

PRISM Support Initiative



PRISM Core Group report to the PRISM Steering Board

*Edited by:
Sophie Valcke*

PSI Management Report
08

May 30, 2007

Contents

1	Scope of document	2
2	METAFOR (Common Metadata for Climate Modelling Digital Repositories)	2
3	Coupling & I/O PAE	4
3.1	Development of the OASIS3 coupler (CERFACS and CRAY)	4
3.2	Development of the OASIS4 coupler (NEC-CCRLE, CERFACS, CNRS, SGI, NEC HPCE)	4
4	Integration and Modelling Environments PAE	5
4.1	Met Office contribution	5
4.2	CERFACS contribution	5
4.3	M&D contribution	6
5	Data processing and Visualisation PAE	7
6	Metadata PAE	7
7	Computing PAE	8

1 Scope of document

This document gives a short overview of the activities in the PRISM Areas of Expertise (PAE) since the last report in September 2006¹.

Intense interaction took place in the PRISM Core Group² during this last year, culminating with the writing of a very strong proposal, the METAFOR project (see section 2). The many years of working together within PRISM did bear their fruits as they allowed very constructive and efficient group dynamics in writing this proposal.

In the document, the activities of the different PAEs are also described as follows:

- Coupling & I/O in section 3
- Data processing and Visualisation in section 5
- Integration and Modelling Environments in section 4
- Metadata in section 6
- Computing in section 7

2 METAFOR (Common Metadata for Climate Modelling Digital Repositories)

METAFOR was submitted May 2nd 2007 to the FP7 infrastructure call INFRA-2007-1.2.1 "Scientific Digital Repositories" (requested EU contribution: 2.2 MEuros).

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 address the fragmentation and gaps in availability of metadata (data describing data) as well as duplication of information collection and problems of identifying, accessing or using climate data that are currently found in existing repositories. 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, private sector).

METAFOR builds on strong expertise within the European Union and beyond to:

- Develop a common information model (CIM) for climate modelling data so that key data and model distinctions can be shared and understood between users of different scientific repositories;
- Use that CIM to develop, deploy, and evaluate a prototype infrastructure that will allow specific data and models to be found and compared between holdings at partner sites;
- Exploit that infrastructure to open existing content to new users communities, and to engender new opportunities for scientific collaboration, policy use and private sector;

¹http://www.prism.enes.org/Publications/Reports/workplan2006_07_b.pdf

²The PRISM Core Group is currently composed of:

- Eric Guilyardi (CNRS/LOCEAN-IPSL), co-coordinator
- Sophie Valcke (CERFACS), co-coordinator and leader of the Coupling & I/O PAE
- Reinhard Budich (MPI), leader of the PRISM User Group
- Mick Carter (UK Met Office), leader of the Integration and Modelling Environments PAE
- Michaël Lautenschlager (M&D), leader of the Data processing and Visualisation PAE
- Lois Steenman-Clark (NCAS-Climate), leader of the Metadata PAE
- Marie-Alice Foujols (CNRS/IPSL), co-leader of the Computing PAE
- René Redler (NEC-CCRLE), co-leader of the Computing PAE
- Nils Wedi (ECMWF), ECMWF representative

- Generate plans for longer term development and sustainability of such an infrastructure, with specific attention to the interactions with national, EU and other international initiatives.

To facilitate this, METAFOR will:

- Provide solutions that will scale with the demands of repositories that will soon measure in tens of peta-bytes and millions of climate simulations datasets across Europe and whose complexity is only increasing as new models are developed and additional climate feedback mechanisms are explored;
- Describe climate models as well as the data they generate. This will improve collaborative exploitation of the available resources.

In doing so, METAFOR will provide direct support for:

- The European contribution for planning for a future Fifth Assessment Report of the IPCC;
- The improved exploitation of existing activities, in particular the international projects ENSEMBLES and CCMVAL;
- Other activities, which may spin up during the project, and for which appropriate information will be identified during the use cases identification phase.

