



# The Common Information Model for Climate Modelling Digital Repositories: the Metafor Project



Sarah Callaghan<sup>1</sup>, Gerry Devine<sup>2</sup>, Eric Guilyardi<sup>3</sup>, Bryan Lawrence<sup>1</sup>, Charlotte Pascoe<sup>1</sup>, Lois Steenman-Clark<sup>2</sup> and the Metafor Project Team  
<sup>1</sup>NCAS-BADC; <sup>2</sup>NCAS-CMS; <sup>3</sup>NCAS-Climate

## The Problem: Discrimination and Documentation

Interest in the results of climate models is no longer limited to those in academia – governments and the private sector also have a need to discover the results in order to prepare for and mitigate against the potentially severe impacts of global climate change.

Climate modelling 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.

Simulations have a key role in climate science in constructing understanding, and in producing predictions. But discriminating between two simulations is not easy, not even when you were responsible for them!

Documentation currently revolves around (at best) the runtime, but not the scientific detail and relevance of the model components. There is little or no documentation of the "simulation context" (the whys and wherefores and issues associated with any particular simulation).

## The Metafor Team

NCAS leadership is an important part of the Metafor project. The Metafor partners are:

NCAS Climate, NCAS CMS. Reading University, UK  
(Coordinator: Dr Eric Guilyardi)

NCAS BADC (STFC CEDA), UK  
(Project Manager: Dr Sarah Callaghan)

