Common Metadata for Climate Modelling Digital Repositories

METAFOR Controlled Vocabularies Month 42

METAFOR Deliverable 2.7 M42

PROJECT	
Project acronym	METAFOR
Project full title	Common Metadata for Climate Modelling Digital Repositories
Grant agreement no:	211753
Funding Scheme	Combination of Collaborative Projects & Coordination and Support Actions
Call Topic	INFRA-2007-1.2.1 Scientific Digital Repositories
DOCUMENT	
Deliverable	D2.7 Month 42
Deliverable Title	METAFOR Controlled Vocabularies
Document Identifier	METAFOR-D2.7_M42
Date	September 13, 2011
Work Package	WP2 Defining the CIM
Authors	UREAD
Document Status	Final
Document Link	http://metaforclimate.eu/documents

Dissemination Level		
PU	Public	
PP	Restricted to other programmes participants	**X**
RE	Restricted to a group specified by the Consortium	
CO	Confidential	

Document History

CAPACITIES

e-infrastructure

Version	Date	Comment	Author/Partner
0.1	September 20, 2011	First Draft	C. Pascoe/BADC and M.-P. Moine /CERFACS/ K. Stott
0.2	October 05, 2011	Final Version	C. Pascoe

Abstract:

Work on the CMIP5 metadata questionnaire has produced a set of climate modelling controlled vocabularies, which are independent of the CIM and are also of significant value to the community. Here we document the controlled vocabularies and the methodology behind their collection and set out a method of governance for them outside the confines of the CIM.

Table of Contents

Purpose

A controlled vocabulary (CV) has been developed by Metafor to describe possible and actual earth system model configurations. The CV is comprised of terms and scientific options that are represented using mind maps, and an XML serialisation consisting of components, parameter groups, parameters, constraints and values. This document describes how the CV was built and used to produce a questionnaire for CMIP5 metadata it also describes how the CV will be managed and governed beyond the end of the Metafor project.

Picture of a climate model and climate experiments

Climate study is a highly interdisciplinary science that historically emerged by the convergence of scientific expertise in the research areas related to the earth system (oceanography, atmospheric physics, hydrology, etc.). As a result, a climate model is a composition of models each one being devoted to a specific domain of the climate system. These models are generally assembled together by a coupling system, e.g. OASIS [1] and BFG [2] which exchange coupling fields at the interface of the domains. The coupling system must also perform spatial remapping and time transformation because each model domain is discretised onto different grids and integrated with distinct time steps. The resulting global model, including realm components and the coupler, is referred to as a “coupled model”. State of the art coupled models contains millions lines of code over thousands of routines.

A given model can be run and integrated in time in a large number of different ways, depending on the temporal and dynamical schemes used, on the physical parameterization schemes turned on and according to the values of the tuneable parameters within each physical scheme of each component of the coupled model. The way the model is scientifically configured is crucial information to interpret and compare results, and it is important to preserve this information along with the data.

Such a model configuration is usually targeted at a specific numerical experiment: to, for example, understand the sensitivity of a climate process to horizontal resolution, or to provide a projection of future climate under a specific emission scenario. Hence, it is important not only to document a particular configuration, but also to document why that configuration was used. This is particularly important where model variants have been deliberately modified by removing processes to understand their impact – obviously integration with such a model should not be used to evaluate the impacts of future climate change, even if the model was projecting future climate! Thus, an important part of simulation documentation is the experiment description, and how the model was configured to conform to the experiment requirements.

The Metafor Controlled Vocabulary

The Metafor CIM is divided into specialised UML packages which cover the constitutive elements of the modelling activity. The CIM packages describe the experiment context, the individual simulations performed, the software used, the data manipulated by this software and the numerical grids. Each individual CIM package needs an associated set of Controlled Vocabulary (CV) that defines allowed attributes (name/value pairs). For some of the packages, the CV was simply based on a list of already existing terms such as the CF conventions for Data, GridSpec for Grids, and ISO standards for Quality.

However, the activity and software CIM packages required new vocabularies, firstly because up to now this area was poorly addressed, and secondly, because models and experiments have their own specificities not covered by existing standards. This report is primarily concerned with the “Model Controlled Vocabulary”, which although guided by the fifth Climate Model Inter-comparison Project (CMIP5) can be applied more generally to any climate model.

Methodology

Building a Controlled Vocabulary for Climate Models

The Model Controlled Vocabulary describes the heart of the climate data production chain which is the numerical model itself. This work started from a nearly blank page and had to go through the early steps of a classical CV building process:

(i) Identify the relevant and discriminating information about the climate models;

(ii) Set an ensemble of appropriate terms (meaningful and non ambiguous naming) efficient to synthetically and faithfully express the information;

(iii) Organize these terms hierarchically, with possible inter-dependencies;

(iv) Attach a definition to each;

(v) Identify allowed/possible values for each.

The resulting CV has a hierarchical structure and is made of possibly embedded elements: single “leaf parameters” (name/value pairs) are gathered within “parameter groups” containers themselves gathered within “components”. Some groups of parameters are “conditional parameter groups” depending on the value(s) taken by another parameter. This way of organizing the CV is both driven by typical structure of the models themselves and by the scientific rationale for gathering ideas into main themes, the two being obviously closely related. In the CMIP5 context, following the data protocol (http://cmip-pcmdi.llnl.gov), the common first level decomposition of a coupled climate model maps onto eight identified realm components: ocean, atmosphere, land surface, land ice, sea ice, atmospheric chemistry, aerosol and ocean biogeochemistry. Each realm component is in its turn made of sub-components, one per main physical or dynamical process. The CV exploits this organization and rearranges scientific concepts in an understandable and balanced way. The current CV granularity is a compromise driven by inter-comparison concern: reach a level of details sufficient to be meaningful and discriminating across the various climate models and avoid at the same time overloading and useless too specific information.