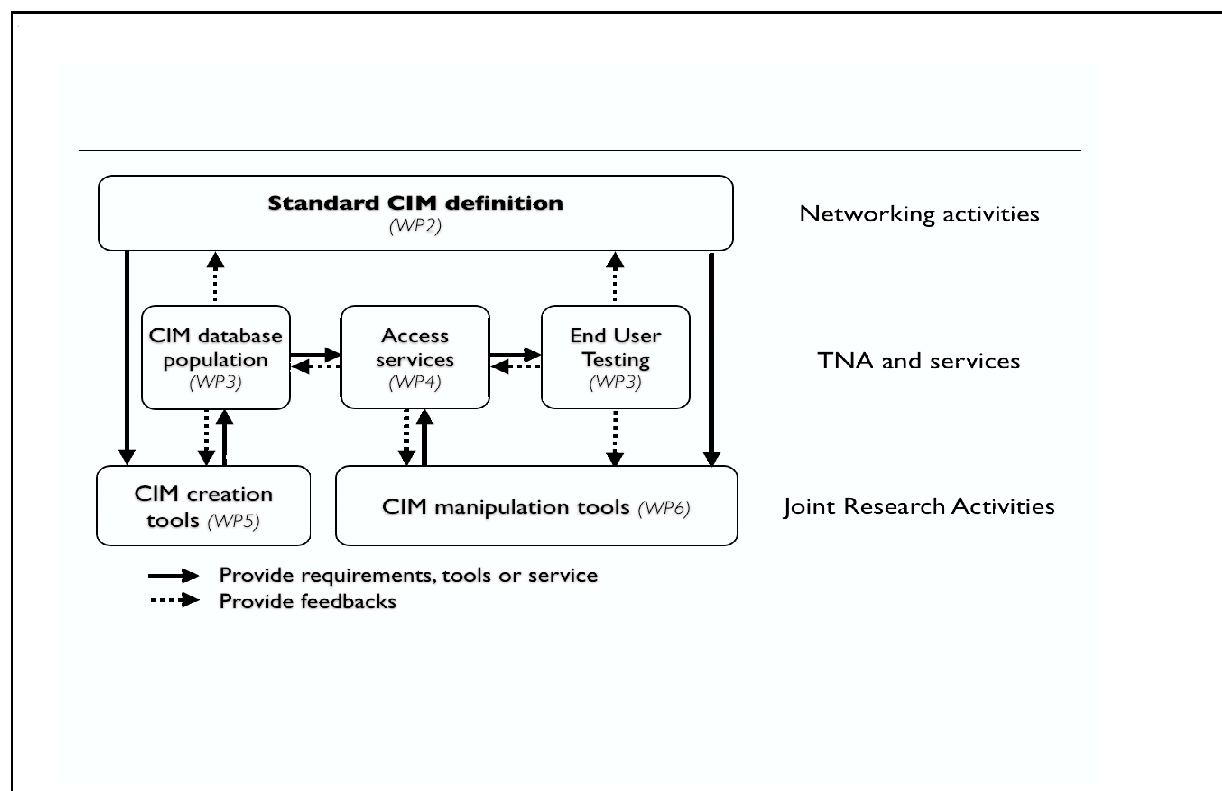


Figure 1: METAFOR activities and work packages (WP) and how they map onto the I3 structure. TNA is Transnational access. Project management, training and dissemination is organised in WP1.

METAFOR will exploit the EU I3 model by carrying out networking activities (in particular, pooling and coordinating existing but distributed metadata developments to establish a common metadata standard, the CIM, in WP2) in support of the establishment of trans-national access and a service infrastructure (WP3, WP4) to make the data and models described by the CIM available to interested parties both within and outside the METAFOR consortium (Figure 1). The CIM design will be iterated throughout the project in WP2, with an early prototype version being used for requirements and concept development in four other work packages. The service infrastructure and population activities will depend on two joint research activities: WP5 and WP6 which will provide tools to manipulate and populate the CIM, respectively. This research activity cannot be done without gathering a wide range of experts to deliver groundbreaking technical developments.

