METAFOR End-User Testing Report
METAFOR Deliverable 3.4 M42

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	D3.4 Month 42
Deliverable Title	End-User Testing Report
Document Identifier	METAFOR-D3.4_M42
Date	September 28, 2011
Work Package	WP3 Service Content Creation & Testing
Authors	IPSL
Document Status	Draft
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	

Document History			
Version	Date	Comment	Author/Partner
0.1	September 29, 2011	Final Draft	S. Denvil/IPSL

Abstract:

There is growing interest among researchers, policy-makers, businesses, and the general public in the potential impacts of climate change and to what extent undesirable consequences of climate change can be mitigated by reducing anthropogenic greenhouse gas (GHG) and specifically carbon emissions. This observation being made we can foresee that the Common Information Model (CIM) End Users represent a broad community. This deliverable will summarize the various points of entry available to access and use CIM information and will highlight the new potentialities offered.

Introduction

The Metafor Common Information Model, i.e. the CIM, is an ontology designed to become the ipso-facto standard for climate modeling related metadata. The CIM eco-system allows institutes to integrate the CIM into their day to day climate modeling processes. It achieves this by supporting various requirements: the ontology itself; validation; search; dissemination; integration (with other metadata platforms such as Earth System Grid).

The preferred access to date to the CIM repository and the underlying information by users includes:

1) **CMIP5 METAFOR QUESTIONNAIRE** (producing CIM instances)

The CMIP5 questionnaire is web a based interactive tool for creating CIM documents about the climate models that contribute to CMIP5, the CMIP5 Questionnaire also captures information about how those climate models were set up to run simulations for CMIP5 experiments. METAFOR has been charged by the WGCM[1] via the CMIP[2] panel to define and collect model and experiment metadata for CMIP5 to be used for the next IPCC assessment (due in 2013). To do this the METAFOR team have developed a web-based questionnaire to collect information from the CMIP5 modelling groups on the details of the models used, how the simulations were carried out, how the models conformed to the CMIP5 experiment requirements and what hardware was used to perform the simulations. The aim is to document models in sufficient detail so that the CMIP5 data can be compared in a scientifically meaningful way. The climate model data will be stored in the CMIP5 archives hosted at PCMDI, BADC and MPI-M. These archives are anticipated to be greater than 1 PB in size, so good metadata is crucial to their effective use and operation.

Help systems and documentation are being developed to support the users of the questionnaire. These include FAQs on the METAFOR site, a dedicated email address solely for questionnaire issues and webcasts and interactive web seminars to publicise and train users of the questionnaire. The development paradigm for the CMIP5 questionnaire required that a working version be available at all times. The questionnaire has evolved through numerous iterations as bugs have been found and fixed and the functionality has been extended and improved following feedback. This paradigm has been really appreciated from the user perspective.

As an example of functionality extension due to user feedback a new addition to the CMIP5 questionnaire has been the introduction of the 'ensemble member' information page. This has been introduced to aid users who need access to this type of information upfront, as opposed to waiting for published CIM documents. This new page, accessible through the front page of the questionnaire, gives an overview of ensemble member information for each simulation that a chosen centre has documented. In particular, it will give a breakdown of the rip values (DRS) for each ensemble member and what this rip value refers to in terms of simulation modification, e.g. an input modification or parameter

[1] http://www.wmo.ch/pages/prog/wcrp/AP_Modelling_WGCM.html
[2] http://www.clivar.org/organization/wgcm/cmip.php

change. An included dropdown help panel gives more detailed guidelines about the information shown.

2) **METAFOR CIM WEB SERVICES** backend (ingesting CIM instances and exposing it to users as resources endpoint, resources being the unit of exchange between a service & client.)

The Metafor web services have been implemented according to the Service Orientated Architecture (SOA) paradigm. SOA is a key tenet in the design of secure, robust & distributed systems. Such architecture deconstructs a system into a set of discrete functional units known as services. Services are said to be supplied by providers (e.g. Metafor) and consumed by clients (i.e. tools, applications, and in some cases other services).