The CV could not be established by exploring the model literature alone. Scientific expertise was required to help identify the model characteristics important to capture and document in an inter-comparison context. The collaboration of a significant number of independent scientists was a key part of the success of the CV development, with many from the international climate community involved. During face meetings or through audio screen-sharing sessions, they were asked about the science and algorithms of the climate models they contributed to, whether by developing or running them. The discussions were captured using mindmaps, one for reach realm, which proved to be very appropriate for capturing structured information and feedback on the fly. The interviewing and reviewing procedure is illustrated in Figure 1. Thus far more than 35 experts from 13 research centres, representing 6 countries have contributed.

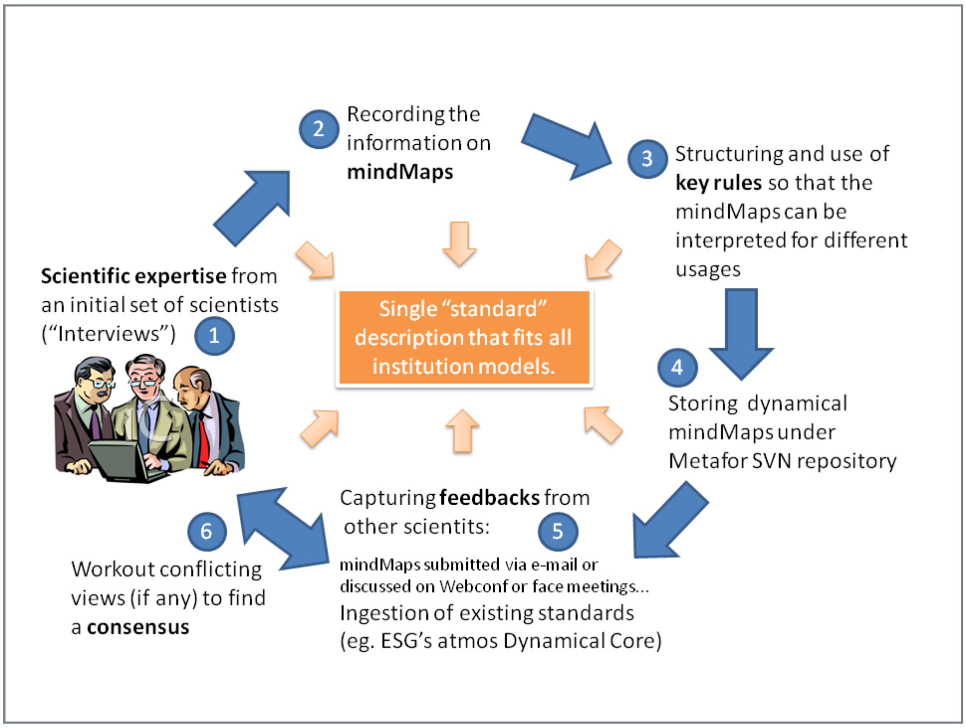

Figure 1: *Interview process with scientists: Following a first round interview with one realm expert (step 1), revision processes were launched with others scientists from other research centres (step 5). We integrated the feedback in a structured way (steps 2, 3), capturing their precise meaning, getting confirmation when necessary, working out possible conflicting views (step 6), and taking care not to introduce inconsistencies with previously collected CV.*

To ensure seamless usability of the model CV within the metadata workflow (in particular the ingestion into the CMIP5 questionnaire discussed later), we defined a set of formal typographic and structure-constraining rules applicable to the mindmap technology we used (freemind, see http://freemind.sourceforget.net). Font formats and icons enabled distinguishing the different CV containers and the different types of choice (exclusive or not) among a set of possible values for parameters along with the type of expected value (numeric or string). Units were prescribed where numeric values were expected, and parameter definitions provided. Rules were defined in the mindmaps to constrain allowable embedding of the CV containers and their relationships. Figure 2 demonstrates how these rules are used to organise CV mindmaps to generate web entry forms for the online metadata creation tool and Figure 3 shows what such a mind map looks like in practice.

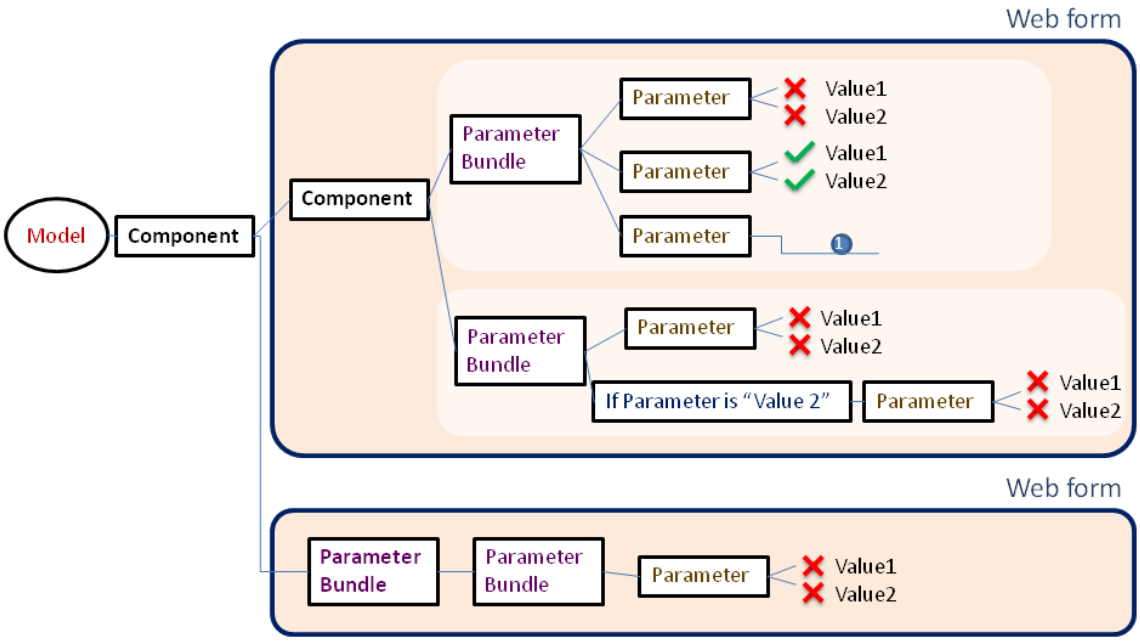

Figure 2: *typographic and structure-constraining rules for mind maps; the enclosed black bold font denotes model components, enclosed purple is for parameter groups; enclosed blue is for conditional parameter groups; enclosed brown is for leaf parameters expecting values; black is for possible values for the leaf parameters; red cross icons mark single choice (XOR), green tick mark icons symbolize possible multiple choice (OR); the circled number 1 and pencil icons (not shown) give the type of the value (numeric or free-text respectively); note-book icons (not shown) indicate where definitions are attached as a foot-notes.*