METAFOR participants are:

1. University of Reading- coordinator (UK)
2. Science and Technology Facilities Council (British Atmospheric Data Centre-BADC) (UK)
3. Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (France)
4. Model and data / Max-Planck Institute for Meteorology (Germany)
5. CNRS/Institut Pierre Simon Laplace (France)
6. University of Manchester (UK)
7. Met Office (UK)
8. Administratia Nationala de Meteorologie (Romania)
9. Météo-France (France)
10. CLIMPACT (France)
11. Princeton University (USA)

Whatever the outcome, the making of the proposal will strongly benefit PRISM activities.

3 Coupling & I/O PAE

During the last year, the following tasks were performed in the PAE Coupling and I/O:

3.1 Development of the OASIS3 coupler (CERFACS and CRAY)

- A new version, Oasis3_prism_2-5, was released in September 2006 and related documentation is currently available as a PRISM report at http://www.prism.enes.org/Publications/Reports/oasis3_UserGuide_T3.pdf. This version includes in particular:
 - few bug corrections;
 - improvements in the SCRIP interpolation library
 - a new routine in the PSMILE interface library, `prism_get_freq` to retrieve the coupling frequency of a particular coupling field;
 - the merging of local versions of the `mpp_io` library used for OASIS3 and OASIS4;
 - diverse adaptations for compiling with the NAG compiler.
- This official release was ported on the CRAY X1, XD1, XT .
- Migration from CVS to Subversion for OASIS3 source management was recently finalized at CERFACS. OASIS3 sources are now also available through http at <http://www.cerfacs.fr/prismsvn/trunk>.

3.2 Development of the OASIS4 coupler (NEC-CCRLE,CERFACS,CNRS,SGI, NEC HPCE)

- A new beta version, OASIS4_0_2, which sources were reorganised to fit the PRISM standard directory structure and Standard Compiling Environment (SCE), was released in September 2006 and related documentation is currently available as a PRISM report at: http://www.prism.enes.org/Publications/Reports/OASIS4_User_Guide_T4.pdf .
- The 2D conservative remapping was implemented and final adjustments are currently underway.
- The implementation of the full parallel global search was finalized for all interpolations except the 2D conservative remapping, for regular, irregular, and gaussian reduced grids and is currently under final quality check.

- A report on the OASIS4 functionalities needed for the realization of the CICLE project funded by the French "Agence Nationale de la Recherche" (ref: ANR-05-CIGC-04) was written (see http://dods.ipsl.jussieu.fr/omamce/CICLE/Documents/Rapport_fonctionsOASIS.pdf)
- Extensive validation of 2D interpolations was started, with some priority for the interpolation needed for the CICLE project.
- Efficiency of the OASIS4 Transformer was improved.
- The development of a Graphical User Interface for OASIS4 is currently underway, involving exclusively CERFACS manpower.
- Active user support and bug fixes were regularly provided for the current beta tester groups: the EU GEMS project, SMHI, and the UK Met Office.
- OASIS4 was presented in November 2006 to the ACCESS workshop in Melbourne.
- OASIS4 and the CICLE project were presented at the European Geosciences Union General Assembly 2007 Vienna, Austria, 15-20 April 2007.
- OASIS4 was presented at the "International workshop Integrated systems of meso-meteorological and chemical transport models on May 21-23, 2007 in Copenhagen, Denmark, organised by the COST Action 728 (see <http://www.cost728.org/>).
- Migration from CVS to Subversion for OASIS4 source management was recently finalized at CERFACS. OASIS4 sources are now also available through http access at <http://www.cerfacs.fr/prismsvn/branches/development> .

4 Integration and Modelling Environements PAE

4.1 Met Office contribution

FCM Integration within the UM has been improved to allow users to extract their required software configuration, build the executable and launch the job with a single button. We have upgraded the version of subversion used to 1.4.3 from 1.3.2 and there have been developments in the light of user feedback. See release notes for version 1.2 at

http://www.metoffice.gov.uk/research/nwp/external/fcm/doc/release_notes/1-2.html

Since that date, there has been a significant code restructure to allow for future developments as well as performance improvements and more error checking added.