CERFACS, France

Models and Data, Max Planck Institute for Meteorology, Germany

Institute Pierre-Simon Laplace, CNRS, France

University of Manchester, UK

Met Office, UK

Administratia Nationala de Meteorologie, Romania

Météo France, CNRM, France

CLIMPACT, France

CICS, Princeton University, USA



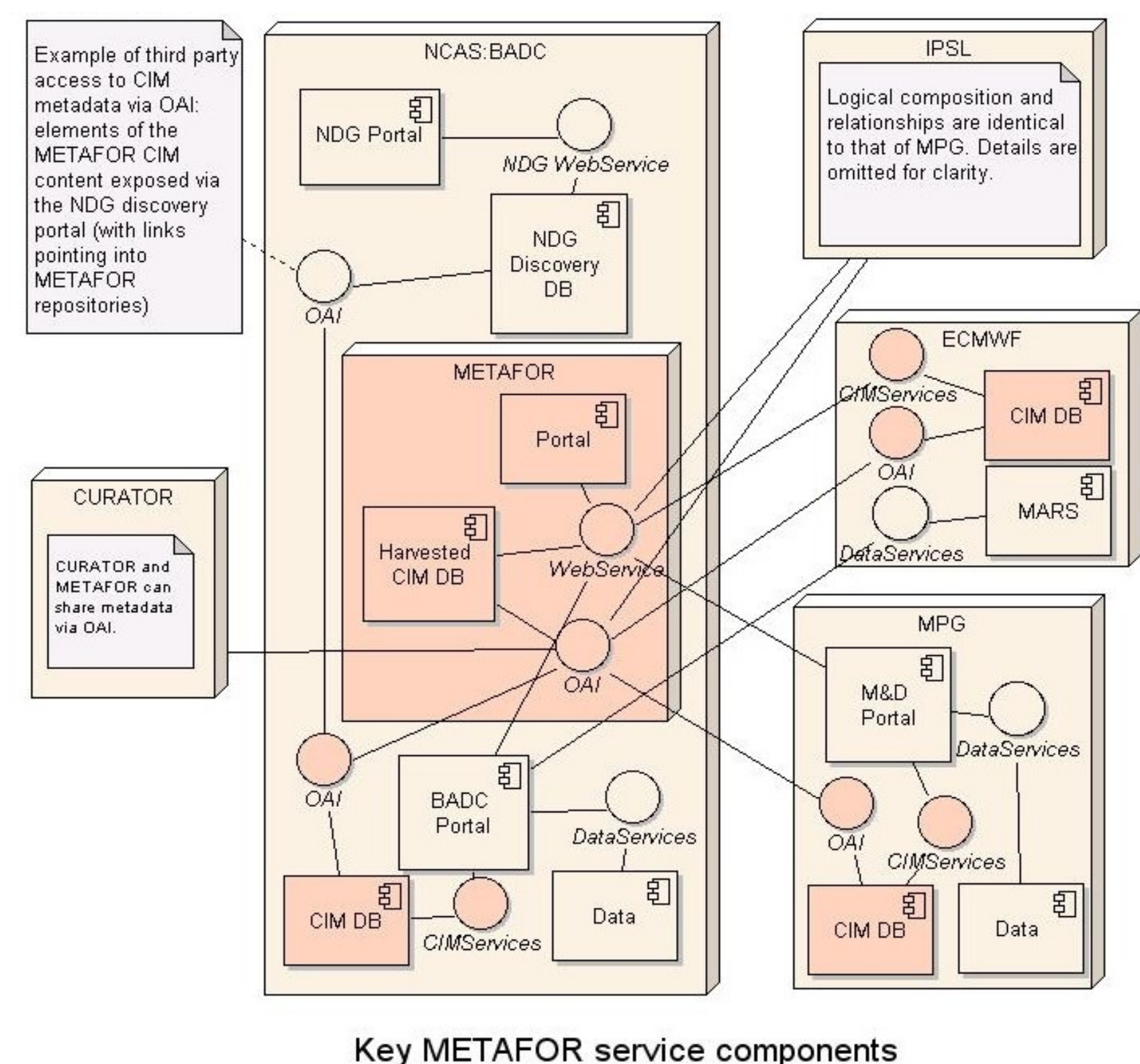
## Metafor Goals

The Metafor project seeks to address 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.

To do this 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.

Metafor will also develop, deploy, and evaluate a prototype infrastructure that will allow (key data and) models to be discovered and compared between distributed digital repositories.

- Develop a "Common Information Model", CIM, to describe all aspects of simulation, from initial activity, to the code, run-time and both input and output data.
- Single sign-on services to populate and manipulate, the CIM metadata
- Centralized CIM content harvested from individual repositories
- Portal to CIM content, links to external (IS-ENES etc) data access services.
- Feed into external portals.

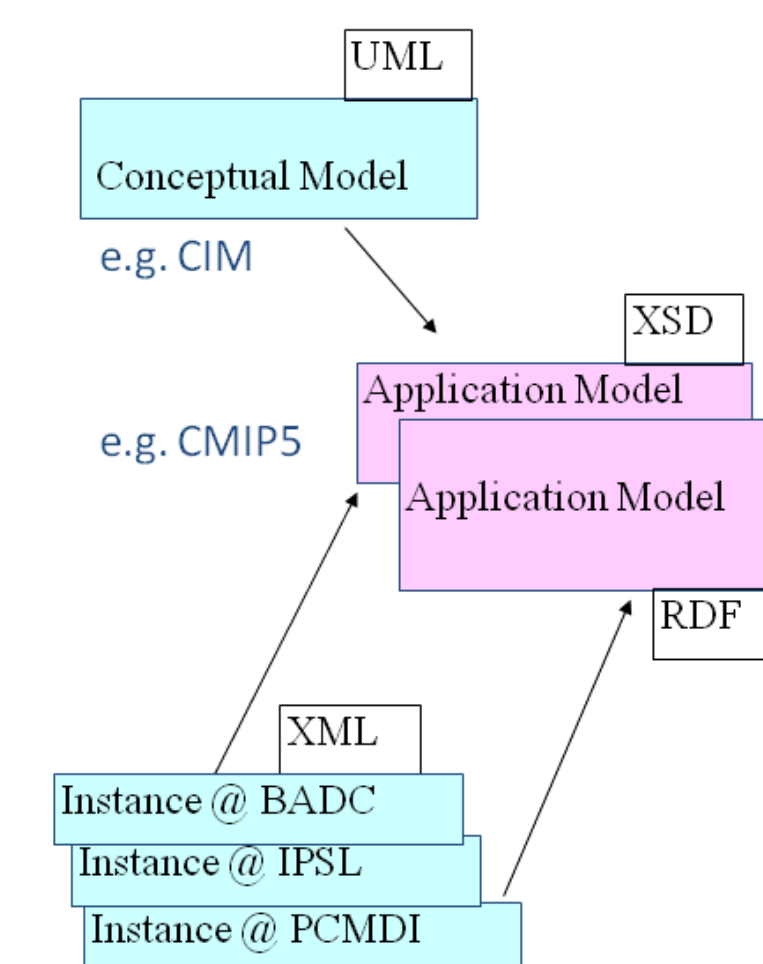


Key METAFOR service components

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).

