



Common Metadata for Climate Modelling Digital Repositories

Newsletter 1 – June 2009

This newsletter is intended to be a quarterly publication of the Metafor project team, and aims to keep interested parties up-to-date with the recent developments in the project.

In this issue:

- **Metafor goals**
- **Development of the CIM v1.0**
- **Defining and collecting CMIP5 model and experiment metadata**

Metafor goals

The main objective of Metafor is to develop a Common Information Model (CIM) to describe climate data and the models that produce it in a standard way, and to ensure the wide adoption of the CIM. The project seeks to address the fragmentation and gaps in availability of metadata (data describing data) as well as duplication of information collection, and problems of identifying, accessing or using climate data that are currently found in existing repositories.

Metafor will optimize the way climate data infrastructures are used to store knowledge, thereby adding value to primary research data and information, and providing an essential asset for the numerous stakeholders actively engaged in climate change issues (policy, research, impacts, mitigation and private sector).

The first year of the Metafor project has concentrated on the definition of a first version the CIM and, as importantly, on a conceptual framework to develop the CIM, independently of implementations (figure 1).

Development of the CIM v1.0:

The CIM is at the heart of the Metafor project and has therefore involved all project partners and significant input from other climate modelling groups in Europe and the US.

Climate modeling is a complex process with a wide degree of variability between different models and different modeling groups. To accommodate this, the CIM has been designed to be highly generic and flexible. The Metafor partners describe the climate modelling process simply as "an activity undertaken using software on computers to produce data." This process has been described as separate UML

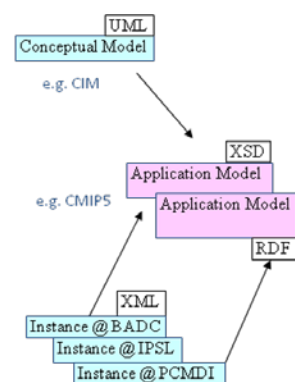


Figure 1: An essential aim of Metafor is that the conceptual model is not changed by the manner in which it is used or applied.



METAFOR is funded by the EU 7th Framework Programme as an e-infrastructure (project # 211753)

packages (and, ultimately, XML Schemas). Figure 2 shows a high-level overview of these packages which include:

- a) activity, the climate modelling simulation/experiments/projects, for example the proposed set of CMIP5 (Climate Model Intercomparison Project) experiments.
- b) software, the climate model as well as any analysis programs used, for example fully coupled atmosphere, ocean, chemistry models.
- c) data, which may be not only the final climate model data served to the community in data centres but could also include data from different stages of the climate modelling process
- d) gridspec, a formal description of the geographic grids modelled by software, required by activities, and mapped to by data
- e) reusable elements, like a quality control mechanism, as well as external standards such as ISO standards (especially the GML series) that need to be used.

This fairly generic structure can be paired with more specific "controlled vocabularies" in order to restrict the range of valid CIM instances. For example, the UML allows for a ModelComponent with child ModelComponents; a controlled vocabulary might restrict that pattern to an atmosphere component with a child radiation component (but not, say, a sea-ice component).

As expected, the high-level CONCIM has greatly facilitated discussion, the result of which CIM 1.0, the abstract CIM, is now available (please get in touch with us, should you be interested to beta test the CIM). The next stage of development will be confronting the CIM with real or virtual *instances* from a range of climate modellers for many different activities. These instances will help explore the flexibility and ability of the CIM to serve many different climate modelling processes.

Defining and collecting CMIP5 model and experiment metadata:

Metafor has been charged by the WGCM¹ via the CMIP² panel to define and collect "type 2" (i.e. model and experiment) metadata for CMIP5 to be used for the next IPCC assessment (due in 2013). To this end, the Metafor team has been actively collecting controlled vocabulary from domain experts to populate the CIM with a standard structure, and is building a web-based questionnaire to collect the associated from the modelling groups participating to CMIP5.

We will shortly be releasing a draft questionnaire for comment to selected individuals. If you are interested in contributing to the questionnaire and/or the controlled vocabulary, or commenting on the draft questionnaire itself, we'd love to hear from you!

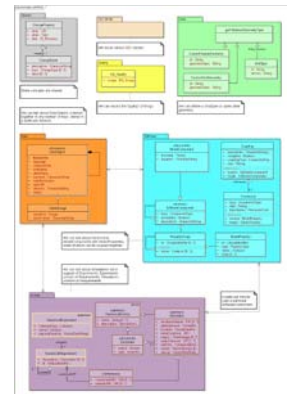


Figure 2: High level overview of CIM packages (full size available from http://metaforclimate.eu/trac/browser/CIM/trunk/packages_abstract.png)

Metafor at a glance:

Project title: [Common Metadata for Climate Modelling Digital Repositories \(Metafor\)](http://metaforclimate.eu)

Web site: <http://metaforclimate.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

¹ http://www.wmo.ch/pages/prog/wcrp/AP_Modelling_WGCM.html

² <http://www.clivar.org/organization/wgcm/cmip.php>