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