## Development of the CIM

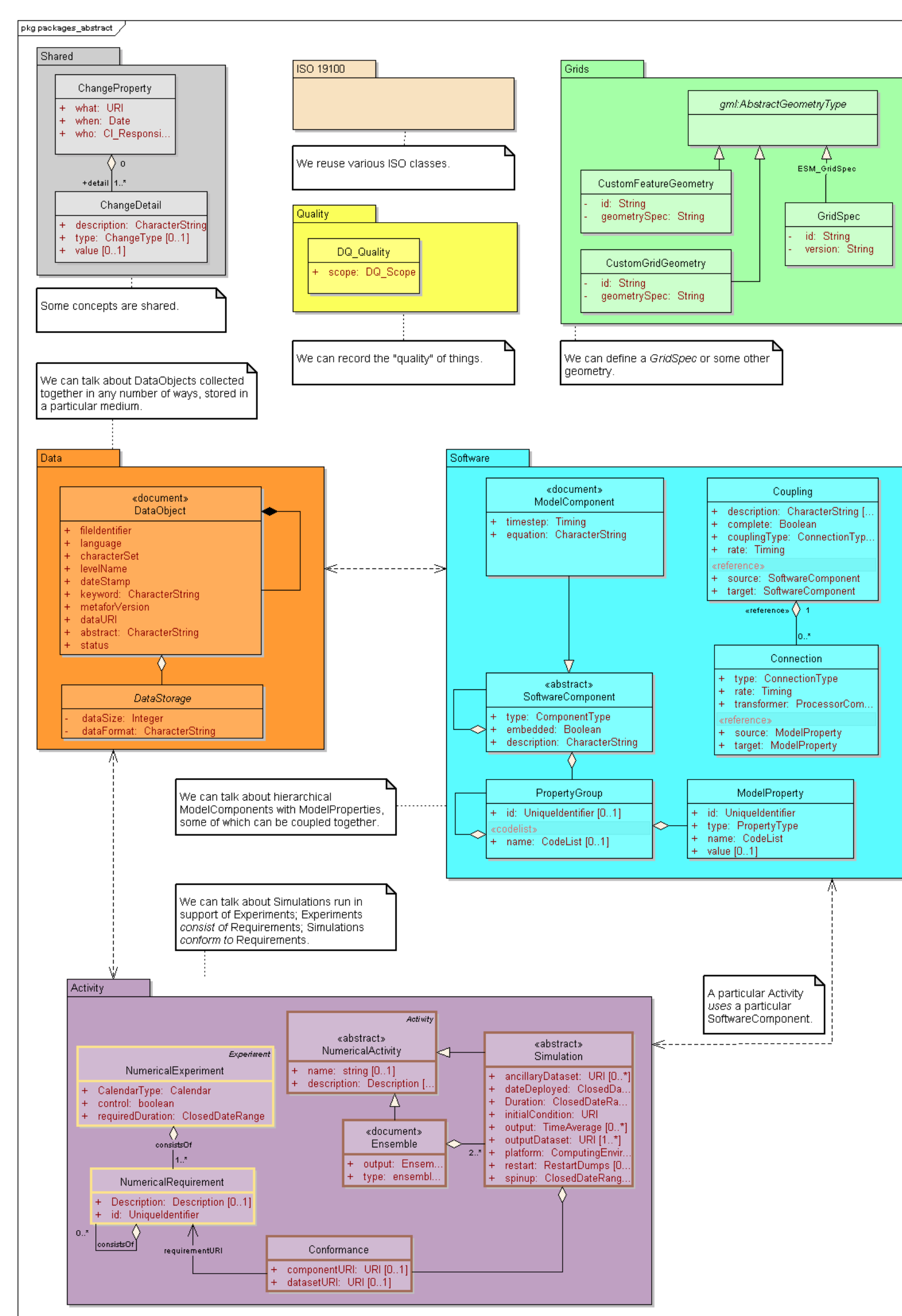
An essential aim of Metafor is that the conceptual model is not changed by the manner in which it is used or applied. Hence multiple application models (APPCIM) can be created using different techniques, but still refer to the same conceptual model (CONCIM).



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

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).



High level overview of CIM packages

The figure above shows a high-level overview of these packages which include:

- activity, the climate modelling simulation/experiments/projects, for example the proposed set of CMIP5 (Climate Model Intercomparison Project) experiments.
- software, the climate model as well as any analysis programs used, for example fully coupled atmosphere, ocean, chemistry models.
- 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
- gridspec, a formal description of the geographic grids modelled by software, required by activities, and mapped to by data
- 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).

The CIM is currently in version 1.3 and is freely available from <http://metaforclimate.eu/trac/browser/CIM/tags/version-1.3>

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.

Further information about the Metafor project can be found at <http://metaforclimate.eu>  
If you have any comments or questions about Metafor or the CIM, we'd love to hear from you!