Common Metadata for Climate Modelling
Digital Repositories

METAFOR EC Dissemination Material METAFOR Deliverable 7.3 M39

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	D7.3 Month 39
Deliverable Title	METAFOR EC Dissemination Material Month 39
Document Identifier	METAFOR-D7.3_M39
Date	June 30th , 2011
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

In the past 6 months, METAFOR has produced 2 more editions in a series of quarterly newsletters, aimed at publicising the work of the project team and encouraging interaction with other climate scientists and EU projects. Also included in this report is a one page document for the general reader about the CMIP5 questionnaire, and a recent general presentation about Metafor.

CAPACITIES

e-infrastructure

Common Metadata for Climate Modelling Digital Repositories

Newsletter 8 – March 2011

In this issue:

Metafor CMIP5 questionnaire splinter meeting at EGU 2011

The Metafor project will be hosting a CMIP5 questionnaire splinter meeting at the European Geosciences Union General Assembly, Vienna, 3-8 April, 2011.
This session will provide an overview and introduction to the CMIP5 model metadata questionnaire developed by the Metafor group. It is particularly aimed at those climate modellers who will be using the questionnaire to capture the metadata for their models and experiments in the context of CMIP5.
All users of the questionnaire and other interested parties are invited to attend.
The meeting will be held in Room SM1 on Thu, 07 Apr, 10:30–12:00
Further details can be found at http://meetingorganizer.copernicus.org/EGU2011/session/8949.

Other presentations about Metafor can be seen at EGU during the following sessions:

- XL162 EGU2011-8459
  - Poster Programme / Hall XL / Attendance Tue, 05 Apr, 15:30–17:00
  - Sarah Callaghan, Mark Morgan, Eric Guilyardi, Sophie Valcke, Charlotte Pascoe, Bryan Lawrence and the METAFOR Project Team "Supporting the climate community by providing common metadata for climate modelling digital repositories: the METAFOR project".
- XL235 EGU2011-7758
  - Poster Programme / Hall XL / Attendance Thu, 07 Apr, 17:30–19:00
  - Gerard Devine, Bryan Lawrence, Charlotte Pascoe, Rupert Ford, Paul Slavin, and Metafor Team "The CMIP5 questionnaire: web-based metadata collection for climate modeling" (solicited)
- ESSI14 Metadata and Data Models, and Markup Languages Oral Programme / Room 19 / Thu, 07 Apr, 17:30–19:00
  - 18:00–18:15 EGU2011-12475 Charlotte Pascoe, Marie-Pierre Moine, Allyn Treshansky, Gerard Devine, Sebastien Denvil, Michel Kolasinski, and Rupert Ford "A Common Information Model paired with scientific Controlled Vocabularies for Climate Models and Statistical Downscaling"

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

Open Workshop about the CIM

The Metafor open dissemination workshop **"Using the Metafor Common Information Model (CIM) to store, discover and locate climate modelling data"**, was held in Abingdon, UK, on Monday 14th March 2011.

The workshop was recorded, and webcasts of the sessions will shortly be available on the Metafor website.

CMIP5 Questionnaire update

The CMIP5 metadata questionnaire was launched in Nov 2010 (http://q.cmip5.ceda.ac.uk), and is now in use by several of the CMIP5 modelling centres. Instructions for gaining access to the questionnaire can be found at that same address. The most obvious update that users will notice is our re-drafting of the decadal experiments. The redraft matches up the list of experiments in the questionnaire with the experiment names that users will be familiar with from the cmip5 experiment design document http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_22Jan11_marked.pdf

The support team are also running live online demonstrations of the questionnaire, to arrange a demonstration for your team please contact cmip5qhelp@stfc.ac.uk.

Filling the model part of the questionnaire has a direct link with IS-ENES for the European ESMs groups as the result will be used to fulfill one IS-ENES deliverable (D4.2), originally due for August 2010 but now delayed until April 2011 (so in a month or so!).

CIM Updates

CIM v1.5 has been released and can be found at http://metaforclimate.eu/trac/browser/CIM/tags/version-1.5. This version will be the operational version used by all the METAFOR tools and services. Work on the CIM will continue to make the CIM fully GML-compatible, and will result in a version 2.0 of the CIM; but this work will not impact the tools and services.

TDS2CIM tool

One main usage of the TDS2CIM tool will be as an integrated tool inside the Metafor portal. The design structure makes it possible to use it locally also as a standalone version. The available functions will be the same, only speed of processing will differ on different configurations. The usage will always be a two path process. The first is to parse at least one TDS (Thredds Data Server) and to store this information in an appropriate postgres data base. This parser software is an output of the development process inside the IS-ENES project. This is an example for a close connection between Metafor and the FP7 EU project IS-ENES. The second path is to map stored information from the data base into the correct CIM XML output format, validate this XML with the Metafor CIM validation class or validation web service and add this reported information to the data description. All this is controlled by an external configuration file which will be used of portal's 'process launcher'.

There is one actual additional approach for the TDS2CIM tool. It will be used to generate metadata from the ECMWF ensembles data and upload them directly into the Metafor portal. To use ECMWF ensembles data inside the downscaling portal of the University of Cantabria this CIM metadata must exist to describe their data content.

Metafor project extension

The Metafor project has been extended for another 6 months, and now will be finishing in September 2011.

Metafor at a glance:

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

Web site: http://metaforclimate.eu

Project coordinator: Prof 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

Common Metadata for Climate Modelling Digital Repositories

Newsletter 9 – June 2011

In this issue:

Metafor project extension

The Metafor project has been extended for another 6 months and will now be finishing in September 2011. The reasons for this extension are to allow us to:

1) Put governance in place, as proposed at M36 in deliverable D2.6, including:
   a. technical developments for AppCIM generation (potentially using the structures already operational in FullMoon) and other CIM processes and tools for easy maintenance and governance
   b. establishment of governance structure and committee for the CONCIM and controlled vocabularies.
2) Continue CMIP5 metadata support, starting with the CMIP5 metadata questionnaire, but also allowing further interaction via use of the questionnaire CIM instances in the Metafor-developed tools and services.
3) Smoothly handover to IS-ENES, who will be taking the software and tools developed by Metafor and developing them for the wider community.

CMIP5 Questionnaire update

The CMIP5 questionnaire (http://q.cmip5.ceda.ac.uk)continues to be used by more modelling groups, and complete questionnaire instances for a number of CMIP5 experiments are now showing up in the atom feed. Currently 37 models are being worked on (coming from 17 of the 24 modelling centres). Recent notable changes have made it easier for a user group to copy previously entered simulations and have conformance information copied at the same time. This should speed up the completion of similar type simulations. The structure of the atom feed page has changed so that the published cim xml documents can be viewed.

Interactions with other projects

Metafor has engaged with a number of projects in order to develop and extend the CIM into areas other than climate modelling.

PIMMS: Portable Infrastructure for the Metafor Metadata System

PIMMS is a consortium proposal lead by the University of Reading with partners including the BADC and the University of Bristol. The aim of the PIMMS project is to package the prototype metadata infrastructure developed by Metafor for CMIP5 so that it can be implemented across multiple institutions and be extended to support other scientific domains.

Ermitage: Enhancing Robustness and Model Integration for The Assessment of Global Environmental Change

Ermitage focuses on the development of interdisciplinary modelling tools and platforms to address the interactions between natural and socio-economic systems. The

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

Metafor CIM will be used to document the models contributing to the Ermitage project by creating new controlled vocabularies for the integrated assessment models that form the bridge between predictions of future climate and their socio-economic impacts. Ermitage will also make use of the CIM portal technology developed by Metafor.

PEG-BOARD: Palæoclimate & Environment data Generation – Building Open Access to Research Data

The PEG-BOARD project focuses on management of palæoclimate data, enabling open access to historical climate data in a systematic, managed environment. PEG-BOARD plan to use the metafor CIM to document the palæoclimate models and simulations in their data repository at the University of Bristol.

http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmi/peg.aspx

We have also developed our close links with the IS-ENES project (https://is.enes.org/) as many of their aims will act in synergy with ours for the benefit of the wider community. Discussions are starting to integrate Metafor next steps in an IS-ENES-2 proposal to further develop the CIM and the associated ecosystem of tools.

TDS2CIM tool

The next extension of the TDS2CIM tool will be the integration of available portal services into the tool. The most important service is the upload service for uploading CIM RecordsSets (here: CIM dataObjects) into the portal. This is not only useful for the portal integrated version of the tool, but it is also useful for the standalone version.

At the DKRZ we use the TDS2XML tool which is an extended version of the TDS2CIM tool. The TDS2XML tool is more generic because it supports the CERA XML format as another output format. At the moment it is used to automatically generate CERA metadata of the published CMIP5 project data.

Metafor coding sprint 4

The end of June has seen the fourth and last of the Metafor coding sprints. A coding sprint is a technique drawn from the Agile software development methodology (http://en.wikipedia.org/wiki/Agile_software_development), where a group of people involved in a project get together to focus on the development of the project. Essentially a team locks itself into a room for a week, blocks out all external interruptions, and endeavors to deliver real and measurable project progress. Coding sprints are particularly useful when team members are scattered across different institutions and countries.

At this coding sprint, held at the DKRZ, Hamburg, good progress was made in delivering the beta-1 version of the CIM Portal (http://www.purl.org/org/esmetadata/cim/portal). New features include a CIM mind-map validation tool and embryonic CIM search web services. Progress was also made with identifying integration points with impact portals run by the University of Cantabria and the KNMI.