Outline information for FCM remains available on the Met Office website at

<http://www.metoffice.gov.uk/research/nwp/external/fcm/>

IPSL still remain interested in the FCM software and have immediate plans to move to Subversion and Trac.

A document has been produced on a process to allow a number of different groups to collaborate using FCM using separate sub-version servers. Comments has been invited on the PRISM wiki at <http://www.cgam.nerc.ac.uk/pmwiki/PRISM/index.php/Main/CollaborationWithFcmProjects>

4.2 CERFACS contribution

The migration from CVS to SVN has been completed for the OASIS3 and OASIS4 repositories in Cerfacs. Access through http is now available at <http://www.cerfacs.fr/prismsvn>. This is considered as a first step to using the Met Office FCM software. The enhanced wiki and issue tracking system for software development projects, TRAC, has also been recently installed at Cerfacs for OASIS3 and OASIS4.

4.3 M&D contribution

- Support activities
 - for using COSMOS models within the SCE/SRE environments.
 - for the Millenium project (MPI-Met, GKSS, AWI, PIK, Univ. Helsinki, CSC Helsinki, etc.) primarily with postprocessing and WDCC DB filling within the SCE/SRE environments.
 - for interfacing with OASIS3: REMO and CLM
- Consortial experiments
 - The German consortial RCM (CLM) experiments (IPCC) are still running using the SCE/SRE environments.
- Repository
 - The PRISM SVN tag prism_2-5 (toyclim(oasis3), toyoa4(oasis4)), announced in December 2006 was ready to be created from branches/prism_2-5_2. Modifications included into branches/prism_2-5_2 and other included later into branches/prism_2-5_3 have been merged meanwhile into trunk. No tag was created. Since then only minor modifies (bugs etc.) were committed to the trunk of the PRISM repository. The developments in SCE/SRE reported on below are not yet committed to the trunk.
 - A script has been written that checks out complete coupled models for specific platforms (e.g. toyclim or toyoa4 on SX at DKRZ) similar to the CVS download of CVS modules done before.
- SCE/SRE development
 - The SCE user interface is revised; input to SCE scripts is now by parameters and (long/short) options (documented but not in PRISM repository yet).
 - All available libraries can be declared central and then are not checked any more.
 - TOYCLIM has been tested, and header files with options are provided, to run on Linux*86 (32 bit machines) for the Portland, Lahey-Fujitsu, NAGWare, Intel, and Sun compilers.
 - SRE is modified to be used interactively on NEC-SX machines in production mode (without queuing system)
 - SRE is enabled to optionally run on machines interconnected by rlogin or ssh or on different platforms (no rlogin or ssh). In the first case the different tasks for model integration, pre- and postprocessing trigger each other in order to run synchronically. In the second case the synchronisation is realized by (non)existence of file.
 - Work is still in progress on graphical monitoring (LE-visualisation) of experiments
 - Work has started to extract model/experiment/data meta data from the SCE/SRE and convert the result into a Cera valid xml file.
- Dissemination
 - 2 poster presentations (IMDI (SCE/SRE/WDCC)) have been presented to the 2007 EGU conference.
 - DKRZ/M&D workshop on best practices (WDCC/SCE/SRE/DKRZ platforms) has been hold in March
- GUI

The GUI activities reported in the last report are dropped since ECMWF is not able to commit any manpower any more for those PRISM activities. Instead, it is planned to continue using a more Cera-specific GUI. The basic xml file to be filled-in for Cera-meta data specification will come from the SCE/SRE if the model that produces the data is set up in the SCE/SRE.

5 Data processing and Visualisation PAE

All activities in the Data processing and visualisation PAE during the last months has been spend on the metadata discussion, preparing METAFOR and continuation of data processing integration in the SRE with respect to COSMOS. One additional point is the discussion of standard names in NetCDF/CF. A presentation has been submitted for the GO-ESSP meeting in Paris in June.