The final model CV has three main categories:

- The CV for the model realm components, including details of the numerical schemes deployed for dynamical processes (advection, diffusion, transport), for time integration schemes and key information about the parameterization schemes used to model sub-grid scale physical processes (e.g. precipitation and clouds in the atmosphere realm; soil hydrology in land surface realm, gas phase processes in atmospheric chemistry realm); this is the heart of the model CV;

- The CV for describing the way components are coupled together, including selected terms for spatial regridding and time transformation; these latter have been derived from vocabulary used for standard configuration of couplers (e.g. OASIS [1], BFG [2]);

- The CV associated with the numerical grids, either the ones used by the models for spatial discretization or the grids onto which the data are projected; this is the result of a conciliation of gridSpec standards (http://www.gfdl.noaa.gov/~vb/gridstd/gridstd.html) used extensively in the CIM Grid package and new Metafor specific vocabulary.

The complete set of CV addressed more than 570 leaf parameters over 8 realms and the complete CV can be found on the metafor subversion repository links to which are listed in appendix A. Figure 3 shows a portion of the sea ice CV showing the Sea Ice Thermondynamics sub-component.

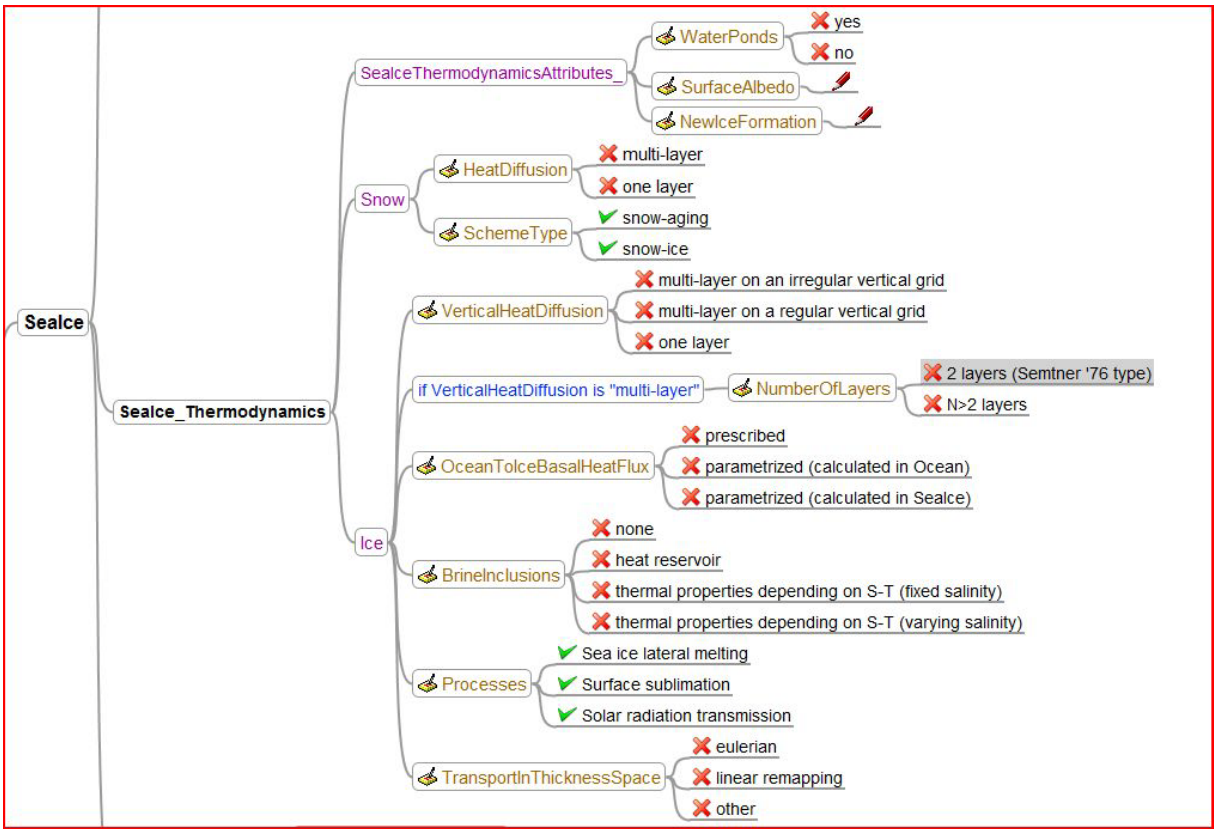

Figure 3: *A portion of the sea ice CV, showing the Sea Ice Thermodynamics sub-component*

From Controlled Vocabulary to Metadata

To collect metadata for CMIP5 numerical models, simulations and experiments, we have constructed what was initially intended to be a "simple questionnaire". However, it rapidly became clear that a traditional questionnaire based linear collection of information would be completely inappropriate for the task as many component descriptions would need to be re-used. A more complex tool was needed, and clearly that tool would have to facilitate the use of the controlled vocabularies. The name has remained, but the "CMIP5 Questionnaire" should be thought of as a complex metadata entry tool!

Figure 4 shows a portion of the resulting tool, built using the python Django web framework (http://www.djangoproject.com/), deployed at the British Atmospheric Data Centre (BADC). The questionnaire allows users to interactively produce CIM metadata documents without any knowledge of CIM structures. The questionnaire provides support for documenting all the major information types a modeler controls when he performs a CMIP5 experiment: the models used, their components, inputs, configuration parameters, the simulation conformance to the specific CMIP5 experiment and the computational platforms used.