There is no doubt that in the absence of coding sprints the Metafor development team would have achieved a far lesser degree of cohesion. Such cohesion has been proved to be directly proportional to the quality of team deliverables. It has been noted that the ESG Curator team have also adopted coding sprints as a means of self organising their development efforts. Hopefully this is an indicator of an increasing maturity in the manner in which software is developed by the Earth Sciences community. The adoption of other Agile techniques such as test driven development, extreme programming, continuous integration ...etc, would undoubtedly reinforce this trend.

Metafor at a glance:

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

Web site: http://metaforclimate.eu

Project coordinator: Prof 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

The CMIP5 Questionnaire

Charlotte Pascoe (28th June 2011)

The CMIP5 Questionnaire[1] web tool captures information about the life-cycle of CMIP5[2] climate simulations. The contextual information it captures explains why and how climate model data was created from the design of experiments (why) to the implementation of experiments via simulations running models (how) (figure 1). The CMIP5 Questionnaire broadens access to climate model data because for the first time researchers can discover the science encoded in the algorithms of climate models without needing to contact the people who wrote the code. With this contextual information, or metadata, climate scientist are able to analyse more deeply the simulated data produced by different modelling groups.

Before metadata about climate models was collected in a systematic way the only documentation available to everyone was that which made it to the scientific literature. However, the scientific literature focused on the latest and greatest things that models could do, the only standard piece of information published about every model was generally its resolution and this was often expressed in different ways depending on the nature of the coordinate system. The only way to really find out about the science encoded in the model was to contact the scientists who ran the model or better still the people who wrote the code. Without personal contacts, climate models were effectively a black box that people and policy makers were expected to trust.

The Metafor[3] project has addressed this issue by creating a new standard for describing simulated data known as the Common Information Model or CIM[4]. The CIM uses UML class diagrams to identify elements that need to be described and the relationships between them. Figure 1 shows a simplified view of the CIM elements that are populated by the CMIP5 questionnaire. The experiments are described as a list of requirements that the simulations must conform to. The simulations run models which are made up of software components and these components can contain child components. The CIM structure is populated using controlled vocabularies which are specific terms, precisely defined that have a common meaning to all climate scientists.

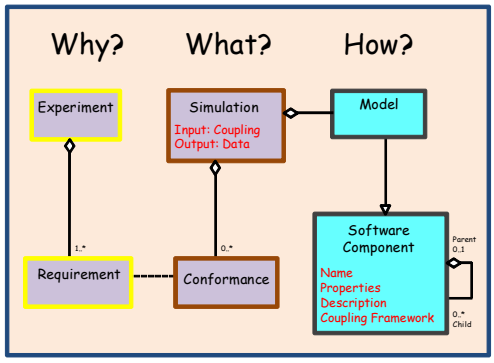


Figure 1, The CMIP5 questionnaire web tool captures metadata about the life-cycle of a climate simulation. Here we see a UML view of the CIM elements that are populated by the CMIP5 questionnaire, they explain why and how the simulated data was created.

The art of the CMIP5 questionnaire has been to bring together the CIM structure with controlled vocabularies that were collected by Metafor from interviews with climate modellers. The aim of the interviews was to find out the information that scientists in different climate disciplines needed to be able to compare climate model simulations. The interviews were interactively summarised using mind map diagrams that not only collated controlled vocabularies but also allowed Metafor to build a hierarchy to structure how the information would be collected in the CMIP5 questionnaire; branches in the mind maps are associated with web forms in the questionnaire and the controlled vocabularies generate drop-down lists.

The CMIP5 questionnaire ensures that a standardised set of metadata is collected across the spectrum of climate modelling domains and includes more than 400 scientific properties. Nevertheless, the hierarchical structure of CIM software components also allows users to describe models in more detail by adding further sub-components to the questionnaire beyond those required by the mind maps. The separation of controlled vocabulary content from the structure of the CIM allows for the questionnaire infrastructure to be specialised for other metadata collection projects by supplying different controlled vocabularies.

The CMIP5 questionnaire is the first attempt to comprehensively describe the science of an Earth System Model. It is a unique community resource that helps scientist to do science and builds trust between scientists and policy makers through openness.

1 CMIP5 questionnaire: http://q.cmip5.ceda.ac.uk
2 CMIP5: 5th Coupled Model Inter-comparison Project http://cmip-pcmdi.llnl.gov/cmip5/
3 METAFOR: Common Metadata for Climate Modelling Digital Repositories http://metaforclimate.eu
4 CIM: Common Information Model developed by Metafor to describe data and the models that produce it in a standard way.

Documenting Earth System Models and Simulations: the Metafor legacy

Eric Guilyardi
and the Metafor team

COMBINE General Assembly
Exeter, May 2011

Climate Model & Experiment Documentation

- What is it ?
  - List of climate model properties
  - Whys and wherefores of simulations
  - Conformance to experimental protocol
  - Standard to describe and compare within a MIP
  - aka “metadata”: data describing data

- What for ?
  - Archive, locate, assess, make sense of climate model data

Example: the CMIP5 Barnum

- 20 modelling centres
- 40+ models
- 60 numerical experiments
- 90,000 years of simulation
- **2 million output datasets !**
- Data to be available from distributed database
- Users need to find datasets, and discriminate between models, and between simulation characteristics.

A real documentation challenge !

Climate Model & Experiment Documentation

- Other reasons why it matters:
  - Share the climate scientist notebook
  - Assess suitability of data for purpose for a widening community of users (impacts,...) - guard against misuse
  - Transparency, quality, accountability insurance
  - Find and understand data in 10-15 years (curation)
  - Add visibility to large institutional effort

Outline

- Why Metafor
- Metafor goals
- How we are doing it
- What we have done
  - Common Information Model (CIM)
  - Controlled vocabulary
  - CMIP5 support and questionnaire
- Next steps: governance and IS-ENES

The life before Metafor

- Finding climate model data is hard
- Understanding data is harder (esp. for non-experts)
- Discriminating between two simulations/models is not easy
- Documentation is “patchy” and specific to modelling centres
- Documentation currently revolves around (at best) the runtime, but not the scientific detail and relevance of the model components
- Little or no documentation of the “simulation context”, i.e. the whys and wherefores and issues associated with any particular simulation.

The Metafor vision

"The open standard developed in Metafor will play a catalytic role in the way next generation climate data repositories, such as IPCC AR5, are organised, preserved and accessed"

"METAFOR will take the first step in doing for climate data and models what search engines have done for the Internet: it will put users of climate data in touch with the information they need"

Metafor goals

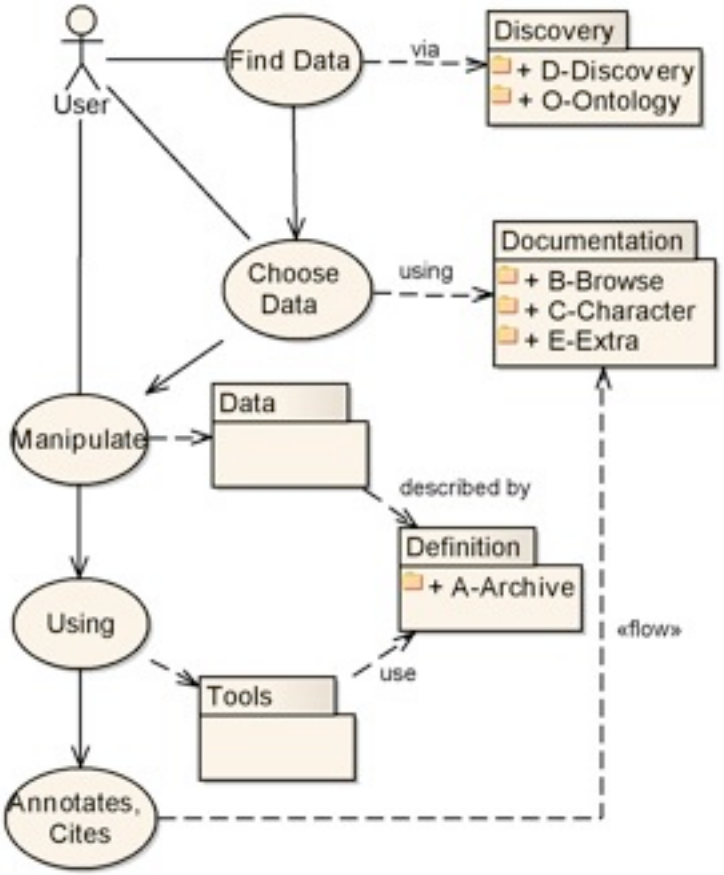
"The main objective of Metafor is to develop a **Common Information Model** (CIM) to describe climate data and the models that produce it in a standard way, and to ensure the wide adoption of the CIM"

