Common Metadata for Climate Modelling
Digital Repositories

METAFOR Deliverable 7.5 - Leaflet presenting findings and recommendations

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	D7.5
Deliverable Title	METAFOR Leaflet presenting findings and recommendations
Document Identifier	METAFOR-D7.5
Date	June 3rd, 2010
Work Package	WP7 Training and Dissemination
Authors	BADC
Document Status	Final
Document Link	http://metaforclimate.eu/documents

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

Abstract

The Metafor team has produced two new leaflets for dissemination purposes. The first provides an overview of the Common Information Model (CIM) and the second discusses the work done by the project on the CMIP5 metadata questionnaire and controlled vocabulary.

It's likely that the CMIP5 questionnaire and controlled vocabulary leaflet will be revised in preparation for the METAFOR final workshop.

http://metaforclimate.eu metafor@metaforclimate.eu

European Commission Information Society and Media

metafor

… addressed the fragmentation and gaps in availability of metadata as well as duplication of information collection, and problems of identifying, accessing or using climate data that are currently found in existing repositories.

The main objectives were to:

- *develop a Common Information Model (CIM) to describe climate data and the models that produce it in a standard way,*
- *to ensure the wide adoption of the CIM.*
- *develop, deploy, and evaluate a prototype infrastructure that allows key data and models to be discovered and compared between distributed digital repositories.*

enes European Network for Earth System Modelling

CLIMATE MODELLING • METADATA • CIM

The METAFOR Project

Project title: Common Metadata for Climate Modelling Digital Repositories (Metafor)

Web site: http://metaforclimate.eu

Project coordinator: Dr Eric Guilyardi (University of Reading, UK and IPSL, France)
email: E.D.A.Guilyardi@reading.ac.uk

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

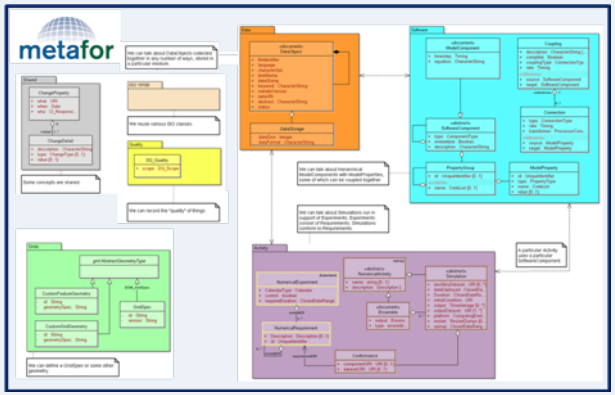
Watch the Metafor and Team Digital Preservation cartoon at http://www.youtube.com/watch?v=76MCRXK4Itc

National Centre for Atmospheric Science · University of Reading · CERFACS · Met Office · BRITISH ATMOSPHERIC DATA CENTRE · UC · M&D · Sciences de l'environnement · climpact · CICS · METEO FRANCE · MANCHESTER 1824

metafor

Common Metadata for Climate Modelling Digital Repositories

The Common Information Model (CIM)



http://metaforclimate.eu

metafor@metaforclimate.eu

SEVENTH FRAMEWORK PROGRAMME

e-infrastructure

METAFOR: The Common Information Model (CIM)

... is a formal metadata model of the climate modelling process

… extends the traditional metadata description of climate data with the provenance of that particular data

… *re-uses* rather than *replaces* existing metadata systems and builds on existing metadata standards used in climate research

... includes descriptions of:

- the experiments being undertaken,
- the simulations being run in support of these experiments,
- the software models and tools being used to implement the simulations
- and the data generated by the software.

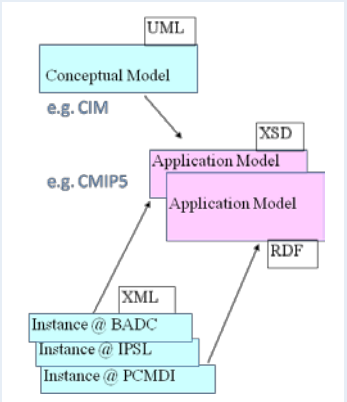
An essential aim of Metafor is that the conceptual model is not changed by the manner in which it is used or applied.