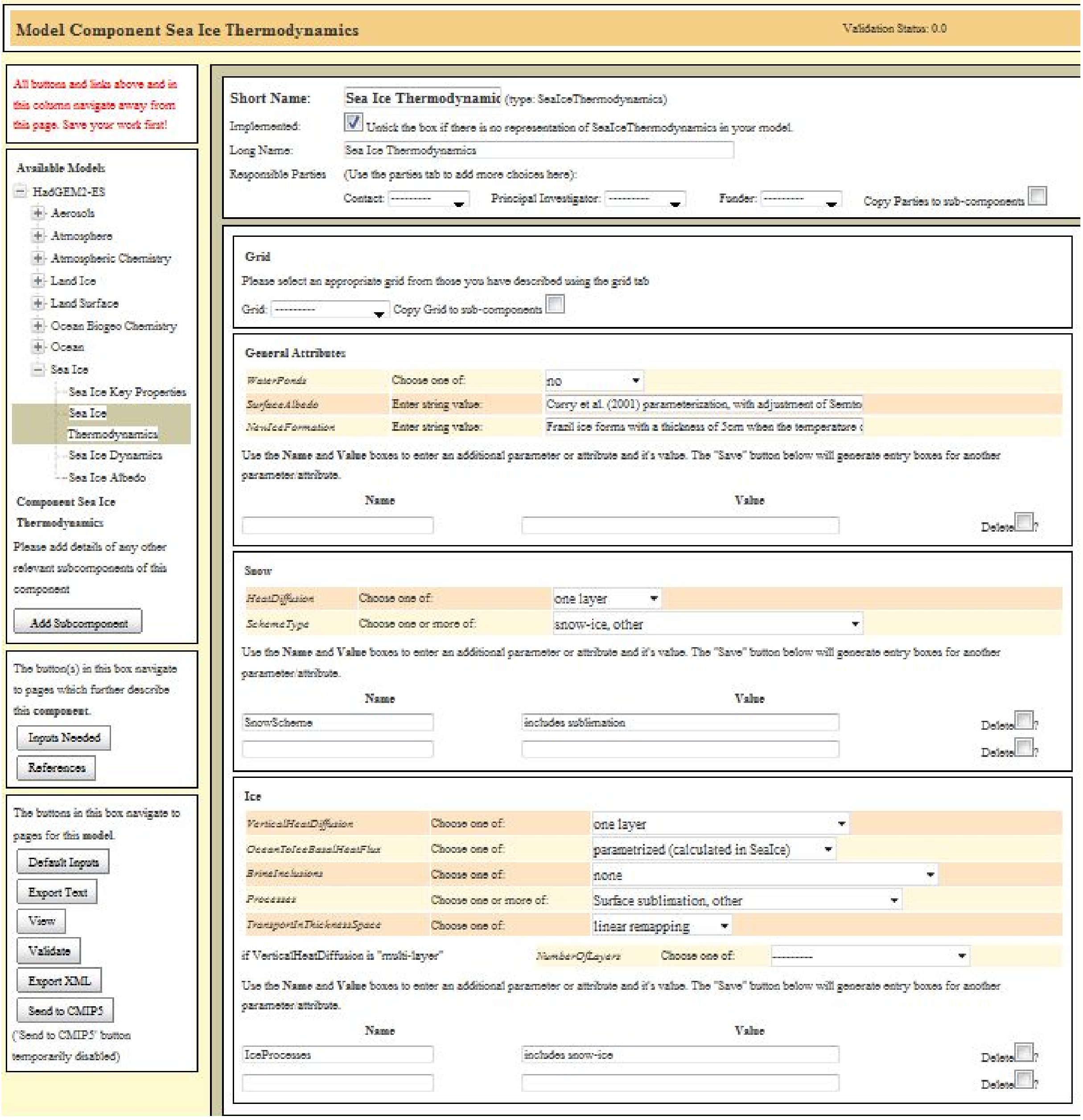

Figure 4: *A page from the CMIP5 questionnaire, showing the use of the CV for Sea Ice Thermodynamics. Note how the first box on the right uses the controlled name for the component, the second box asks which grid the component uses, and the next three boxes ask questions which are driven by the content of the CV for that component. The navigation tree on the left exploits the CV's structure to provide a hierarchical view of the possible components within an earth system model. Other buttons in the middle left provide the ability to navigate around the Questionnaire (e.g. file inputs, references, simulation descriptions, etc.).*

Once complete, the questionnaire supports three levels of validation: (i) the CV constraints are directly enforced within the component descriptions; (ii) when documents are exported, XML schema validation is automatically enforced; (iii) a Schematron (http://www.schematron.com) based validation is performed to check coherency between the different parameters.

The CIM5 questionnaire requires the modeller to describe their model using the CV if possible, but is also extensible supporting user defined parameter-value attributes for each component, and indeed arbitrary additional component structures.

Under the hood: the metadata pipeline into the Questionnaire

The methodology used for the construction of the CV was intended not only to ensure that metadata documents are populated with comparable terms, but also so that the Questionnaire can exploit the CV easily, especially since the CV is expected to evolve more rapidly than the CIM itself.

To that end, we have developed the software to support an information pipeline (Figure 5). A simple XML CV structure was defined to encode the CV along with the mind map rules and constraints described earlier. A mind map validator, written in XSLT and invoked by Python, was implemented to check that a specific mind map instance conforms to the defined rules. The resulting Metafor CV XML documents are then imported into the Django tables and are used to drive the Questionnaire graphical interface. The end result is that the structure of the component pages -in terms of, for example, hierarchy presented, order of the parameters asked about- is completely under the control of the originating mind map. This flexibility has of course been crucial in the development of the questionnaire. Other questionnaires are under development which will be able to exploit new CVs such as that required to describe various downscaling algorithms (e.g. [3]), and Integrated Assessment Models [4].

Other tools support other transformations: for example, a tool has been developed to convert the CV into an OWL ontology which is used within the ESGF to support gateway interfaces and guide faceted browsing. This ontology is also used to guide the mapping tool which allows the conversion of CIM documents into gateway triples. These tools (and all Metafor software and schema) can be found on the METAFOR SVN repository at http://proj.badc.rl.ac.uk/metafor.