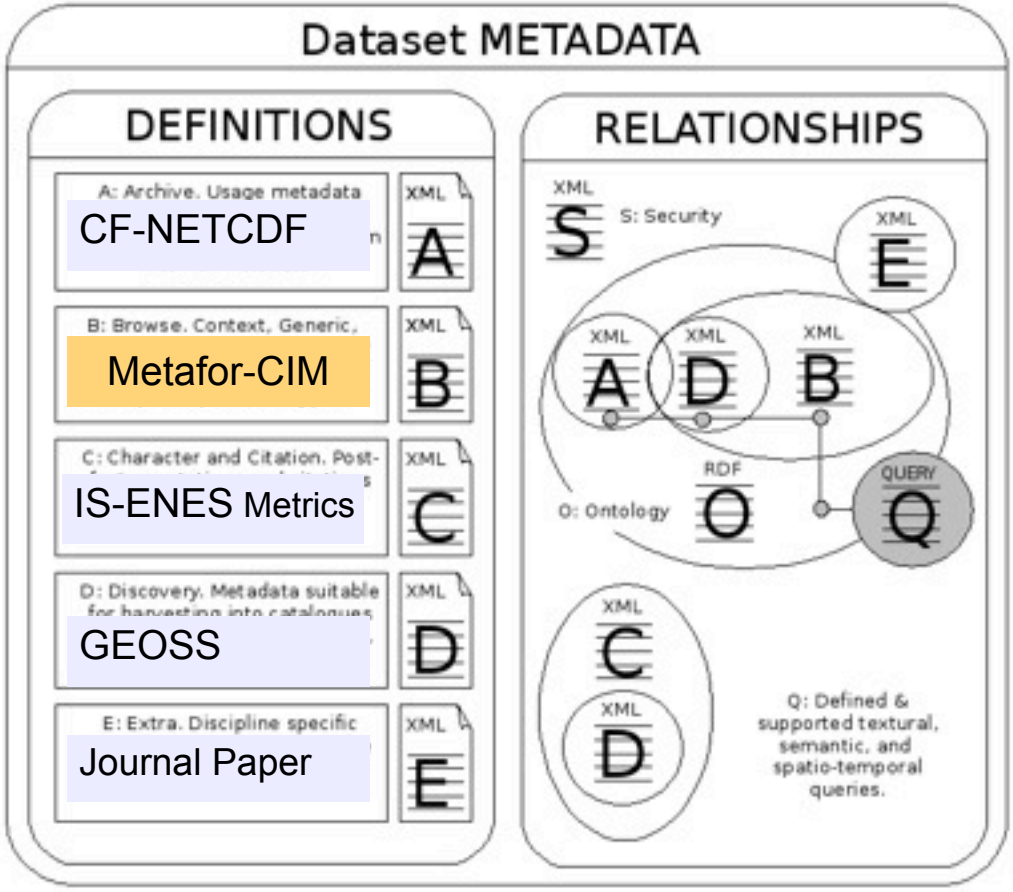
Requirements for success:

- Clearly define scope
- Gather top field experts
- Engage with similar existing activities
- Work towards community adoption
- Capture wider community needs
- Ensure post-project governance

Metafor scope

Discovery, Documentation, Definition





Courtesy Bryan Lawrence (NCAS/BADC)

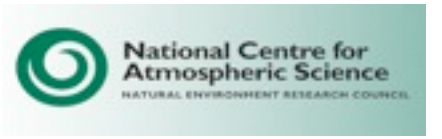
The Metafor Project

e-infrastructure

INFRA-2007-1.2.1
Scientific Digital Repositories

12 partners
EU contribution of 2.2M€
Started March 2008, duration 3.5 years

- NCAS, University of Reading, UK (Coordinator)
- BADC, Science and Technology Facilities Council, UK
- 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 Meterologie, Romania
- Météo France, CNRM, France
- CLIMPACT, France
- CICS, Princeton University, USA
- University of Cantabria, Spain



metafor

Gather top field experts !

METAFOR Year 1 meeting in Abingdon
Feb. 2009

Where Metafor came from

- PRISM project (FP5 2001-2004), ENES
- PRISM Sustained Initiative (PSI)
  - Code coupling and I/O
  - Integration and modelling environments
  - Data processing and management
  - Meta-data standards (key !)
  - Computing issues

Where Metafor came from

- PRISM project (FP5 2001-2004), ENES
- PRISM Sustained Initiative (PSI)
  - Code coupling and I/O
  - Integration and modelling environments
  - Data processing and management
  - Meta-data standards (key !) → Metafor
  - Computing issues

Where Metafor came from

- PRISM project (FP5 2001-2004), ENES
- PRISM Sustained Initiative (PSI)
  - Code coupling and I/O → IS-ENES
  - Integration and modelling environments → IS-ENES
  - Data processing and management → IS-ENES, Metafor
  - Meta-data standards (key !) → Metafor
  - Computing issues → IS-ENES

Where Metafor came from

- PRISM project (FP5 2001-2004), ENES
- PRISM Sustained Initiative (PSI)
  - Code coupling and I/O → IS-ENES
  - Integration and modelling environments → IS-ENES
  - Data processing and management → IS-ENES, Metafor
  - Meta-data standards (key !) → Metafor
  - Computing issues → IS-ENES

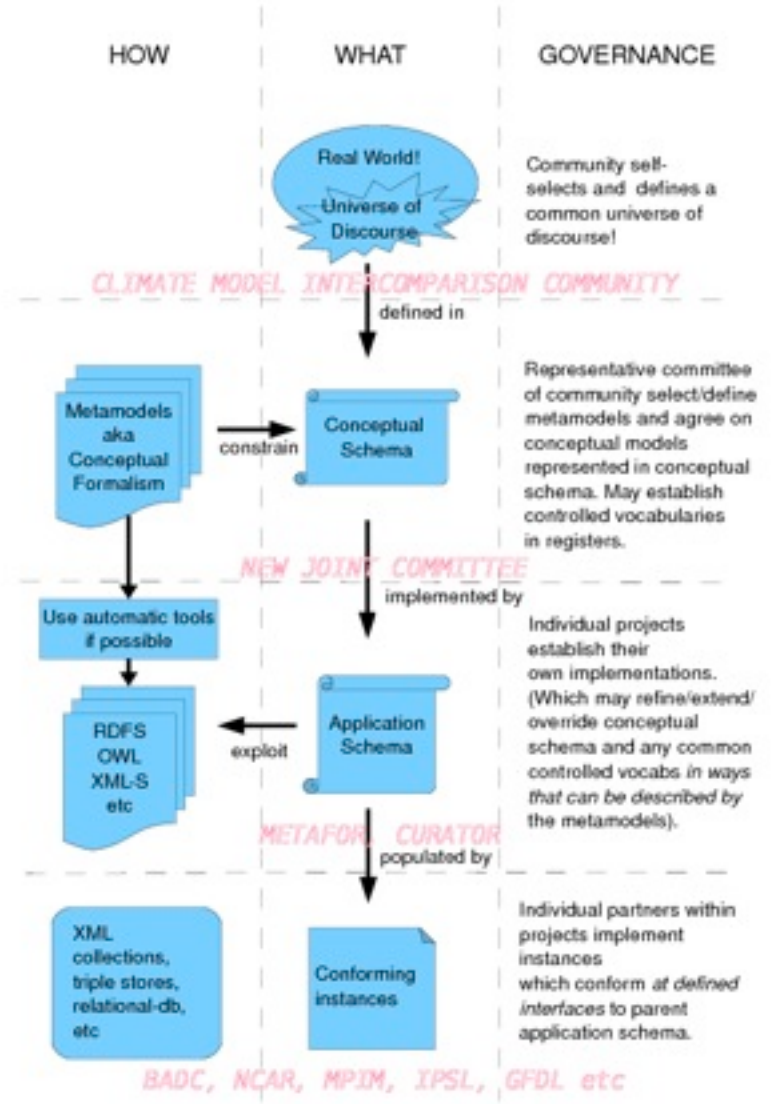
ENES coordination

The Common Information Model CIM

- the CIM builds on existing metadata standards used internationally in climate research (CF, NMM, Curator, FLUME, ISO-standards, etc.) + new bits
- the CIM defines a general structure over which a specific Controlled Vocabulary (CV) can be applied
  - a CV consists of the terms (and their relationships) used to build the content of CIM instances.

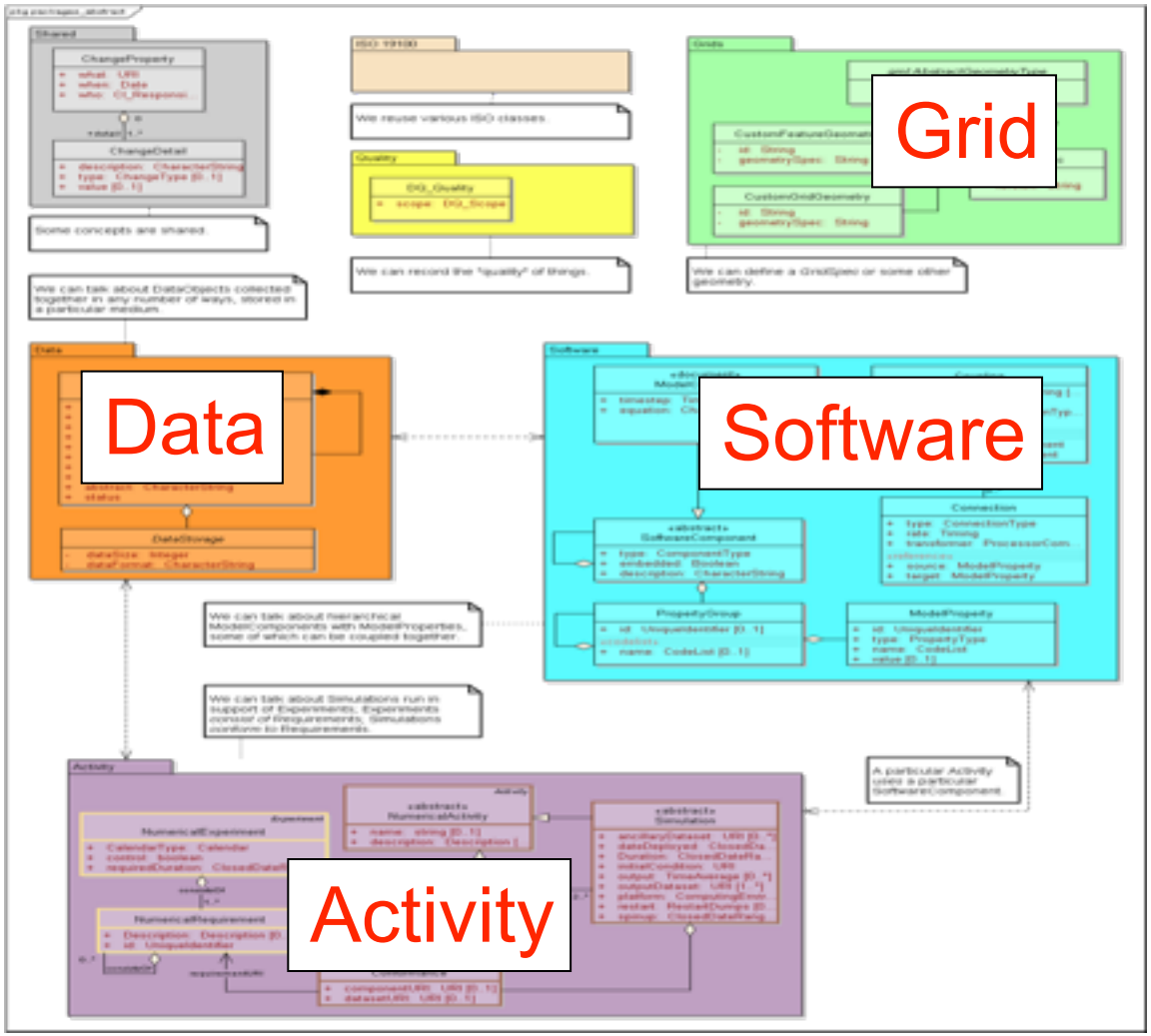
The CIM "meta-model"