A web service is such a discrete functional unit deployed upon a web server and thus consumable via the HTTP protocol. Web services leverage inherent HTTP features such as security & caching. By decomposing a system into web-services, rich & diverse informational eco-systems can be incubated, nurtured & supported. Web services can be implemented in several different styles, but an architectural style known as REST (REpresentational State Transfer) is particularity well suited to the HTTP protocol. REST services posit resources as the unit of exchange between a service & client. A resource is an informational unit supporting at least one representation (i.e. an encoding such as XML)

In order to deconstruct a system into a set of discrete REST services, one identifies the collection of resources flowing through the system. Each type of resource is then managed by a distinct REST service implementation.

These web services are RESTful in style and support a diverse array of CIM resources: documents; validation reports; queries, id's, controlled vocabularies. Collectively CIM web services constitute a platform upon which CIM clients such as web-portals or a standalone tools can be developed. At the heart of a CIM client is the consumption of one or more CIM web services, at the heart of a CIM web-service is a CIM resource.

3) **METAFOR PORTAL** (rely on METAFOR CIM web services backend)

Organisations (research institutes, companies … etc) are encouraged to integrate the CIM into their day to day climate modelling processes. This integration process will be achieved by providing a supporting infrastructure: the ontology itself; validation tools; search engine; outreach programs; integration (with other metadata platforms such as Earth System Grid) … etc.

One important aspect of this eco-system is the CIM Web Portal. The portal is a website designed to deliver a diverse set of use cases to a diverse set of actors. The set of supported use cases can be placed into several categories: outreach; technical support; search; tools; publishing. The actors can be placed along a continuum of meta-data expertise ranging from members of the public to meta-data experts.

Satisfying the array of use cases in a user friendly fashion is a key objective of the CIM Web Portal. To achieve this the portal is developed as a modern rich internet application (RIA). RIA's typically render rich user interfaces that asynchronously consume data provided by back-end web services, i.e. the CIM web services. Thus the CIM Web Services (see EU deliverable D5.5) act as the source of all information rendered in the CIM Web Portal.

4) **EARTH SYSTEM GRID (ESG) GATEWAY PORTAL** (integration of METAFOR CIM content from the CMIP5 questionnaire into existing portal. Could ingest content from METAFOR web services as well)

Model metadata and related pages in the ESG-CET gateway has been collaboratively developed between CURATOR and METAFOR team. Model metadata display that is part of ESG release 1.3.1 includes:

- Metadata versioning
- Model metadata to data connections
- Component tree and navigation
- Inputs/couplings
- Grids
- Model modifications
- Ensemble information

This has been a community vetted development; based on regular live demo and regular email feedbacks including domain specialists' representative. This has been the result of an international network of experts forming a recognized focal point for metadata development in the climate modelling community, including close collaboration with US led Curator project.

For the 1.3.1 release, an ESG gateway can download files from the Metafor Atom feed and harvest and display the results in the ESG gateway trackback. Items that can currently be harvest and display include basic properties, technical properties, conformances, platforms, citations, grids, inputs, experiments, and scientific properties.

5) **UNIVERSITY OF CANTABRIA (UoC) DOWNSCALING PORTAL** (integration of METAFOR web services into existing portal).

The University of Cantabria downscaling portal (UoCDP, https://www.meteo.unican.es/downscaling) allows users from the climate impact community to calibrate/downscale the model outputs in the region of interest using historical observed records. The portal includes public observation datasets (e.g. GSOD) and allows uploading new historical data (including private datasets, not available for other users). One of the major drawbacks from the user perspective is discover the climate model and simulations been performed. For this reason the UoCDP has integrated the different services developed by WP4. This integration allows the portal user to access model components and simulations metadata provided by modelers through.

As a first approach the linkage has been made with some CMIP5 CIM instances which corresponds to climate models simulations currently available in UoCDP. Because those model has been ingested into the downscaling portal from the DKRZ CERA database, all

the effort made with TDS2CIM tool uploading dataObjects, can be accessed directly from the portal, which constitutes valuable information about spatio-temporal coverage and climate fields outputs from models available in the CERA database. For example, the UoCDP is based on daily and sub-daily at surface and upper-level fields, therefore what is important for the user, is to know which datasets are available with these requirements.

