

## METAFOR Service Report and Revised Infrastructure METAFOR Deliverable 4.5 M42

<b>PROJECT</b>	
Project acronym	METAFOR
Project full title	Common <u>Metadata</u> for <u>Climate Modelling</u> 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
<b>DOCUMENT</b>	
Deliverable	D4.5 Month 42
Deliverable Title	Service Report and Revised Infrastructure
Document Identifier	METAFOR-D4.5 M42
Date	September 30 <sup>th</sup> 2011
Work Package	WP4 CIM Services
Authors	IPSL
Document Status	Draft
Document Link	<a href="http://bit.ly/qXxe4r">http://bit.ly/qXxe4r</a>

<b>Nature 1</b>		
R	Report	X
P	Prototype	
D	Demonstrator	
O	Other	

<b>Document History</b>			
Version	Date	Comment	Author/Partner
0.1	Sept 20 <sup>th</sup> 2011	Draft	M. Morgan/IPSL
1	Sept 30 <sup>th</sup> 2011		M. Morgan/IPSL

# Table of Contents

<b>1. SUMMARY.....</b>	<b>3</b>
<b>2. WEB SERVICES.....</b>	<b>4</b>
OVERVIEW.....	4
SERVICE OPERATIONS.....	4
<b>3. WEB APPLICATIONS.....</b>	<b>5</b>
CMIP5 QUESTIONNAIRE.....	5
CIM WEB PORTAL.....	5
CMIP5 QC TOOL.....	5
CIM GEONETWORK.....	6
<b>4. COMMAND LINE UTILITIES.....</b>	<b>7</b>
CIMVALIDATOR.....	7
CMIP5QINGESTOR.....	7
CMIP5QCINGESTOR.....	7
CIMGEONETWORKINGESTOR.....	7
TDS2CIM.....	7
CIMFROMCODE.....	7
<b>5. SOFTWARE CATALOG.....</b>	<b>8</b>
30/09/11.....	8

# 1. Summary

The METAFOR Common Information Model, i.e. the CIM, is an ontology designed to become the ipso-facto standard for climate modelling related metadata. METAFOR is responsible for incubating a CIM ecosystem, i.e. a set of useful CIM compliant tools & services. Such an eco-system is essential to encouraging adoption of the CIM by the global climate modelling community.

CIM compliant tools & services consist of a collection of software artefacts designed to standardise the operational activities of climate modelling institutes. Such activities include model documentation, output data discovery, quality control, experimental design, code documentation, coupling configuration, etc. The software artefacts have been developed in the Python programming language and can be divided into 3 categories:

1. Web services;
2. Web applications;
3. Command line utilities.

Such tools & services are necessarily inter-dependent as in the large part they either share data or common components. For example, the CIM search web service executes queries against a back-end database populated with CIM compliant XML instances ingested from the CMIP5 Questionnaire web application. Another example is the THREDDS to CIM command line utility that publishes it's output to the CIM instance web service. Thus whist each individual tool/service delivers a functional subset, collectively the tools & services deliver a platform, at the heart of which is the CIM. This document will pay particular attention to the CIM Portal web application which integrates together several web services and provides a focal point for users wishing to find and interact with CIM instances.

During the 6 month project extension phase the METAFOR tools & services in support of the CIM all became operational. Some deployment issues related to the CIM Web Portal & Services occurred due to security concerns at the BADC. Operational support will continue under the auspices of the IS-ENES. Further development (as driven by community demand) will also take place under the auspices of both IS-ENES and Earth System Grid Federation (ESG-F).