- Lots of people talk about climate models and data;
- Some people even agree about those things ("the yolk");
- We have a formal way of describing that (UML, CONCIM);
- That UML is constrained to follow a particular meta-model...
- ...so that it can be transformed into something usable (XSD, APPCIM) for particular users;
- Metadata instances conform to an APPCIM



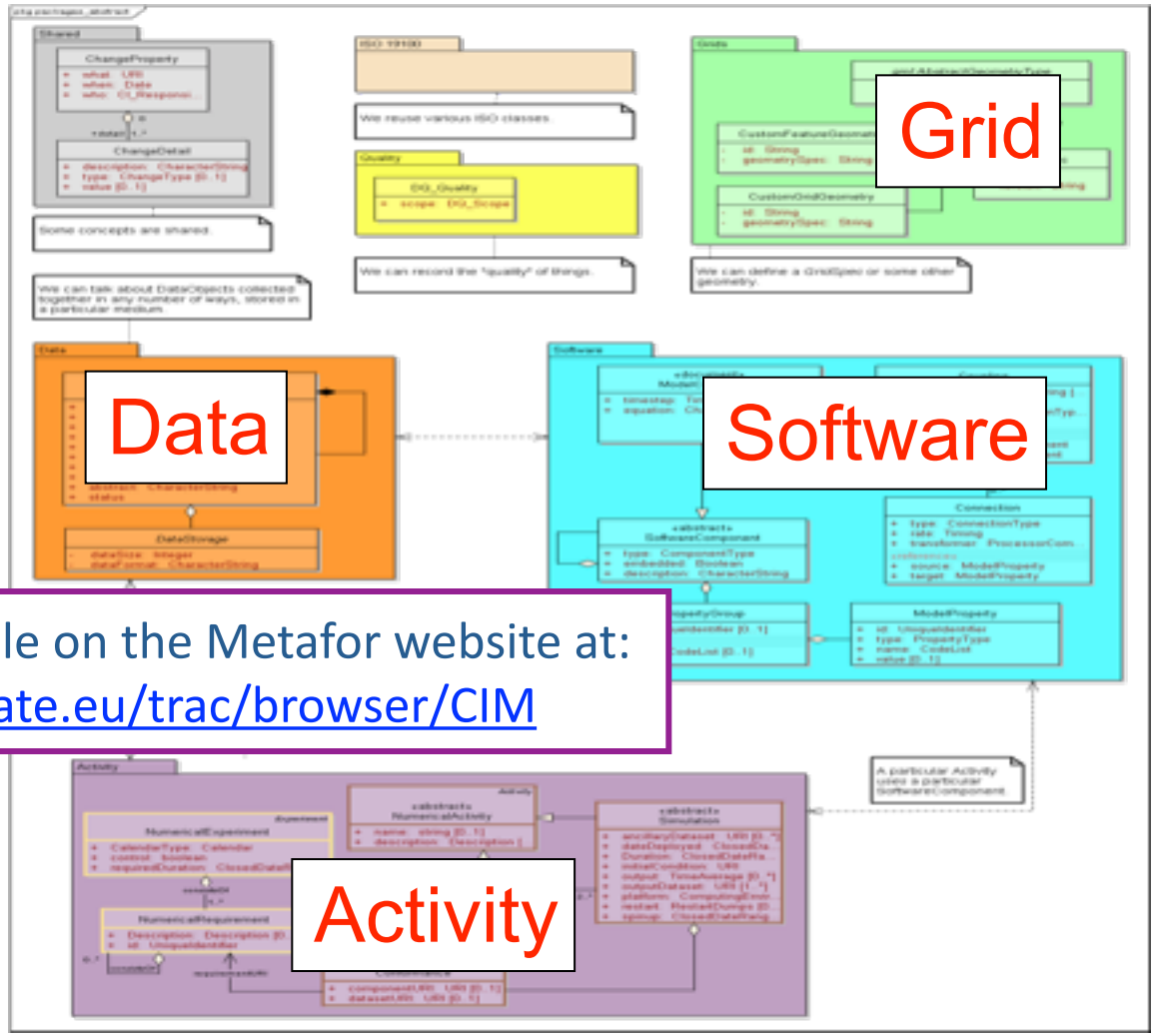
The CONCIM

Climate Modelling = an activity using a software to produce data on a grid to be archived in a repository.



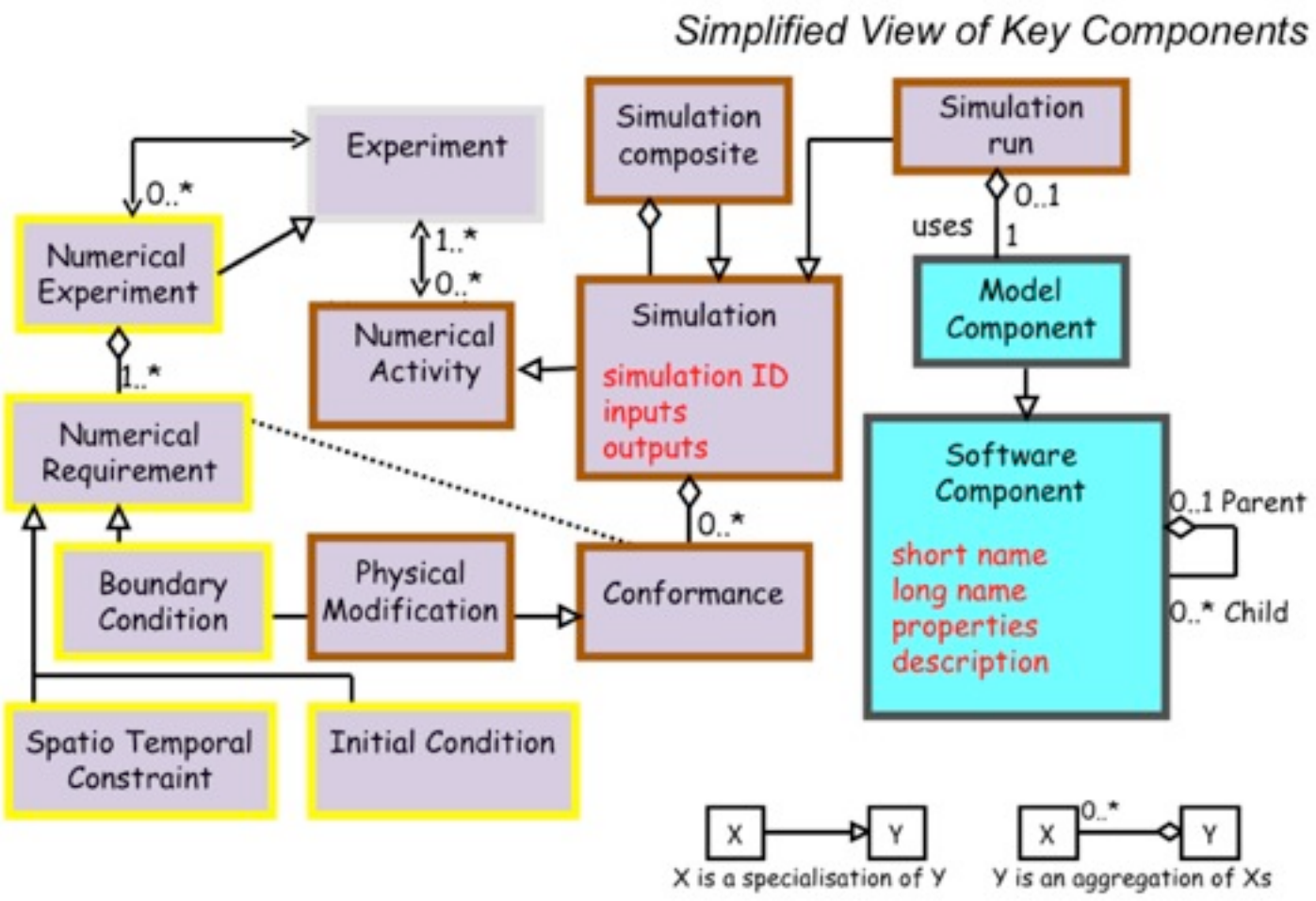
The CONCIM

Climate Modelling = an activity using a software to produce data on a grid to be archived in a repository.



