



Common Metadata for Climate Modelling Digital Repositories Newsletter 4 – March 2010

In this issue:

- *Progress with the CMIP5 metadata questionnaire*
- *Grid controlled vocabulary*
- *CIM updates*
- *Plans for Metafor's final year*
- *New members of the Metafor team*

Progress with the CMIP5 metadata questionnaire

Work on the CMIP5 questionnaire is continuing, and the questionnaire is currently in its beta4 version. Feedback from beta testers from modelling groups in the USA, France and UK was collected and used to improve the functionality and usability of the questionnaire.

The development paradigm for the CMIP5 questionnaire required that a working version be available at all times. The questionnaire has evolved through numerous iterations as bugs have been found and fixed, and the functionality has been extended and improved following feedbacks. Initially the regular Alpha deployments of the questionnaire were released only to the Metafor community, but in late 2009 the first Beta deployment was made available to the wider climate community. The iteration cycle of Beta deployments occurred on a longer timeframe than the Alpha deployments and substantial beta upgrades of the questionnaire were given new URLs. The long iteration cycle and the persistence of old versions of the questionnaire ensured that the beta testers were offered a stable product in which to test their model descriptions.

If you are interested in beta testing and commenting on the questionnaire then please let us know by emailing sarah.callaghan@stfc.ac.uk. The questionnaire can be found at <http://q.cmip5.ceda.ac.uk>. Selecting the Example modelling centre will allow you to explore the questionnaire without previously registering for a modelling centre account.

Grid controlled vocabulary

During the past quarter a special effort was done to reach a higher level of standardization for the vocabulary used to describe the numerical grids of the models. The objective is to make the model grid description independent of the description of model components themselves (atmosphere, ocean, land surface, sea ice, etc.) in order to fully exploit the CIM grid schema class capabilities and allow further standard queries for the search tools.

The work consisted of (i) extracting current information about horizontal and vertical grids from existing standards such as the gridSpec, CMOR and the CF convention and additionally from scientist interviews, (ii) expressing and organizing it in a generic way and (iii) identifying and resolving gaps and redundancies. This new structured and controlled vocabulary for grids forms a separated page of the CMIP5 Questionnaire, allowing a model component using the same (horizontal/vertical) grid as another to simply point to a pre-existing grid description. For example, in the situation where Land Surface uses the same horizontal grid as the Atmosphere, the user won't have to describe the horizontal grid twice.



Screen shot of the beta version of the questionnaire, which can be found at <http://q.cmip5.ceda.ac.uk>



CIM updates

Significant changes from the last version of the CIM include:

- The use of IDs and versions within the CIM has been rationalised. A GUID type has been introduced which ensures unique instances. This, along with a version (internal and external) and a metadata schema ID and version are associated with a CIM Record. Other IDs are used throughout the CIM, but simple character strings are used rather than GUIDs (to correspond to how modelling centres name their products).
- The use of codelists to capture controlled vocabularies (CVs) has been expanded. In the soon to be released v1.4, a CIM codelist not only includes all of the vocabulary terms but also details on the vocabulary server. However, these details are not currently being used and a future version of the CIM will integrate external CVs with the CIM.
- CIM documents can reference other documents or parts of other documents. In the latest version of the CIM, a reference includes the concept of a modification. This allows the CIM to refer to a “base” document or element and describe how it has been changed in the current context. This is particularly relevant in light of how the CMIP5 questionnaire handles ensembles; ensemble members are treated as a set of changes made to a single simulation rather than a set of separate simulations. This makes the resultant CIM documents much more compact and efficient.
- The concept of conformance – as in a simulation must *conform* to particular experimental requirements – has been fleshed out based on examples coming from CMIP5. It is now possible to associate a set of numerical requirements with a simulation and describe how the simulation has been modified (or not, as the case may be) in order to meet those requirements. It is also possible to record that a simulation is purposefully non-conformant with certain requirements.

The output of the CMIP5 Questionnaire are actual CIM instances. Many of these changes are the result of ensuring that the CIM structure can accommodate the information provided by the questionnaire (while still retained enough generality to be useful for future user groups beyond CMIP5). The current version of the CIM is available from the subversion repository on the Metafor website at: <http://metaforclimate.eu/trac/browser/CIM/tags/version-1.3>

Plans for Metafor's final year

The Metafor project recently had its year 2 All Partners meeting, which means that the project is now entering its third and final year of work. In year 3 the project will:

- Produce and host an open dissemination workshop (in Autumn 2010). If you would like to attend, please let us know by emailing sarah.callaghan@stfc.ac.uk
- Finalise the CIM, and produce a plan for its governance and development past the end of the Metafor project.
- Develop and operate the tools and services such as the query tool and presentation tool to allow the CIM to be used in an effective and useful way by climate modellers.
- Release the CMIP5 questionnaire and support the questionnaire users in order to capture the metadata for the CMIP5 archive.

New members of the Metafor team

Metafor are pleased to welcome Catalan Institute for Climate Sciences (IC3) and Allworlds Geothinking as subcontractors to the project. IC3 will be joining the work done in WP3 for service content creation and testing, and Allworld Geothinking will use model driven architecture techniques to ensure that the CIM is compliant with GML-UML encoding rules

Metafor at a glance:

Project title: [Common Metadata for Climate Modelling Digital Repositories \(Metafor\)](#)

Web site: <http://metafordclimate.eu>

Project coordinator: Dr Eric Guilyardi (University of Reading, UK and IPSL, France)
email: Eric.Guilyardi@locean-ipsl.upmc.fr

Project manager: Dr Sarah Callaghan (BADC- UK)
email: sarah.callaghan@stfc.ac.uk
tel.: +44 1235 445770
fax.: +44 1235 446140

Project participants:

UREAD	UK
BADC	UK
CERFACS	FR
MPG	DE
CNRS/IPSL	FR
UNIMAN	UK
UKMO	UK
NMA	RO
MeteoF	FR
CLIMPACT	FR
PrinceU	US
Univ.Cantabria	ES