



... 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.



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](mailto:E.D.A.Guilyardi@reading.ac.uk)

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

Project participants:

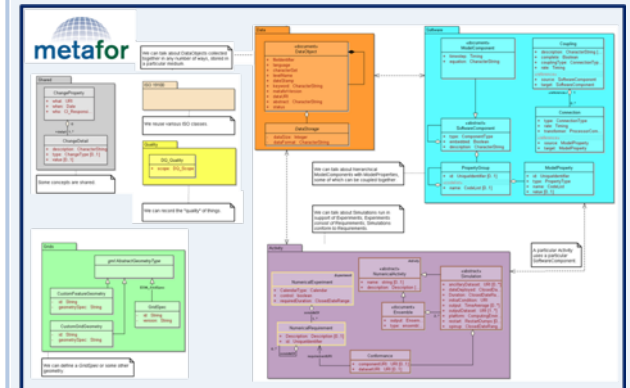
UREAD	UK
BADC	UK
CERFACS	FR
MPG	DE
CNRS/IPSL	FR
UNIMAN	UK
UKMO	UK
NMA	RO
MeteoF	FR
CLIMPACT	FR
PrinceU	US
Univ.Cantabria	ES

Watch the Metafor and Team Digital Preservation cartoon at <http://www.youtube.com/watch?v=76MCRXX4Itc>



Common Metadata for Climate Modelling Digital Repositories

## The Common Information Model (CIM)



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



e-infrastructure



# METAFOR: The Common Information Model (CIM)

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.

... 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

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

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)

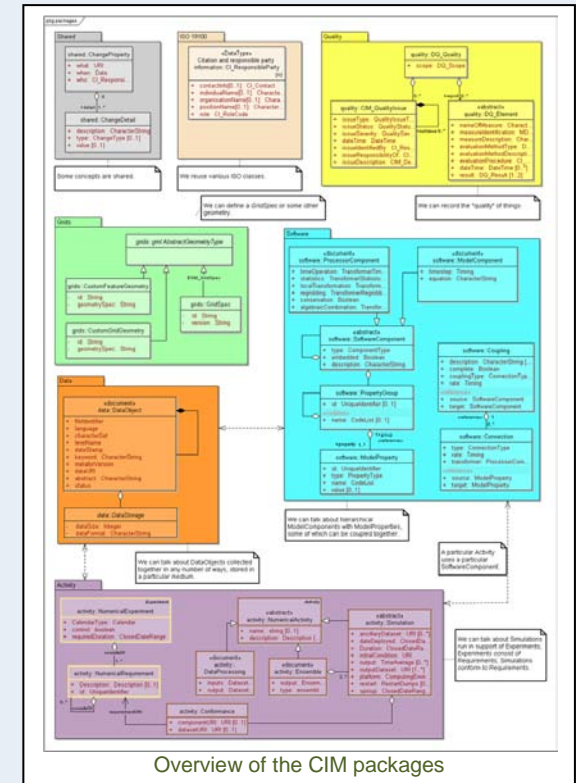
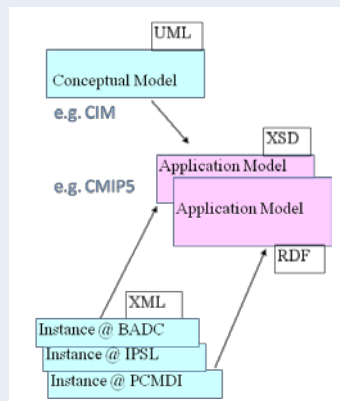
- An attempt to standardise our understanding of climate modelling at a very high level.
- A "common" model that all interested parties can agree on
- Written in UML
- Gets modified in response to user needs

## 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

... 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.



Overview of the CIM packages