CIM v 1.5 (UML) available on the Metafor website at:
http://metaforclimate.eu/trac/browser/CIM

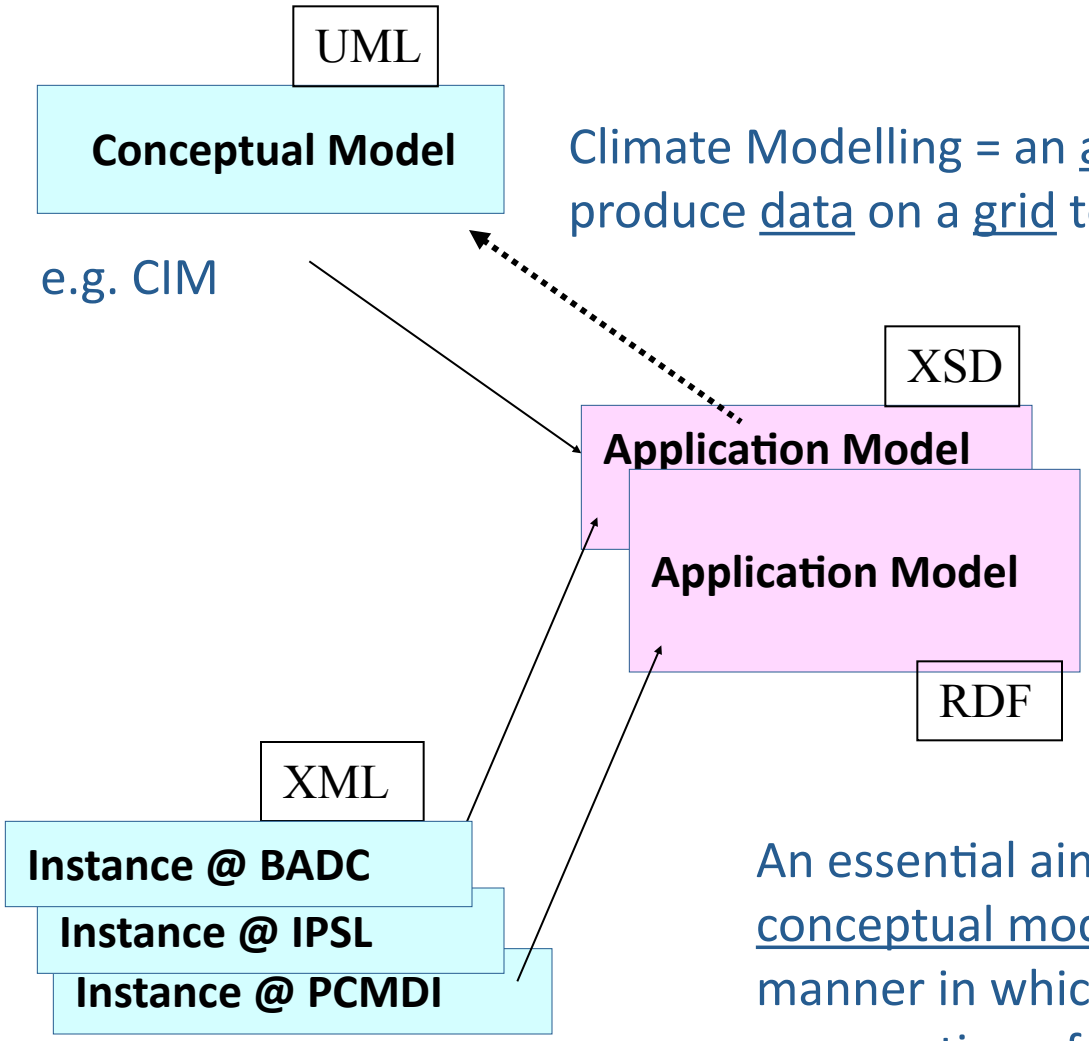
The CONCIM



metafor

National Centre for Atmospheric Science

The CONCIM to APPCIM

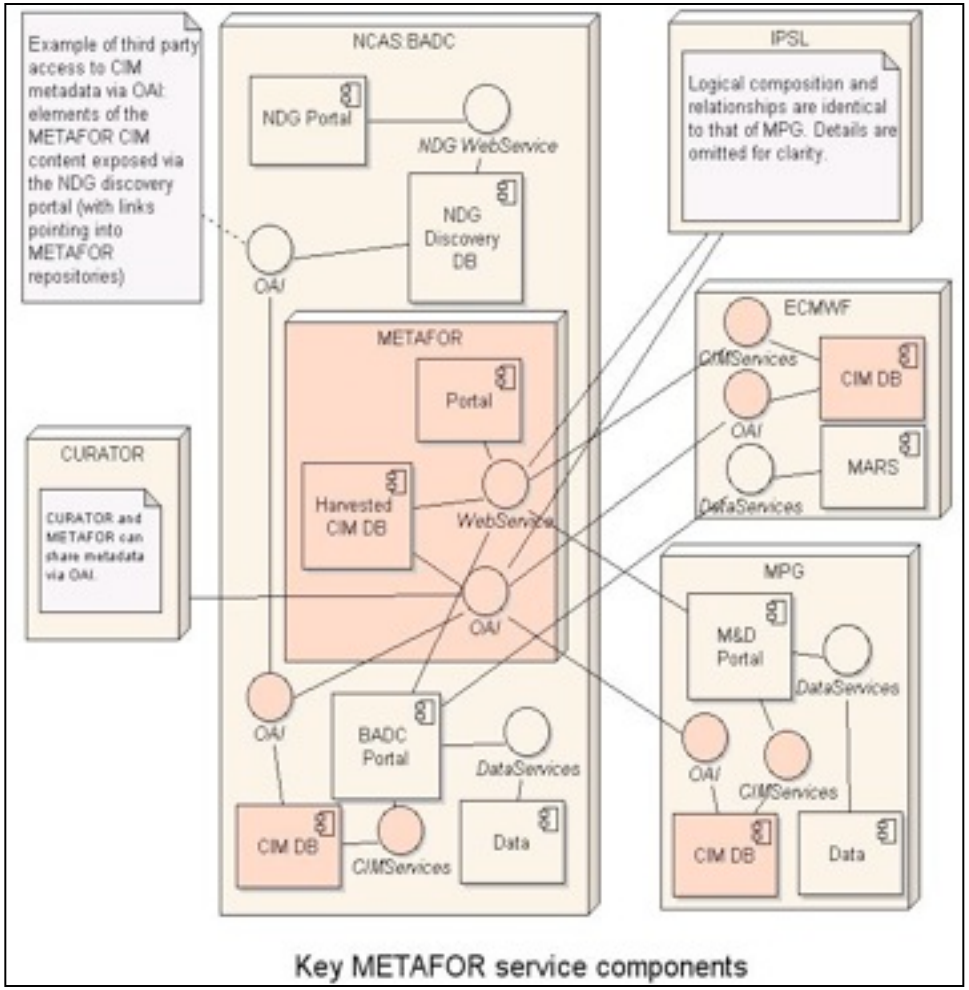


Climate Modelling = an activity using a software to produce data on a grid to be archived in a repository.

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

The Query and other CIM Tools

- **search** interface for CIM instances
- CIM instance **viewer**
- CIM instance **comparer**



The CMIP5 Model & Experiment Documentation

- A year into the project, METAFOR became "*a major international focal point for earth system modelling metadata definition*" (Karl Taylor, PCMDI)

- Metafor was tasked by WGCM/CMIP to define, collect and provide the CMIP5 model metadata

- This is when the real life stuff started !

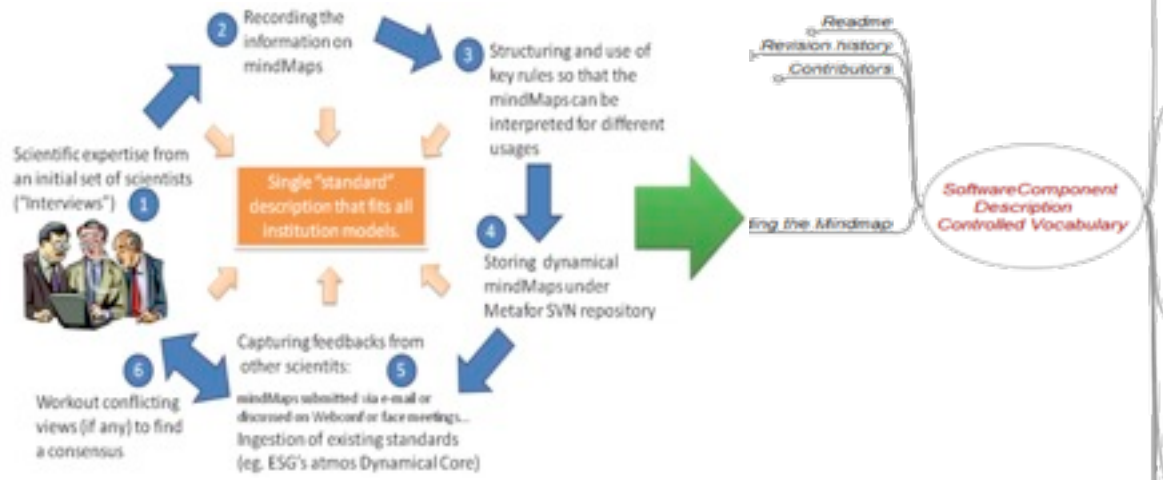
The CMIP5 Model & Experiment Documentation

1. Community undertaking (EU Metafor, US Curator,..)
2. Get the concepts right, organise separation of concerns:
   - The Common Information Model (CIM)
   - The application schema: e.g. CMIP5
3. Define controlled vocabulary (CV):
   - What matters to understand model differences ?
   - Interview experts (>35) to define CV ("mind maps")
   - Get experiments CV + experimental protocol from CMIP

Controlled Vocabulary

Novelty in community: Software CV

Creating the CV mind map



Defined a way of describing (in a community consistent way) the scientific properties of model subcomponents