The CIM has been broken down into several packages:

- **Shared** – contains those elements that are used in many different packages.
- **Quality** – contains elements used to express diverse quality metrics for CIM metadata or the artefacts that metadata describes.
- **Grids** – provides a complete description of the horizontal and vertical discretisation of modelling elements: this may refer to grids that data is mapped onto, software adheres to and/or activities constrain.
- **Activity** – specifies the experimental design including the experimental requirements and descriptions of how simulations conform to these requirements.
- **Software** – specifies all the modelling software components used within the experiment process.
- **Data** – describes the data output from the climate modelling process as well as that for any input data.

Climate modelling is a complex process with a wide degree of variability between different models and different modelling groups. To accommodate this, the CIM has been designed to be highly generic and flexible.

We describe the climate modelling process simply as "an activity undertaken using software on computers to produce data." This has been described as separate UML packages (and, ultimately, XML schemas).

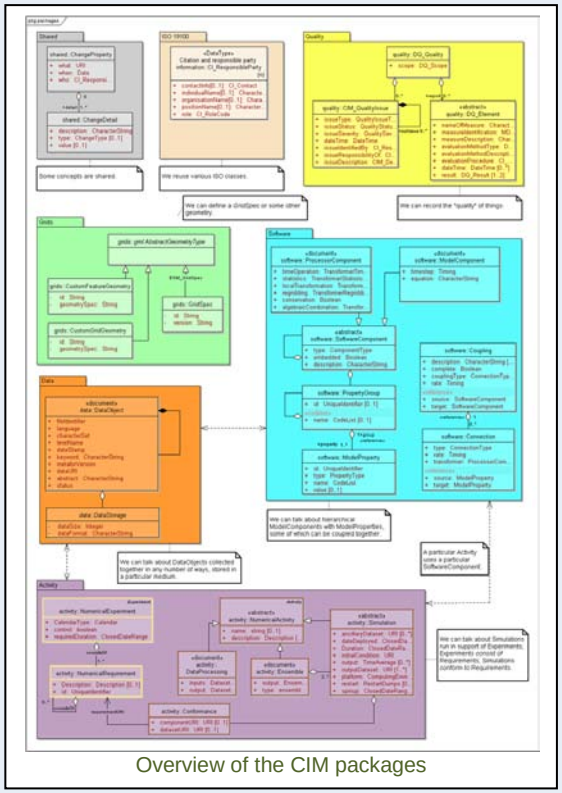


Conceptual CIM (ConCIM)

- An attempt to standardise our understanding of climate modelling at a very high level.
- A "common" model that all interested parties can agree on
- Written in UML
- Gets modified in response to user needs

Application CIM (AppCIM)

- An "application-specific" model (derived from the CONCIM) that is used for a particular community
- Can be written in XSD, RDF or other appropriate language



Overview of the CIM packages

metafor@metaforclimate.eu

http://metaforclimate.eu

European Commission
Information Society and Media

metafor

More people than ever now have a need to discover the results of climate models in order to prepare for and mitigate against the potentially severe impacts of global climate change. But climate modeling is a complex process, which requires accurate and complete metadata (data describing data) in order to identify, assess and use the climate data stored in digital repositories.

The main objective of Metafor was 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. Metafor also developed, deployed, and evaluated a prototype infrastructure that allows key data and models to be discovered and compared between distributed digital repositories.

enes European Network for Earth System Modelling

CLIMATE MODELLING • METADATA • CIM

The METAFOR Project

Project title: Common Metadata for Climate Modelling Digital Repositories (Metafor)

Web site: http://metaforclimate.eu

Project coordinator: Dr Eric Guilyardi (University of Reading, UK and IPSL, France)
email: E.D.A.Guilyardi@reading.ac.uk