This development constitutes a first prototype, providing a huge potential of improvement of the metadata information available and the corresponding services. Life beyond Metafor will see integration of the CIM with IS-ENES and ESGF initiatives, this will provide to the UoCDP almost immediately an access to those metadata referring model simulations, which will be available to impact users from the UoCDP. This will create end-to-end linkage from modelers to impact users, where historically a gap exists.

Actors

The CIM eco-system must be able to support a wide array of actors; these actors can be sorted by their relative expertise with climate related meta-data. These actors will have different preferential relations with CIM repository access point.

Member of Public

A member of public arrives at the METAFOR portal as a result of a news item they may have seen/read. Is curious to gain a deeper understanding of the rationale behind the Metafor project. It's very likely that he won't access CIM information from the other access point as they are now (specialized access point). Attractor: for each controlled vocabulary element make a link to "Wikipedia" type of content.

Policy Maker

A policy maker arrives at the METAFOR portal as a result of a generalized outreach program. He is curious to gain a deeper understanding of the rationale behind the Metafor project. He wants to integrate information held on the portal into their policy making process. May wish to experiment with simple search functions. It's likely that he won't access CIM information from the other access point as they stand now (specialized access point).

Attractor : provide high level statistics of the CIM repository content (coupled model per continent; total number of different PI; statistics of computer architecture used by climate community, by downscaling community,…; number of model per class of mesh spatial resolution (100 km and above, between 50 km and 100 km, ….)).

Scientific User

A scientific user arrives at the METAFOR portal as a result of either a scientific collaboration or a Metafor publication/presentation. He is curious to gain a deeper understanding of the rationale behind the Metafor project. He wants to understand some of the science encoded in the CIM ontology. It's likely that he will access CIM information from METAFOR portal, ESG gateway portal and/or UoC portal if he is using climate data during her scientific activity.

Attractor: build domain specific controlled vocabulary so as to be able to build specialized scientific domain application taking advantage of the CIM eco-system. Promote METAFOR CIM web services integration into scientific portal when applicable.

Climate Scientist

Climate scientist arrives at the METAFOR portal as a result of a focused community outreach program. He requires a very deep understanding of the rationale behind the Metafor project. He will experiment with the search engine and may provide input into portal evolution. He will access CIM information from the METAFOR portal, the ESG gateway portal and/or the CMIP5 questionnaire. He will build software development plan and scientific use cases so as to take advantage of METAFOR CIM web services (populating new CIM contents and querying it).

Attractor: build networking bringing together Climate Scientists and CIM experts so as to ensure that "new science" will be capture by the CIM-ecosystem

Climate Modeler

Climate modeler arrives at the METAFOR portal as a result of a focused community outreach program. He requires a very deep understanding of the rationale behind the Metafor project and the CIM. He will leverage the search engine in her daily working life. He needs to be notified of changes in the portal and may provide input into the portal evolution. He will access CIM information from the METAFOR portal, the ESG gateway portal and/or the CMIP5 questionnaire. He will build software development plan and scientific use cases so as to take advantage of METAFOR CIM web services (populating new CIM contents and querying it).

Attractor: build networking bringing together Climate Scientists and CIM experts so as to ensure that "new science" will be capture by the CIM-ecosystem. Integrate part of the CIM-ecosystem in the climate modeling software framework so as to generate CIM instances during models runtime.

Data Analyst

Data analyst arrives at the METAFOR portal as a result of a focused community outreach program. He requires a deep understanding of the rationale behind the Metafor project and the CIM. He will leverage the search engine in her daily working life. He needs to be notified of changes in the portal and will provide input into the portal evolution. He will access CIM information from the METAFOR portal, the ESG gateway portal and/or the CMIP5 questionnaire. He is one of the CIM content heavy users. He will provide feedbacks so as to enhance the way scientific portals expose CIM information.

Attractor: quality and ergonomics of scientific portals will be key attractor factor. Also data analyst can be interested in many scientific digital repositories. So the existence of a client side library able to query the METAFOR CIM back end web services easily is key to success.