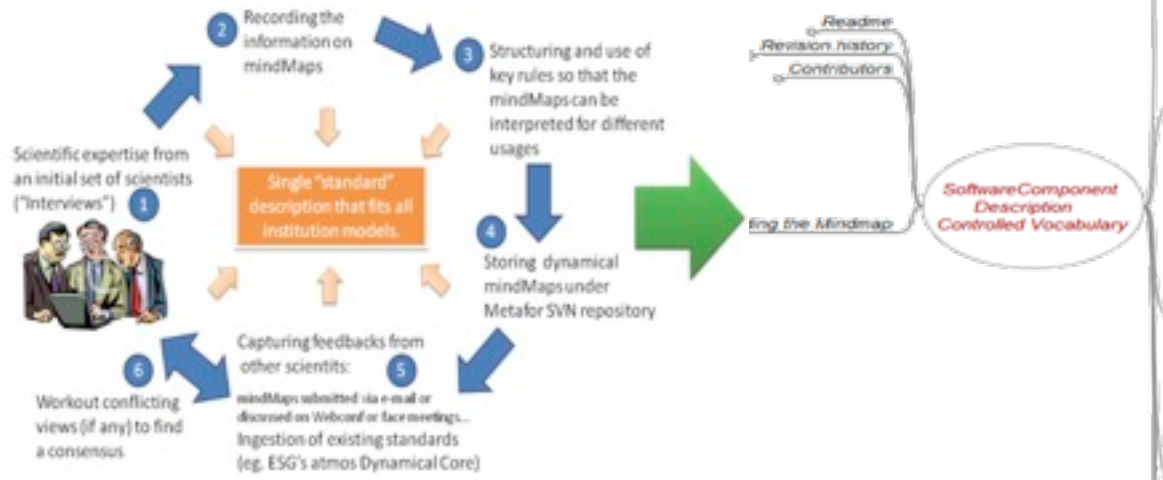
metafor

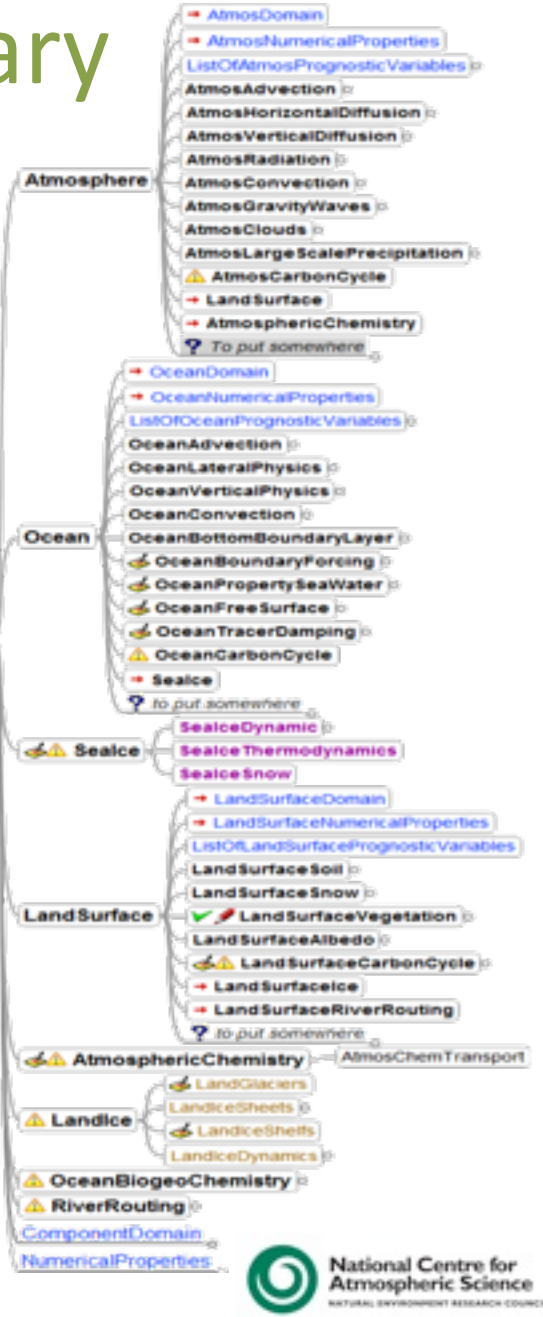
National Centre for Atmospheric Science

Controlled Vocabulary

Novelty in community: Software CV

Creating the CV mind map





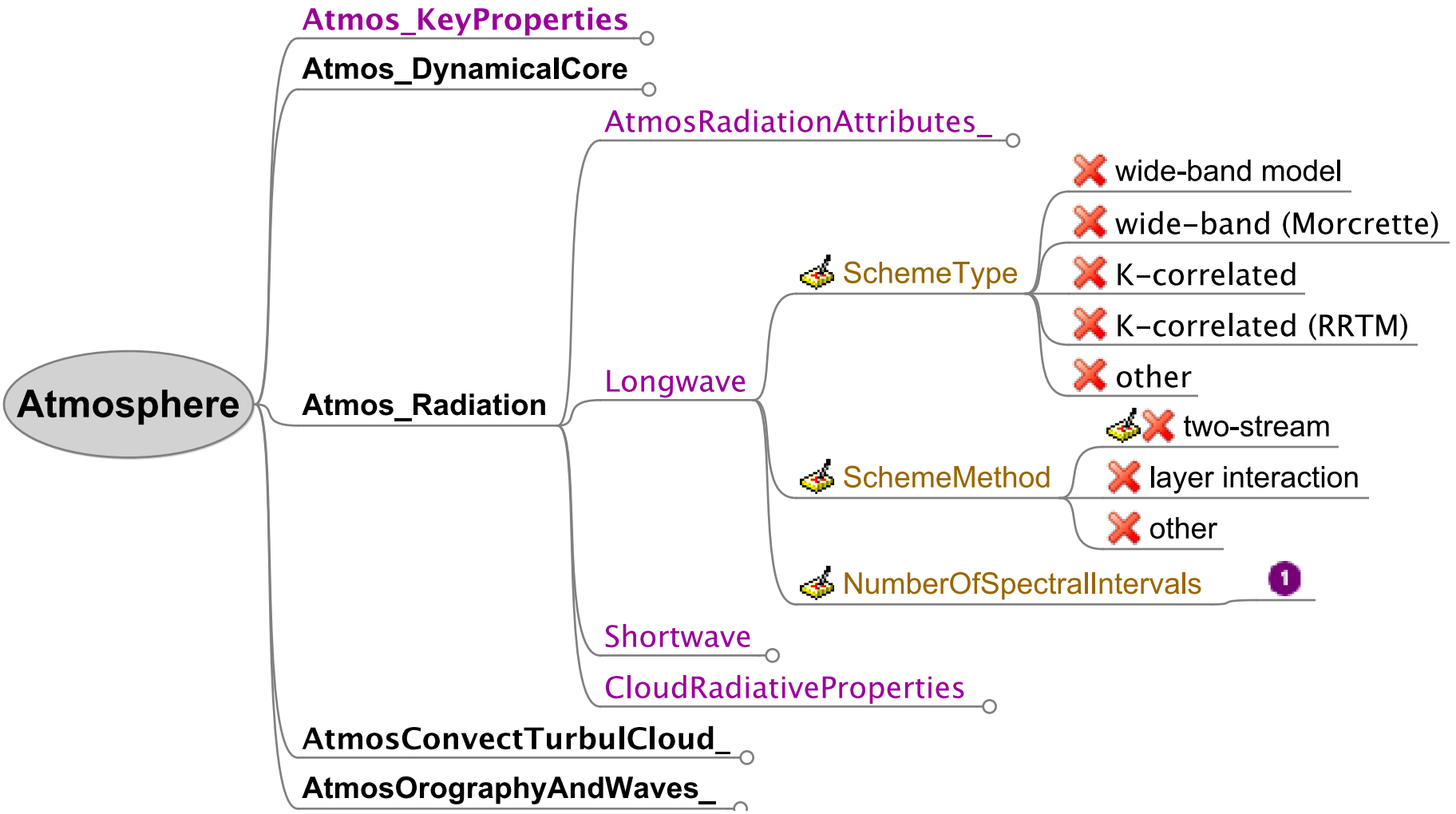
Defined a way of describing (in a community consistent way) the scientific properties of model subcomponents