6 Metadata PAE

The metadata schemas developed within PRISM for communicating and controlling exchanges between the models and the coupler OASIS4 were ahead of their time with respect to other multi-model coupling environments. The metadata schemas outlined in PRISM documentation are used to:

- describe each application (Application Description, AD)
- describe the relationship between input and output between the different components of the application (Potential Model Input and Output Description, PMIOD)
- specify the general characteristics of the coupled model run (Specific Coupling Configuration, SCC)
- specify the relations between the components in a specific experiment with the application (Specific Model Input and Output Configuration, SMIOC)

Since the development of these innovative metadata schema, PRISM has also lead the development of the Numerical Model Metadata (NMM, <http://ncas-cms.nerc.ac.uk/NMM/>). NMM is now an evolving international metadata standard intended for the exchange of information about numerical model codebases, the associated components and the models/simulations done using them.

Other metadata projects have been initiated or developed; the most relevant are:

- “gridpsec” (<http://www.gfdl.noaa.gov/vb/pdf/gridstd.pdf>) a complete description of the horizontal discretisation of numerical models
- the re-organisation and the development of the CF convention for climate and forecast metadata (<http://www.cgd.ucar.edu/cms/eaton/cf-metadata/CF-1.0.html>) designed to promote the processing and sharing of files created with the NetCDF API. The convention define metadata that provide a definitive description of what the data in each variable represents, and of the spatial and temporal properties of the data. This enables users of data from different sources to decide which quantities are comparable, and facilitates building applications with powerful extraction, regridding, and display capabilities.
- new model coupling projects like the Flexible Unified Model Environment (FLUME, <http://www.metoffice.gov.uk/research/interproj/flume/>) project aiming to generate a completely new system architecture for the Met Office Unified Model system, and ESMF in the USA (<http://www.esmf.ucar.edu/>), which aims to enable and facilitate multi-component model coupling.
- other coupled modelling projects that are keen to exploit existing metadata schema like the Earth System Curator (<http://www.earthsystemcurator.org/>) creating a software environment for assembling, running, and archiving information about climate models. The idea is to make it easier for scientists to perform modeling experiments, and to coordinate with each other on efforts such as Model Intercomparison Projects (MIPs) and Intergovernmental Panel on Climate Change (IPCC) assessments. One of the first goals of the Curator project is to help supplement and refine metadata schema such as the Climate and Forecast (CF) and Numerical Model Metadata (NMM) conventions so that they better serve model analysis projects, and so that they are complete and precise enough to serve as the foundation for software tools that can automate workflow tasks.”
- other data repository projects like COSMOS in Germany and NERC Data Grid (NDG) in UK needing to ingest metadata

The PRISM metadata initiative has followed these developments and, aware that further developments need to be undertaken to ensure that PRISM metadata uses and exploits current metadata schema and methodologies, has proposed the METAFOR project described above in section 2.

Therefore if METAFOR is funded, effort will be available to address the metadata issues in PRISM: the PRISM metadata schema will be fully integrated with other metadata schema, new methodologies of management, maintenance and governance will be implemented, and the PRISM schema will be extended and adapted for use within other coupling frameworks. The development of the CIM, with PRISM metadata schema at its heart, and its general adoption would provide the important next steps for the PRISM metadata. However should METAFOR not be funded, then plans need to be made to make progress in the PRISM metadata PAE. The immediate plans could include:

- consideration of the governance of PRISM metadata so that changes and amendments can be considered and implemented as well as ensuring proper reference to PRISM metadata standard
- the development of a complete glossary embedded within the PRISM metadata schema to promote its broad and general use
- the integration of the gridspec schema into the PRISM metadata
- the development of PRISM schema within Curator and ESMF

Potential new funding initiatives to support this effort continue to be investigated from local and national funding bodies.

7 Computing PAE

The Wiki pages about PAE Computing have been set up and filled with contents (see <http://www.cgam.nerc.ac.uk/pmwiki/PRISM/index.php/Main/PAEComputing>). Preliminary information is provided about file I/O. A selection of links is provided which addresses various topics in the computing area. People have been invited to contribute to PAE Computing and a mailing list has been set up.

2007/2008 will be an interesting period with a large number of new computers through PRISM partners. We will be happy to share technical information regarding sustained performances for climate applications, coupler OASIS, IO, files, post-processing, problems and solutions, ... We know it's important and we hope in input from partners.