Controlled Vocabulary for Simulations and experiments

Although the Model CV discussed is valid for any climate model, the vocabulary necessary for experiments, particularly model inter-comparison, is dependent on the experiment context and aims. The CIM requires that Experiments are identified by a label, a title, a description and its associated list of requirements. We encoded the entire set of CMIP5 experiments using these categories to build a "CMIP5 experiment controlled vocabulary".

In the CIM, the way a simulation conforms to a given experiment is captured by a "Conformance", which describes how a model configuration (inputs, coupling, parameters, etc.) meet the experimental requirements. The development of the model and experiment CV was a key for capturing metadata information using the CMIP5 questionnaire. The full suite of CMIP5 experiments can be described by their experiment requirements, each requirement is given a unique name but where a requirement is the duplicate of an identical requirement in a different experiment both are given the same identifier. The full list of CMIP5 experiment requirements is available from the Metafor subversion repository[1].

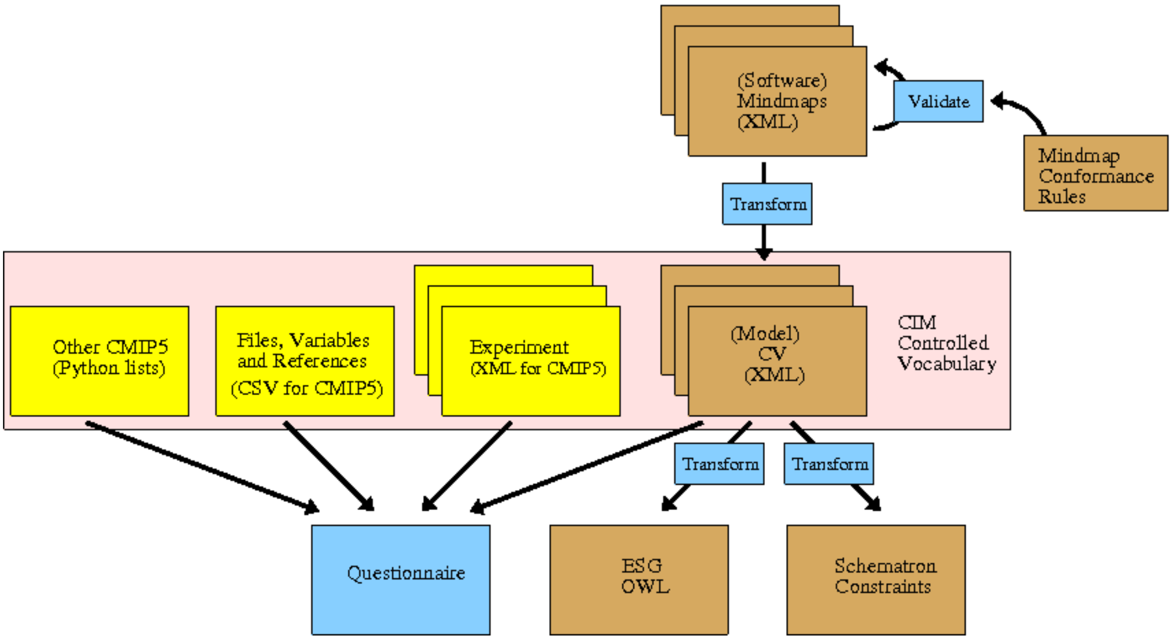

Figure 5: *Key components of the CV and information pipeline. Brown boxes denote generic CIM CV elements while the yellow boxes are for, CMIP5 specific controlled vocabularies. A CV XML is derived from original mind map format (Transform blue box at top) after validation check against the set of rules (Validate blue box at top-right) for ingestion into the Questionnaire (Blue box at bottom-left). Additional translation tools (remaining Transfom blue boxes) are for conversion into ESG ontology and into other ancillary formats.*

Future Development of Controlled Vocabularies

The CV is independent of the CIM and is expected to evolve more rapidly than the CIM itself. We expect that experience using the tool that we developed to harvest metadata for models contributing to the CMIP5 exercise (discussed above) will engender many suggestions of improvements from modelling groups. Thus, while the current version of the CV is the first one encompassing all the components of a coupled climate model, we will have to manage its evolution and ensure its preservation.

Extending the CMIP5 Controlled Vocabulary

The online web interface driven by the CV enables users to enter additional name and value pairs in places where they decide to describe their model components in more detail than that which is specified by the Controlled vocabulary. Also the online interface allows users to select “other” where the list of controlled values is insufficient to describe the nature of their particular configuration. Where users select “other” they are encouraged to enter a verbose description of what that other should be. The additional

[1] http://metafortrac.badc.rl.ac.uk/trac/browser/cmip5q/trunk/cmip5q/cmip5q/data/experiments

information collected by the online interface for CMIP5 will be used to inform the initial expansion of the CV, this process will need to be governed. To that end, an international governance committee is currently being set-up and is discussed later in this document.

Building new Controlled Vocabularies

The JISC funded PIMMS (Portable Infrastructure for the Metafor Metadata System) project will provide individual institutions with the ability to build new controlled vocabularies. The stand-alone PIMMS infrastructure will enable existing CV to be extended and new CV to be created to drive new questionnaires. PIMMS class CV will require a metric to define the scope of the community it supports. A clear measure of the scope of any CV is essential to understand how widely the CIM documents created with it can be shared. PIMMS will develop guidelines for the creation of new controlled vocabularies to ensure that their scope is clearly defined. In preparation for PIMMS the Metafor mind map validator code has been incorporated into the CIM portal so that it can be used to ensures that the structure of the new mind maps are able to drive the questionnaire.

Governance

Any plans for the future of the Metafor CV must include specification for the hosting of the CV as well as protocols and guidelines to govern CV development. Here we address both of these issues.

Hosting Controlled Vocabularies

The controlled vocabulary developed by Metafor for CMIP5 will require more than a simple vocabulary server to host it because it encodes relationships between variables and parameters including conditional statements and XOR and OR assignments (figure 2 and figure 3). Where possible the method of hosting of CV beyond Metafor should also make use of existing standards and conventions.