Data Expert

Data expert arrives at the portal as a result of a focused community outreach program. He requires a deep understanding of the rationale behind the Metafor project. He will experiment with the search engine and will imagine what could be done with respect to the kind of data he is using. He will try to find out if her data can be described by the CIM with the same level of quality as what he is doing now. He wants a clear picture so as to decide whether the CIM-ecosystem could be interoperable with other major scientific digital repository. He wants to know if the METAFOR CIM-ecosystem is a sustainable approach for his case.

Attractor: Build a strong outreach program together with a strong documentation strategy. The CIM-ecosystem needs to approach selected domain specific data expert so as to engage domain specific expansion/development.

Meta-Data Expert

He arrives at the portal as a result of a focused community outreach program. He requires a very deep understanding of the rationale behind the Metafor project and the CIM. He will leverage the search engine in her daily working life. He needs to be notified of changes in the portal. He will provide input into the portal evolution. He will work on CIM-ecosystem integration and/or development. He can work on conceptual CIM and on application CIM. He needs a documented CIM-ecosystem.

Attractor: Build a strong outreach program together with a strong documentation strategy. CIM-ecosystem components being base on several standards they must be precisely describe and should evolve with time. Produced code will have essentially to search and manipulate CIM instances. Thus for each language (Javascript, Python & Java) having a serialization and validation component with native objects representing the CIM would simplified downstream developers work and thus the community could be more creative.

Data/Meta-Data Administrator

He arrives at the portal as a result of a focused community outreach program. He needs to be certain that meta-data generated by their organization appears in and is accessible from the search engine results. He may leverage the search engine in her daily working life. He will wish to be notified of changes in the portal and he will provide input into the portal evolution.

Conclusions

Due to the complexity of the climate system and the variety of experimental design a climate modeling software can be configured to conform to; it is crucial to offer to end users a comprehensive representation of this information. With respect to this point description of the experimental design and description of the available data are critical point. Future CIM governance process will be key to permit new end-to-end linkage from modelers to users, where historically a gap exists.

Furthermore METAFOR provides essential pieces and framework to support parties building climate information system. It is also essential to support this activity by providing clear documentation and ease the life of downstream developers to encourage integration of CIM-ecosystem into third parties software infrastructure.

Bibliography

[1] ESG publisher scripts: http://www2-pcmdi.llnl.gov/Members/bdrach/.personal/esg-publication-scripts/

[2] Metafor CMIP5 questionnaire: http://q.cmip5.ceda.ac.uk/cmip5/

[3] CERA data model: http://www.mad.zmaw.de/wdc-for-climate/cera-data-model/

[4] World Data Center for Climate (WDCC): http://www.mad.zmaw.de/wdc-for-climate/

[5] CERA database: http://www.mad.zmaw.de/wdc-for-climate/cera-database/

[6] CERA metadata with internal experiment id: http://anticyclone.dkrz.de:8080/xsql/cera_map_cim.xsql?id=2035543

[7] CIM controlled vocabulary, CV: http://metaforclimate.eu/trac/wiki/ticket/192

[8] MindMap CV capturing: http://metaforclimate.eu/trac/wiki/tickets/245

[9] Linux MindMap tool (freemind): http://freemind.sourceforge.net/wiki/index.php/Main_Page

[10] CF (Climate and Forecast) description: http://cf-pcmdi.llnl.gov/

[11] Metafor CIM tools and services: http://metaforclimate.eu/trac/browser/cim_software/trunk/docs/design/METAFOR_D5_5.odt

[12] THREDDS: http://www.unidata.ucar.edu/projects/THREDDS/

[13] DRS: http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf

[14] CMIP5 questionnaire roadmap: http://metaforclimate.eu/trac/attachment/ticket/721/CMIP5QuestionnaireRoadmapBeta5.pdf

[15] TDS2CIM tool / service: http://metaforclimate.eu/trac/wiki/tickets/707

[16] ENSEMBLES final report: http://ensembles--u.metoffice.com/docs/Ensembles_final_report_Nov09.pdf