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

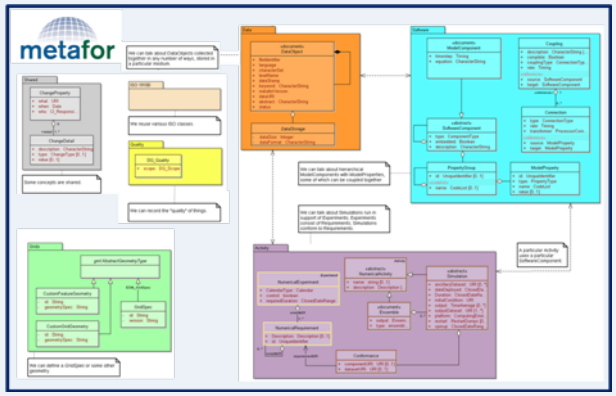
Watch the Metafor and Team Digital Preservation cartoon at http://www.youtube.com/watch?v=76MCRXK4Itc

National Centre for Atmospheric Science · University of Reading · CERFACS · Met Office · BRITISH ATMOSPHERIC DATA CENTRE · UC · M&D · Sciences de l'environnement · climpact · CICS · METEO FRANCE · MANCHESTER 1824

metafor

Common Metadata for Climate Modelling Digital Repositories

The Common Information Model (CIM) and the CMIP5 Metadata questionnaire



http://metaforclimate.eu

metafor@metaforclimate.eu

METAFOR: The Common Information Model (CIM) and the CMIP5 metadata questionnaire

METAFOR

… has been charged by the Working Group on Coupled Modelling (WGCM) via the Coupled Model Inter-comparison Project (CMIP) panel to define and collect model and experiment metadata for the CMIP Phase 5 (CMIP5) project.

…has developed a web-based questionnaire to collect information and metadata from the CMIP5 modeling groups on the details of the models used, and how they conform to the CMIP5 experiment requirements.

The questionnaire is primarily a tool to document models in sufficient detail so that the CMIP5 data can be compared in a scientifically meaningful way.

The CMIP5 questionnaire

…is an ambitious metadata collection tool and will help scientists to provide the most comprehensive metadata of any climate model inter-comparison project.

… allows users to enter descriptions of components which are not already specified by the questionnaire controlled vocabulary.

… XML output complies with the Metafor Common Information Model (CIM), allowing tools and services developed using the CIM to be applied to the questionnaire outputs

...aims to collect enough detail to allow users to easily...

- browse the archive & find desired datasets
- easily differentiate between the "genealogy" (related models & experiments) of datasets

... will populate the CIM repository with CIM instances complying with the CIM ontology

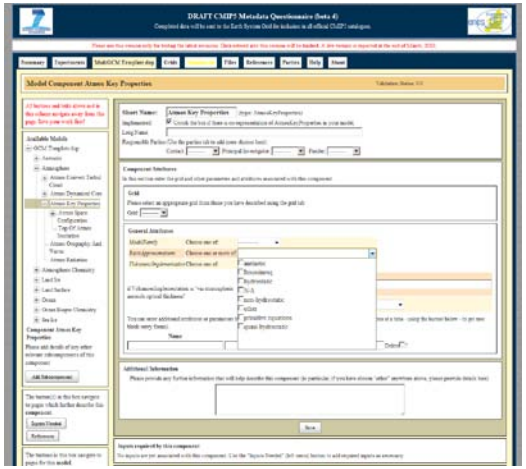
The core archive of model data produced by CMIP5 will be used for the next Intergovernmental Panel on Climate Change (IPCC) assessment, due in 2013.

In METAFOR, we have engaged with the climate modeling community by collecting controlled vocabulary from domain experts, via a series of interviews with climate modelers.

The results of the interviews are interactively summarised in mind maps, allowing us to not only build up the lists of controlled vocabulary, but also build a structure for the way the information is collected.

The controlled vocabulary lists and structure resulting from these interviews will eventually be governed independently from the Metafor project, providing a valuable resource for the climate modelling community.

The mind maps then feed directly into the questionnaire and feedback from scientists about the questionnaire content can be integrated quickly without exposing the questionnaire code.



Screenshot of the questionnaire

SEVENTH FRAMEWORK PROGRAMME
e-infrastructure