The ISO TC211 set of standards for geospatial information includes a standard for the description of Feature Catalogues (ISO 19110). These are intended to describe the domain of discourse for an application schema. Most commonly, these are applied to describing the classes and properties (attributes, associations and possibly operations) that are contained in an application schema, as these usually reflect the domain of discourse. In the case of Metafor, the application schema is mostly dedicated to representing the data that is used to describe various aspects of climate models, focussed around the use of these models for simulations of earth system processes. However, the Metafor controlled vocabulary describes the real world aspects of the domain of discourse relating to Metafor: the actual earth system processes whose modelling and simulation the Metafor CIM describes.

Thus for Metafor, in addition to the conventional Feature Type Catalogues to describe the classes, attributes, associations and operations in the CIM, a Domain Feature Catalogue (DFC) has been designed as a separate application schema (http://proj.badc.rl.ac.uk/svn/metafor/DomainFeatureCatalogue), to provide all of the functionality of the Metafor controlled vocabulary, with the addition of semantic relationships that can be used for a foundation for semantically-enriched description and discovery. The DFC could be automatically populated from the existing Metafor controlled vocabulary.

Appendix A shows how the current Metafor XML controlled vocabulary maps to an ISO 19110 Feature Type Catalogue, and Appendix B contains a sample XML encoding of part of the Metafor controlled vocabulary (part of the Ocean realm). This sample uses imaginary URIs and values from the CF controlled vocabulary for illustrative purposes. Table 1 shows the general UML to XML mapping rules that were used.

UML	**XML**
Class	Element
Attribute	Element with the name of the attribute, being a child element of the element representing the class of which it is an attribute.
Association	Element with the name of the association role, being a child element of the element representing the class of which it is an association, with the associated class being a child element of the element representing the association.
Aggregation/Composition	Element with the name of the aggregated/composed class, being a child element of the element representing the aggregating/composing class.
Association class	Element with the name of the association role, being a child element of the element representing the class of which it is an association, with attributes and associations of the association class being attributes of the element representing the association role.
Generalisation/Inheritance	Classes at the more general end of all generalisation associations are treated as abstract classes, whose properties are inherited by the child (more specialised) class, cascading to the most specialised class.

Table 1: *UML to XML Mapping for Domain Feature Type Catalogue*

It would also be possible to develop an OWL ontology from the DFC to provide additional abilities to express the semantics of the domain concepts, which may be appropriate for future applications of the CIM that require a semantically richer representation, with some possibility of inference or other kinds of reasoning. An OWL ontology could be generated automatically from the Domain FTC. A number of mappings from an XML instance document to an OWL instance document have been published, and the xml2owl[2] tool is available to generate OWL from XML. Alternatively, an xslt stylesheet could be developed to accomplish the same goal.

The CIM contains a number of enumerations (which are not extensible) and code lists (which may have new values added). The code lists are referred to as either open or closed (as shown in tagged values attached to the code lists). Open code lists may have new values added by users in the course of their metadata entry, while closed code lists require some kind of community agreement to have new values added. The enumerations and code lists included in the CIM are stored in the CIM FTC (and in the

[2] http://xml2owl.sourceforge.net/index.php

UML model), and are distinct from the sets of values that are stored in the DFC. The DFC represents the domain features whose simulation is described in the CIM, and the properties thereof, while the CIM FTC represents the feature types in the CIM, and the properties of those feature types, together with the code lists and enumerations that support them and describe the metadata contained in the CIM (for example, types of relationship between observations; types of ensemble).

Governance

The maintenance and extension of Metafor Controlled Vocabularies is ultimately the responsibility of the CIM governance committee. However, it is expected that governance of the CV will be relatively non-contentious such that day-to-day vocabulary maintenance can be carried out by a separate CV group.

The CV group will be made up of members from the BADC and the University of Reading who have been working on the questionnaire for CMIP5. The initial role of the CV group will be to extend the CMIP5 CV by making use of the additional information provided by the modelling groups when they completed the CMIP5 questionnaire.

Although the CV will be hosted in the form of a Domain Feature Catalogue (DFC) all communication of CV between the CV group and the community and the CV group and the CIM governance committee will be done using the mind-map interface.

As part of the PIMMS project, metrics will be defined that will define the scope of the community that existing, extended and new CV are built to support. These metrics will be offered to the CIM governance committee for approval.

Conclusions

The Metafor project has established a new controlled vocabulary for the description of climate models and simulations. In doing so the project has also defined a methodology for collecting this information and guidelines for interaction with domain experts. We have developed a prototype system for collecting controlled vocabularies and tools to exploit it. Further to that we have also established a protocol for storing controlled vocabularies that is both independent of the CIM and standards compliant.

References

[1] Redler, R., S. Valcke and H. Ritzdorf (2010) A Coupling Software for Next Generation Earth SystemModelling. Geosci. Mod. Dev., 3, 87-104

[2] Ford, R.W., G.D. Riley, M.K. Bane, C.W. Armstrong and T.L. Freeman (2006). GCF: a general coupling framework. Concurrency Computat. Pract. Exper. 18:163–181

[3] Boé, J., L. Terray, F. Habets, et E. Martin, 2006: A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling. J. Geophys. Res., 111, D21106

[4] Ackerman, F. and DeCanio, S.J. and Howarth, R.B. and Sheeran, K (2009). Limitations of integrated assessment models of climate change. Clim. Change 95, 297-315.

Appendix A: Controlled vocabulary mind maps

Links to the controlled vocabulary mind maps developed in consultation with climate scientists in support of CMIP5 and held in the Metafor subversion repository.

