features available for others to subscribe to is a significant factor in reducing the cost of interoperability between information systems and the costs of designing and realising data products or services that span the coastal zone.

REFERENCES

Atkinson, R. (2004), “Next steps to interoperability – Mechanisms for semantic interoperability”, EOGEO Workshop, University College London, 23–25 June 2004, at http://y.eogeo.org/files/eogeo2004/eogeo-2004-full-atkinson.ppt , 31 January, 2007]

Booz Allen Hamilton (BAH). 2005. *Geospatial Interoperability Return on Investment Study*, April 2005. National Aeronautics and Space Administration, Geospatial Interoperability Office.

Cox, S., ed., 2003, *Observations and Measurements*, OpenGIS Consortium document 05-08r4, at http://portal.opengeospatial.org/files/?artifact_id=17038 , 31 January, 2007].

Ferreira, M., Lucius, I., Longhorn, R., Millard, K., 2007, *Data Harmonisation: The Needs and Aspirations of the EU Coastal Management Community*. MOTIIVE Deliverable D7b (21/02/2007): MOTIIVE_D7b_UserPositionPaper_R1r0.pdf at www.motiive.net

Longhorn, R., *Cost Benefit Worksheet*, MOTIIVE Deliverable D11b (24/05/2007) MOTIIVE_D11b_CBAWorksheet_R2r0.xls at www.mottive.net

Longhorn, R., Osborne, M., Pepper, J., 2007, *Data Harmonisation: Analysis of UK land-sea data integration projects MOTIIVE Deliverable D7a* (28/05/2007), MOTIIVE_D7a_IntegratedMapping_R2r0.pdf at www.motive.net

Longhorn, R., *Report on Cost-Benefit Methodologies*, MOTIIVE Deliverable D11a (24/05/2007) MOTIIVE_D11a_CBAReport_R2r0.pdf at www.motiive.net

Millard, K., et. al. 2005, “Using XML Technology for Marine Data Exchange – Position Paper of the MaxineXML initiative”, *Proc. 4th EuroGOOS, Conference*, Brest, France, at http://books.hrwallingford.co.uk

OASIS/ebXML Registry Technical Committee, 2002, OASIS/ebXML *Registry Information Model v2.0*. Approved OASIS Standard at http://www.oasis-open.org/committees/regrep/documents/2.0/specs/ebrim.pdf

Open Geospatial Consortium, 2005, *OGC Catalogue Services Specification*. OGC 04-021r3 at http://www.opengis.org

Open Geospatial Consortium, 2005a, OGC Catalogue Services – ebRIM (ISO/TS 15000-3) profile of CSW. OGC 05-025r3 at http://www.opengis.org

Stock, K., Small, M., Higgins, C., Atkinson, R., Woolf, A., 2007, OGC *Compliant Registry to Accommodate a Rich Feature Type Catalogue*, MOTIIVE Deliverable D10

UK MetOffice, 2006, *OGC Web Services and GML Modelling for Operational Meteorology, Communiqué – workshop findings and outcomes,* OGC_Web_Services_and_GML_Modelling_for_Operational_Meteorology --communique_V1.pdf, 31 January 2007, at http://www.metoffice.gov.uk/informatics

Woolf, A. et. al. (2003), *Data Virtualisation in the NERC DataGrid*, UK e-Science All-Hands Meeting, Nottingham, UK at http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/094.pdf , 26 July, 2004]

Woolf, A., 2007, *Climate Science Modelling Language – Reference Manual*. MOTIIVE Deliverable D9b. 22nd May 2007 MOTIIVE_D9b_ClimateScienceModellingLanguage_R1r0.pdf at www.motiive.net

Woolf, A., Lawrence, B., Lowry, R., Kleese van Dam, K., Cramer, R., Gutierrez, M., Kondapalli, S., Latham, S., Lowe, D., O'Neill, K., Stephens, A., 2006, "Data integration with the Climate Science Modelling Language", *Advances in Geosciences, 8, 83 – 90*. at http://direct.sref.org/1680-7359/adgeo/2006-8-83 , 31 January, 2007

Woolf, A., Lawrence, B., Lowry, R., Kleese van Dam, K., Cramer, R., Gutierrez, M., Kondapalli, S., Latham, S., and O'Neill, K., 2005, "Standards-based data interoperability for the climate sciences", *Met. Apps., 12(1), 9–22*, doi:10.1017/S1350482705001556.

Woolf, A., Lawrence, B., Tandy, J., Millard, K., Lowe, D., Pepler, S., 2006, "'Feature types' as an integration bridge in the climate sciences", AGU Fall Meeting, San Diego, Dec 2006, *Eos Trans. AGU, 87(52)*, Fall Meet. Suppl., Abstract IN53C-02.

Woolf, A., Millard, K., 2007, Harmonisation Framework Testing, MOTIIVE Deliverable D9a (31/05/2007). MOTIIVE_D9a_HarmonisationFrameworkTesting_R1r0.pdf at www.motiive.net