570 controlled questions - hundred of choices

metafor

National Centre for Atmospheric Science

Atmosphere radiation CV

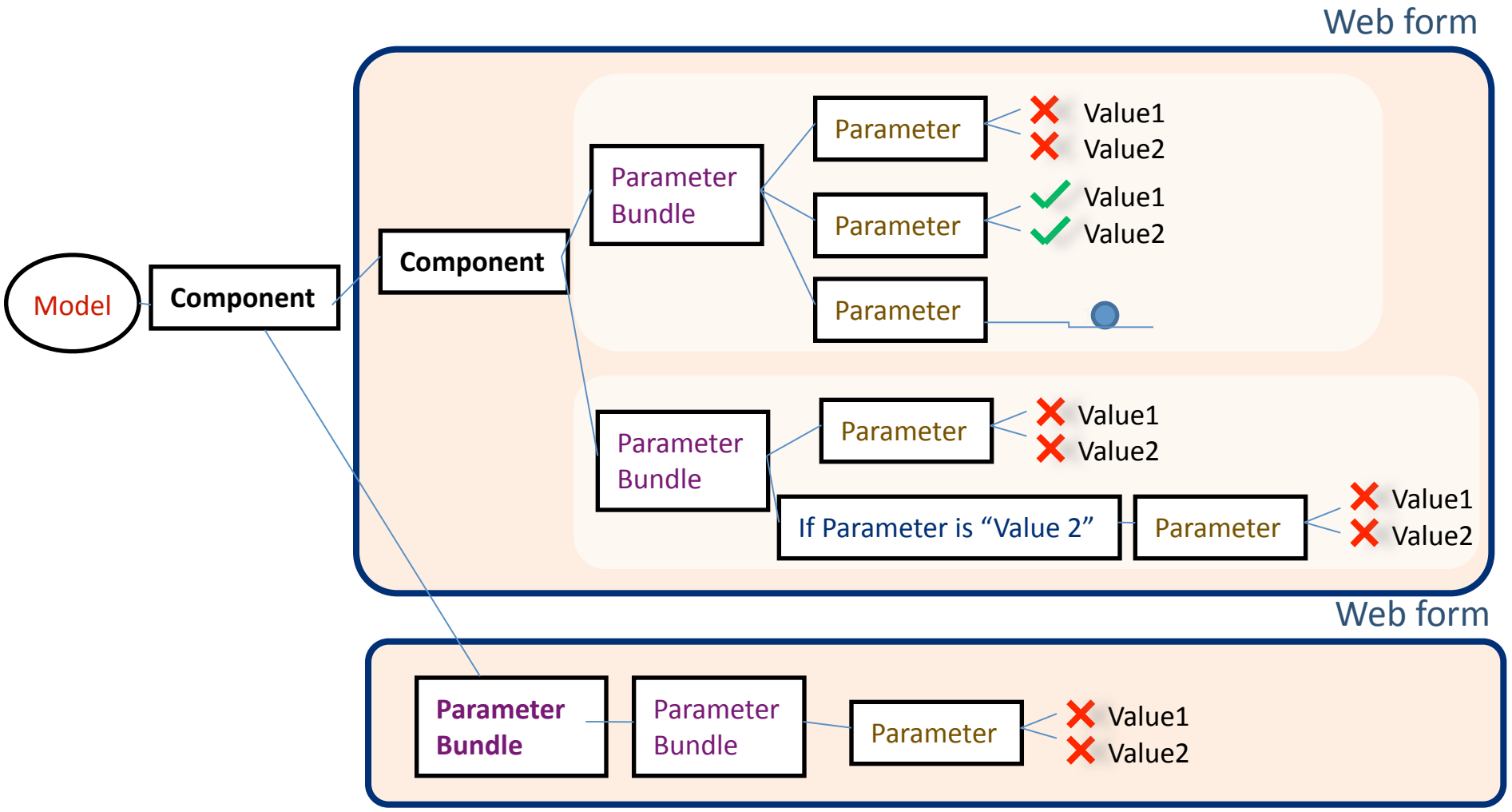


The CMIP5 Model & Experiment Documentation

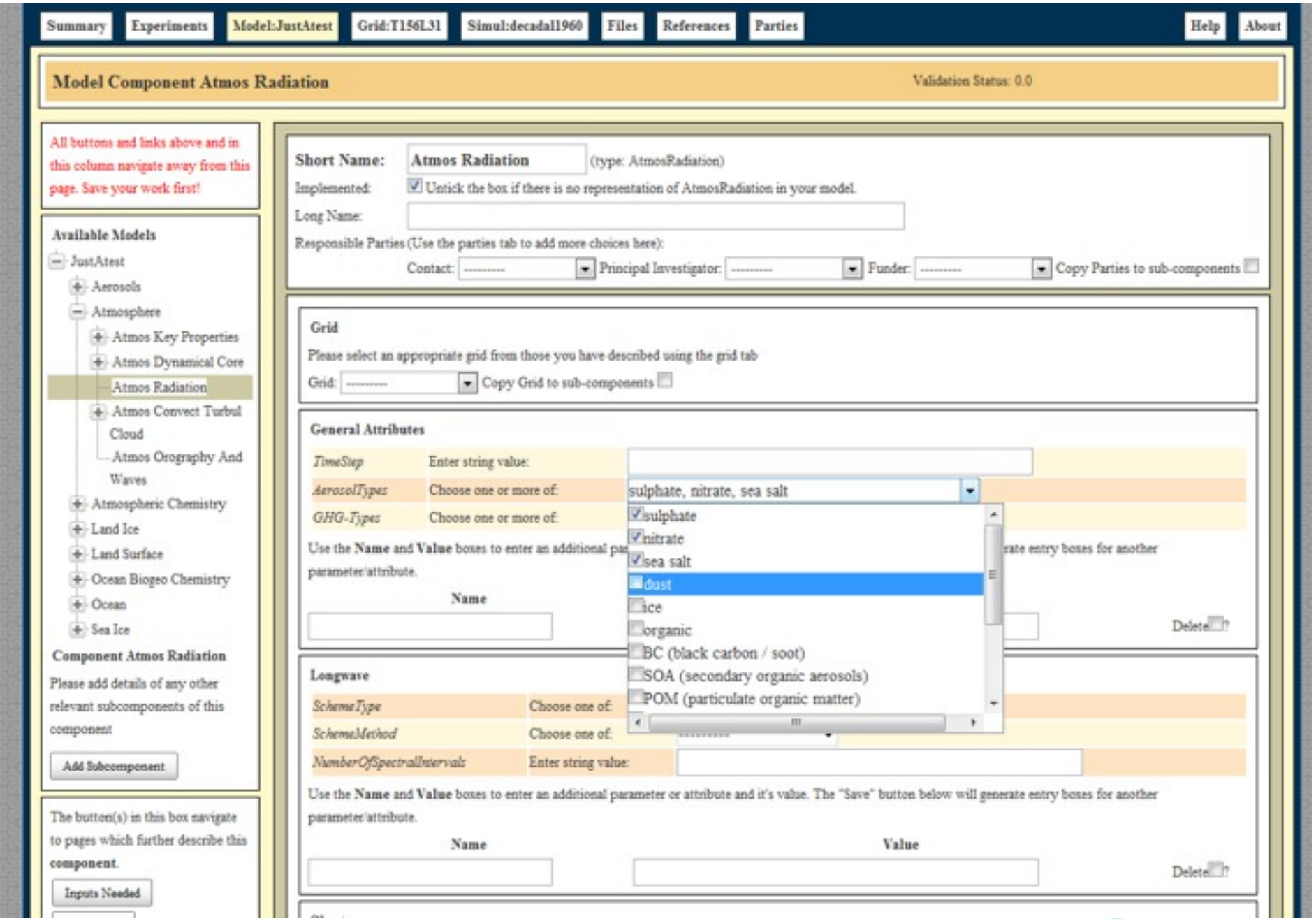
1. Community undertaking (EU Metafor, US Curator,..)
2. Get the concepts right, organise separation of concerns:
   - The Common Information Model (CIM)
   - The application schema: e.g. CMIP5
3. Define controlled vocabulary (CV):
   - What matters to understand model differences ?
   - Interview experts (>35) to define CV ("mind maps")
   - Get experiments CV from CMIP
4. Build a metadata entry tool:
   - aka the "CMIP5 questionnaire"

Mind Map to Questionnaire

Processing mind maps for a questionnaire



CMIP5 metadata entry tool (aka questionnaire)



The CMIP5 Model & Experiment Documentation

1. Community undertaking (EU Metafor, US Curator,..)
2. Get the concepts right, organise separation of concerns:
   - The Common Information Model (CIM)
   - The application schema: e.g. CMIP5
3. Define controlled vocabulary (CV):
   - What matters to understand model differences ?
   - Interview experts (>35) to define CV ("mind maps")
   - Get experiments CV from CMIP
4. Build a metadata entry system:
   - aka the "CMIP5 questionnaire"
   - validation (QC), display, distribution tools

The CMIP5 Model & Experiment Documentation

1. Community undertaking (EU Metafor, US Curator,..)
2. Get the concepts right, organise separation of concerns:
   - The Common Information Model (CIM)
   - The application schema: e.g. CMIP5
3. Define controlled vocabulary (CV):
   - What matters to understand model differences ?
   - Interview experts (>35) to define CV ("mind maps")
   - Get experiments CV from CMIP
4. Build a metadata entry system:
   - aka the "CMIP5 questionnaire"
   - validation (QC), display, distribution tools

The result is complex...

so is climate modelling !

Bringing it all together for CMIP5

The players:

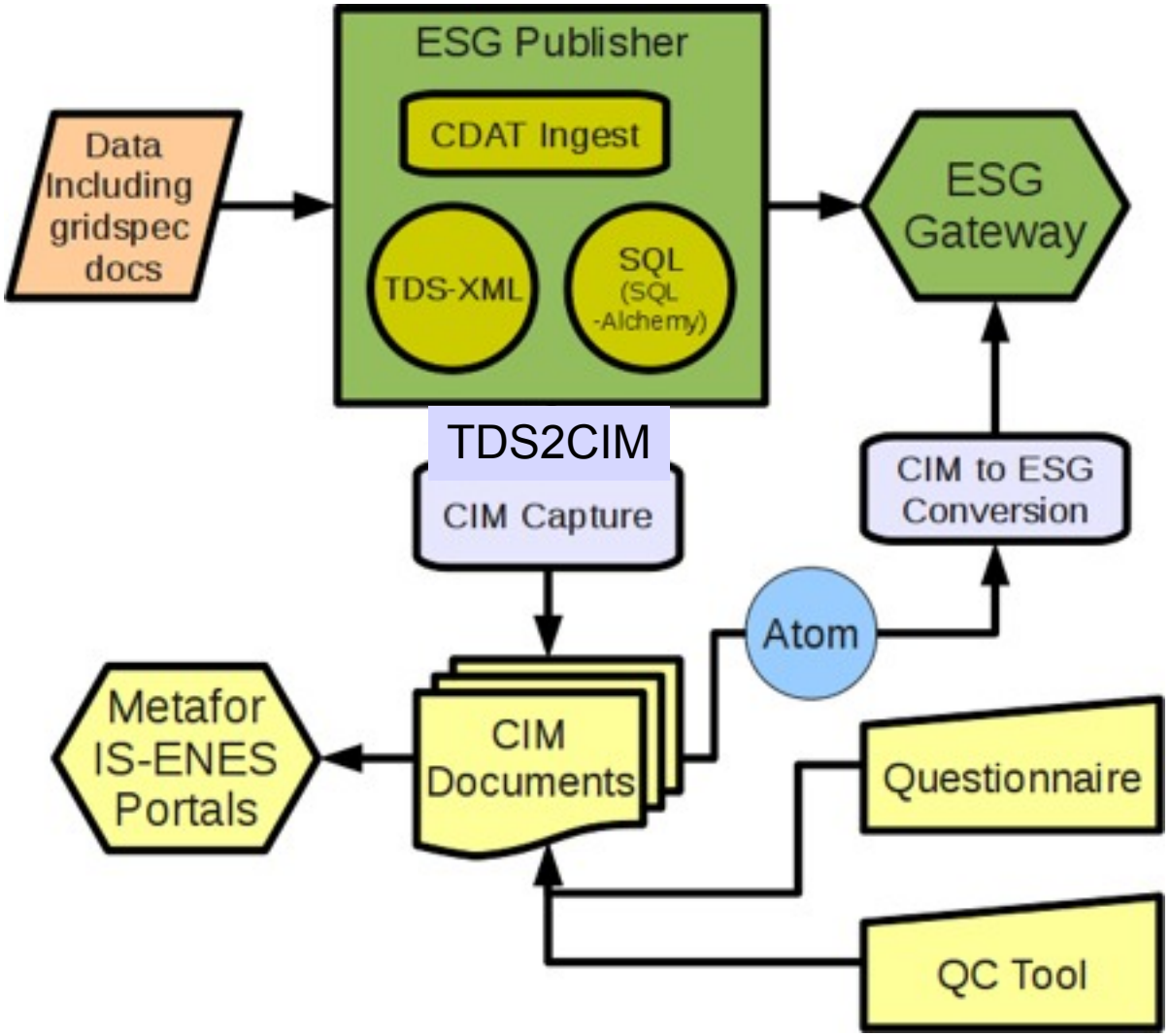
1) NetCDF
   CF Conventions +
   CMIP5 extensions
   (orange)

2) Earth System Grid +
   Earth System Curator
   (green)

3) Metafor
   (yellow)

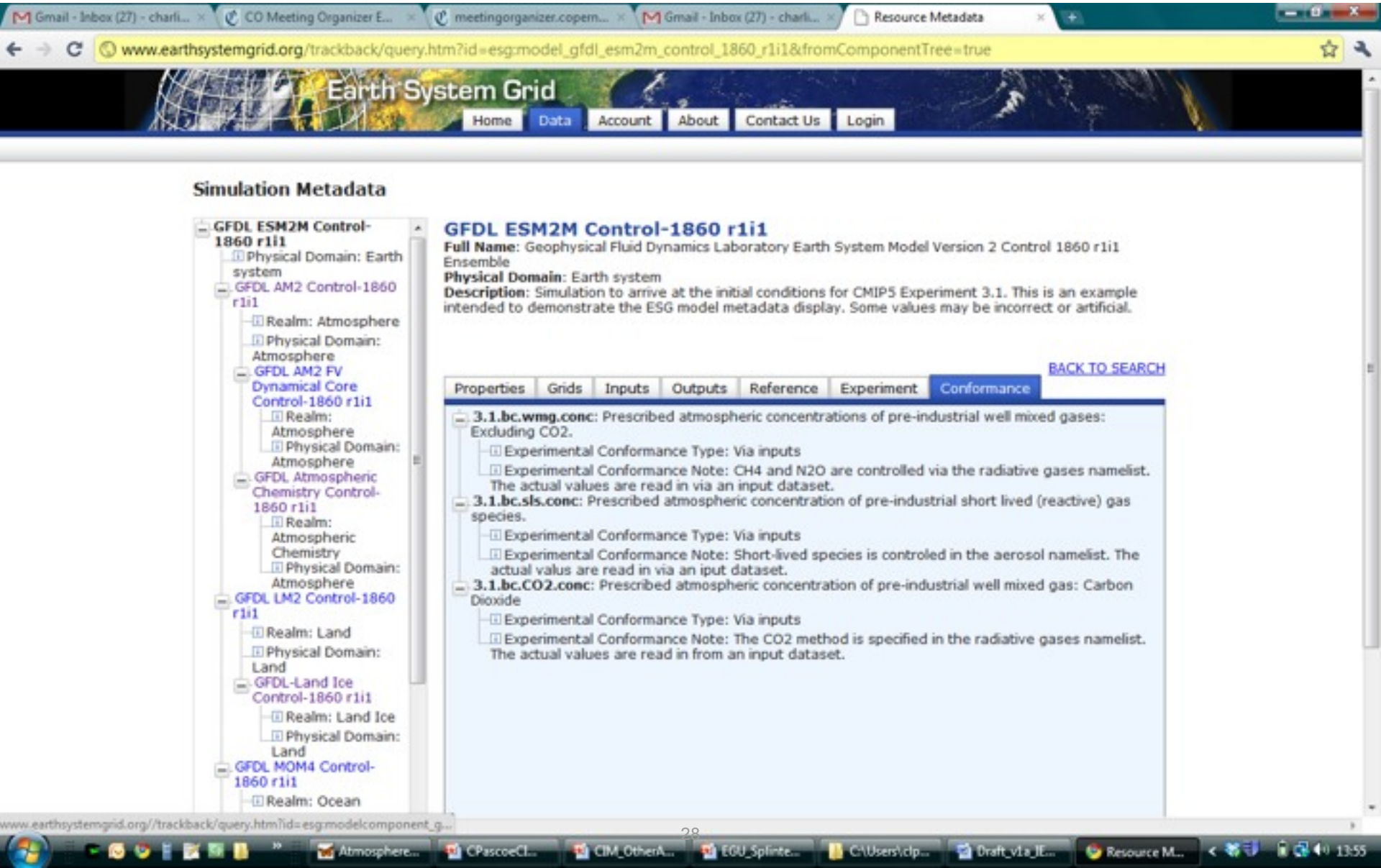
Conversion code
(light blue/grey)

Glue: Atom



ESG display of CIM instances

http://www.earthsystemgrid.org/



We've come a long way !

"The main objective of Metafor has been 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"

Requirements for success:

- Clearly define scope (separation of concerns)
- Gather top field experts (IT+climate, 124 telcos !)
- Engage with similar existing activities (US Curator)
- Work towards community adoption (CMIP5)
- Capture wider community needs
- Ensure post-project governance

Next steps

- CMIP5 questionnaire support (BADC, DKRZ)
- Provide first CIM tools and services:
  - hand over to IS-ENES
  - ecosystem of tools to flourish
- CIM use beyond CMIP5:
  - statistical downscaling CV
  - CMIP5 metrics
  - library to ingrain CIM generation within GCMs
- Community governance for CIM and CV:
  - "Standards Committee" under WCRP/CMIP

More from

http://metaforclimate.eu

http://cmip-pcmdi.llnl.gov/cmip5/modeling_getting_started.html

Guilyardi E., V. Balaji, S. Callaghan, C. DeLuca, S. Denvil, R. Ford, M. Lautenschlager, B. Lawrence, L. Steenman-Clark, S. Valcke and the METAFOR group (2011) **The CMIP5 model and simulation documentation: a new standard for climate modelling metadata**, *CLIVAR Newsletter, submitted*

Lawrence, B.N., V. Balaji, P. Bentley, S. Callaghan, C. DeLuca, S. Denvil, G. Devine, M. Elkington, R. Ford, E. Guilyardi, M. Lautenschlager, M. Morgan, M.-P. Moine, S. Murphy, C. Pascoe, H. Ramthun, P. Slavin, L. Stenman-Clark, F. Toussaint, A. Treshansky and S. Valcke (2011). **Describing Earth System Simulations**, *IEEE, submitted*.

Moine, M.P., C. Pascoe, A. Alias, V. Balaji, P. Bentley, G. Devine, R. Ford, E. Guilyardi, B. N. Lawrence, S. Valcke (2011) **Development and Exploitation of a Controlled Vocabulary in support of Climate Modelling**, *IEEE, submitted*.

The CMIP5 questionnaire costs

1. For a modelling centre:
   - Several weeks of organising and identifying in-house and external experts to answer the 570 questions and describe conformance of simulations to CMIP5 protocol
   - Few days to fill up the online entry tool ("questionnaire")
   - Frustration and sweat (but 3 groups already made it !)
2. For the Metafor and Curator projects:
   - More than 350 man.months of effort over 3 years
   - Building on many community initiatives and existing standards
3. For BADC, WDCC and ESG:
   - Commitment for operation and maintenance of metadata service systems, including entry tool, quality insurance, search and display tools, ...

The CMIP5 questionnaire benefits

1. For a modelling centre:
   - Identification, location and use of climate model data (now and in 10 years) beyond climate scientists (“climate models google”)
   - Showcase for in-house climate models (public description)
   - Unique and comprehensive container for model description
   - New standards for internal development (e.g. CF/netCDF)
2. For a CMIP5 analyst:
   - Locate and make sense of data - assess suitability for own purpose (“climate models google”)
3. For the community:
   - a new open standard, alongside CF and netCDF
   - Evidence of maturity, credibility and openness of our science
   - Ease and strengthen model assessment in IPCC AR5