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/Aerosols_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/Atmosphere_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/AtmosphericChemistry_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/LandIce_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/LandSurface_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/Ocean_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/OceanBioChemistry_bdl.mm

http://metafortrac.badc.rl.ac.uk/trac/browser/controlled_vocabularies/trunk/Software/SeaIce_bdl.mm

Appendix B: Mapping from Controlled Vocabularies to ISO 19110 and the Domain Feature Catalogue

Controlled Vocabulary (XML)	**ISO 19110**
Realm	DFC_DomainFeature
parent-child relationship between realms	DFC_DomainFeatureAssociation (this expands the FC_InheritanceRelation class from ISO 19110).
component	DFC_DomainProperty
component nesting within other components	DFC_DomainPropertyAssociation (this expands the attributeOfAttribute role included in ISO 19110). Note that the characterizes relationship is intended to be used for properties that are subproperties of another.
parametergroup	DFC_DomainProperty
parameter	DFC_DomainProperty. Components, parameter groups and parameters are all different levels of attributes/properties of a realm (feature type), rather than feature types themselves.
value	Multiple attributes of FC_FeatureAttribute inherited by DFC_DomainProperty
value format	FC_FeatureAttribute.valueType inherited by DFC_DomainProperty
value units	FC_FeatureAttribute.valueMeasurementUnit inherited by DFC_DomainProperty
value name	FC_FeatureAttribute.listedValue and FC_ListedValue, inherited by DFC_DomainProperty. Alternatively, valueType may refer to a URI which could contain the set of possible values instead of storing them in FC_ListedVale.
choice (XOR, OR, keyboard) (for parameter)	DFC_LogicalChoiceType
definition (for component, parameter, value)	FC_FeatureType.definition (for component), inherited by DFC_DomainFeature; FC_PropertyType.definition (for parameter) and FC_ListedValue.definition (for value), inherited by DFC_DomainProperty.
definition status (missing, present)	Not included, because it can be derived from place holder text?
name (applies to parameter, parametergroup and component)	Can be derived from the rawName. Alternatively, FC_FeatureAttribute.code may be used.
rawName (applies to parameter, parametergroup and component)	FC_PropertyType.memberName
esgName (applies to parameter, parametergroup and component)	Can be derived from the rawName
Constraint (condition)	FC_Constraint
Constraint parameter	constrainedBy role of association with FC_FeatureAttribute, (via FC_CarrierOfCharacteristics), which may contain a set of listed values, and associations attached to DFC_DomainConstraint.
componentView	Not part of the CV, so not modelled. Could be added as an additional property of component if needed.

Appendix C: Sample XML Encoding of the Domain Feature Catalogue (using part of the Ocean realm from the Metafor Controlled Vocabulary)


```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Sample file of possible XML encoding of the Domain Feature Catalogue for part of the Ocean Realm in the Metafor Controlled Vocabulary-->
<FC_FeatureCatalogue>
    <name>Metafor Domain FTC</name>
    <scope>Climate Science</scope>
    <versionNumber>0.1</versionNumber>
    <versionDate>10-Jun-2011</versionDate>
    <producer>
        <organisationName>BADC</organisationName>
        <role>owner</role>
    </producer>
    <DFC_DomainFeature>
        <typeName>Ocean</typeName>
        <definition>Definition of component type Ocean required</definition>
        <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#Ocean</code>
        <!-- made up example relations with other processes -->
        <relatedFeature>
            <DFC_DomainFeatureAssociation>
                <featureAssociationRole>subType</featureAssociationRole>
                <relatedFeature>
    <linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#EarthSystemModel</linkage>
                </relatedFeature>
            </DFC_DomainFeatureAssociation>
        </relatedFeature>
        <relatedFeature>
            <DFC_DomainFeatureAssociation>
                <featureAssociationRole>equivalent</featureAssociationRole>
                <relatedFeature>
                    <linkage>http://a.made.up.url/cf_controlledvocab#Sea</linkage>
                </relatedFeature>
            </DFC_DomainFeatureAssociation>
        </relatedFeature>
        <DFC_DomainProperty>
            <memberName>OceanKeyProperties_</memberName>
    <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyProperties_</code>
        </DFC_DomainProperty>
        <DFC_DomainProperty>
            <memberName>OceanKeyPropertiesAttributes_</memberName>
    <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyPropertiesAttributes_</code>
            <relatedProperty>
                <DFC_DomainPropertyAssociation>
                    <propertyAssociationRole>characterizes</propertyAssociationRole>
                    <relatedProperty>
```

<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyProperties_</linkage>
</relatedProperty>
</DFC_DomainPropertyAssociation>
</relatedProperty>
</DFC_DomainProperty>
<DFC_DomainProperty>
<memberName>ModelFamily</memberName>
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#ModelFamily</code>
<definition>Type of oceanic model.</definition>
<valueChoice>XOR</valueChoice>
<listedValue>
<FC_ListedValue>
<label>OGCM</label>
</FC_ListedValue>
<FC_ListedValue>
<label>1D (SCM)</label>
</FC_ListedValue>
<FC_ListedValue>
<label>1D (ocean mixed layer)</label>
</FC_ListedValue>
<FC_ListedValue>
<label>shallow-water</label>
</FC_ListedValue>
<FC_ListedValue>
<label>1 layer</label>
</FC_ListedValue>
<FC_ListedValue>
<label>slab</label>
</FC_ListedValue>
<FC_ListedValue>
<label>other</label>
</FC_ListedValue>
<FC_ListedValue>
<label>N/A</label>
</FC_ListedValue>
</listedValue>
<relatedProperty>
<DFC_DomainPropertyAssociation>
<propertyAssociationRole>characterizes</propertyAssociationRole>
<relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyPropertiesAttributes_</linkage>
</relatedProperty>
</DFC_DomainPropertyAssociation>
</relatedProperty>
</DFC_DomainProperty>
<DFC_DomainProperty>
<memberName>BasicApproximations</memberName>

```
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#BasicApproximations</code>
                    <definition>Basic approximations made in the oceanic model.</definition>
                    <valueChoice>OR</valueChoice>
                    <listedValue>
                         <FC_ListedValue>
                              <label>primitive equations</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>non-hydrostatic</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>Boussinesq</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>other</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>N/A</label>
                         </FC_ListedValue>
                    </listedValue>
                    <relatedProperty>
                         <DFC_DomainPropertyAssociation>
                              <propertyAssociationRole>characterizes</propertyAssociationRole>
                              <relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyPropertiesAttributes_</linkage>
                              </relatedProperty>
                         </DFC_DomainPropertyAssociation>
                    </relatedProperty>
               </DFC_DomainProperty>
               <DFC_DomainProperty>
                    <memberName>ListOfPrognosticVariables</memberName>
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#ListOfPrognosticVariables</code>
                    <definition>List of the prognostic variables of the model.</definition>
                    <valueChoice>OR</valueChoice>
                    <listedValue>
                         <FC_ListedValue>
                              <label>potential temperature</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>conservative temperature</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>salinity</label>
                         </FC_ListedValue>
                         <FC_ListedValue>
                              <label>U-velocity</label>
                         </FC_ListedValue>
```

```
                <FC_ListedValue>
                    <label>V-velocity</label>
                </FC_ListedValue>
                <FC_ListedValue>
                    <label>W-velocity</label>
                </FC_ListedValue>
                <FC_ListedValue>
                    <label>SSH</label>
                </FC_ListedValue>
                <FC_ListedValue>
                    <label>other</label>
                </FC_ListedValue>
                <FC_ListedValue>
                    <label>N/A</label>
                </FC_ListedValue>
            </listedValue>
            <relatedProperty>
                <DFC_DomainPropertyAssociation>
                    <propertyAssociationRole>characterizes</propertyAssociationRole>
                    <relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyPropertiesAttributes_</linkage>
                    </relatedProperty>
                </DFC_DomainPropertyAssociation>
            </relatedProperty>
        </DFC_DomainProperty>
        <DFC_DomainProperty>
            <memberName>SeaWater</memberName>
            <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#SeaWater</code>
            <relatedProperty>
                <DFC_DomainPropertyAssociation>
                    <propertyAssociationRole>characterizes</propertyAssociationRole>
                    <relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#OceanKeyProperties_</linkage>
                    </relatedProperty>
                </DFC_DomainPropertyAssociation>
            </relatedProperty>
        </DFC_DomainProperty>
        <DFC_DomainProperty>
            <memberName>EquationOfState</memberName>
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#EquationOfState</code>
            <definition>Formulation for equation of state of sea water.</definition>
            <valueChoice>XOR</valueChoice>
            <listedValue>
                <FC_ListedValue>
                    <label>linear</label>
                </FC_ListedValue>
                <FC_ListedValue>
```

<label>McDougall et al.</label>
</FC_ListedValue>
<FC_ListedValue>
<label>Jackett et al. 2006</label>
</FC_ListedValue>
<FC_ListedValue>
<label>other</label>
</FC_ListedValue>
<FC_ListedValue>
<label>N/A</label>
</FC_ListedValue>
</listedValue>
<relatedProperty>
<DFC_DomainPropertyAssociation>
<propertyAssociationRole>characterizes</propertyAssociationRole>
<relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#SeaWater</linkage>
</relatedProperty>
</DFC_DomainPropertyAssociation>
</relatedProperty>
</DFC_DomainProperty>
<DFC_DomainProperty>
<memberName>FreezingPoint</memberName>
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#FreezingPoint</code>
<definition>Hypothesis on sea water freezing point.</definition>
<valueChoice>XOR</valueChoice>
<listedValue>
<FC_ListedValue>
<label>fixed</label>
</FC_ListedValue>
<FC_ListedValue>
<label>varying</label>
</FC_ListedValue>
<FC_ListedValue>
<label>other</label>
</FC_ListedValue>
<FC_ListedValue>
<label>N/A</label>
</FC_ListedValue>
</listedValue>
<relatedProperty>
<DFC_DomainPropertyAssociation>
<propertyAssociationRole>characterizes</propertyAssociationRole>
<relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#SeaWater</linkage>
</relatedProperty>
</DFC_DomainPropertyAssociation>

```
                </relatedProperty>
            </DFC_DomainProperty>
            <DFC_DomainProperty>
                <memberName>SpecificHeat</memberName>
                <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#SpecificHeat</code>
                <definition>Hypothesis on sea water specific heat.</definition>
                <valueChoice>XOR</valueChoice>
                <listedValue>
                    <FC_ListedValue>
                        <label>fixed</label>
                    </FC_ListedValue>
                    <FC_ListedValue>
                        <label>varying</label>
                    </FC_ListedValue>
                    <FC_ListedValue>
                        <label>other</label>
                    </FC_ListedValue>
                    <FC_ListedValue>
                        <label>N/A</label>
                    </FC_ListedValue>
                </listedValue>
                <relatedProperty>
                    <DFC_DomainPropertyAssociation>
                        <propertyAssociationRole>characterizes</propertyAssociationRole>
                        <relatedProperty>
<linkage>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#SeaWater</linkage>
                        </relatedProperty>
                    </DFC_DomainPropertyAssociation>
                </relatedProperty>
            </DFC_DomainProperty>
            <DFC_DomainProperty>
                <memberName>FreezingPointValue</memberName>
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#FreezingPointValue</code>
                <definition>Sea water freezing temperature.</definition>
                <valueMeasurementUnit>
                    <unitOfMeasure>
                        <uomName>degrees Centigrade</uomName>
                    </unitOfMeasure>
                </valueMeasurementUnit>
                <valueType>Number</valueType>
                <constrainedBy>
                    <DFC_DomainConstraint>
                        <conditionProperty>
                            <DFC_DomainProperty>
<code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#FreezingPoint</code>
                            </DFC_DomainProperty>
                        </conditionProperty>
```

```
                    <conditionValue>
                        <FC_ListedValue>
                            <label>fixed</label>
                        </FC_ListedValue>
                    </conditionValue>
                </DFC_DomainConstraint>
            </constrainedBy>
        </DFC_DomainProperty>
        <DFC_DomainProperty>
            <memberName>FreezingPointComputation</memberName>
    <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#FreezingPointComputation</code>
            <definition>Method used to compute the sea water freezing temperature.</definition>
            <valueType>CharacterString</valueType>
            <constrainedBy>
                <DFC_DomainConstraint>
                    <conditionProperty>
                        <DFC_DomainProperty>
    <code>http://proj.badc.rl.ac.uk/svn/metafor/controlled_vocabularies/trunk/Software/DFC#FreezingPoint</code>
                        </DFC_DomainProperty>
                    </conditionProperty>
                    <conditionValue>
                        <FC_ListedValue>
                            <label>varying</label>
                        </FC_ListedValue>
                    </conditionValue>
                </DFC_DomainConstraint>
            </constrainedBy>
        </DFC_DomainProperty>
    </DFC_DomainFeature>
</FC_FeatureCatalogue>
```