## 2. Web Services

### Overview

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

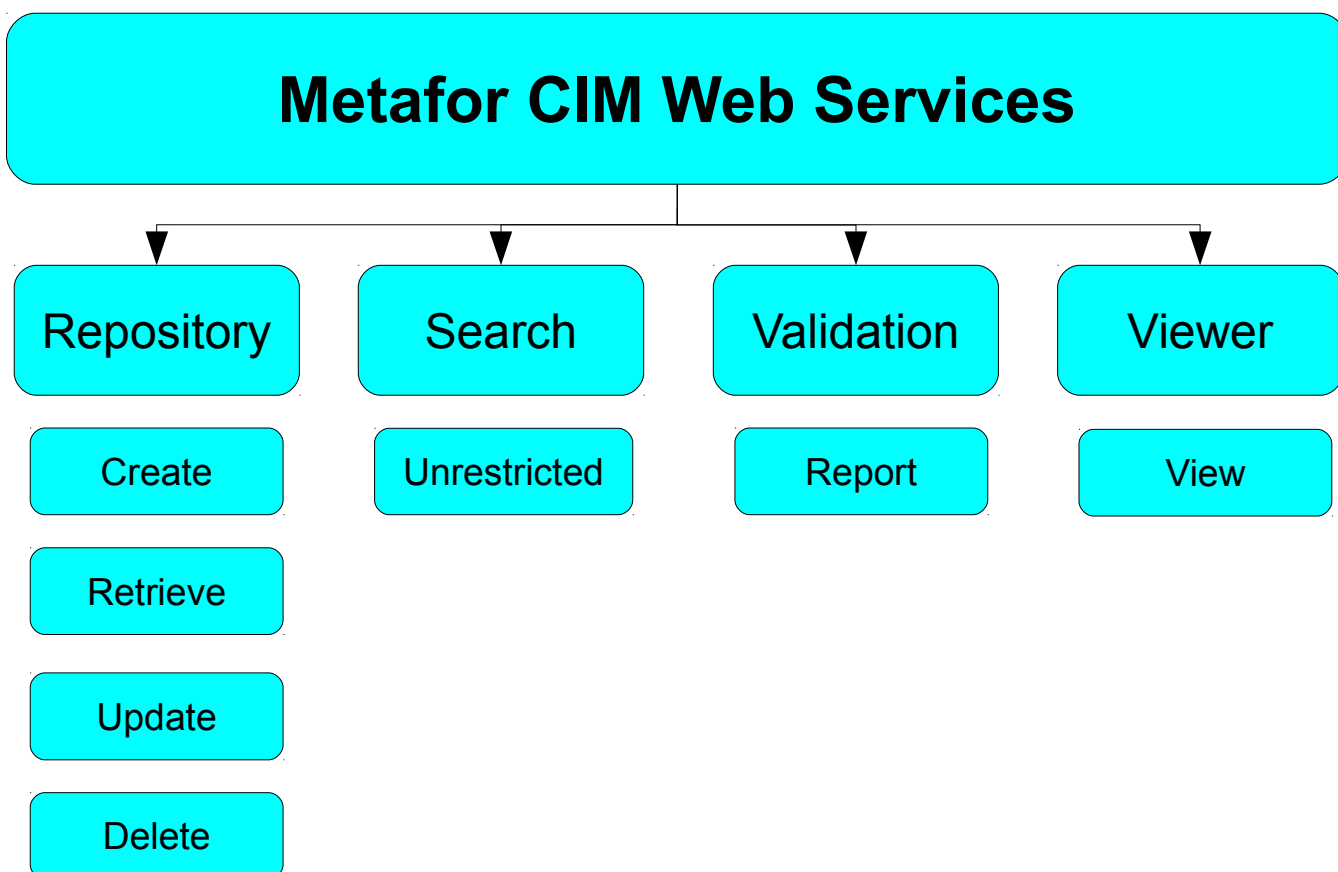
The CIM web services are designed so that institutes can gradually over time integrate them into their own climate modelling workflows. For example an institute may wish to automatically publish meta-data to the CIM web services so that this meta-data will subsequently appear in the CIM Portal (see below).

The web services were refined in response to community demand during the 6 month project extension phase and are thus:

<b>Repository</b>	Supports persistence of CIM documents within CIM repository.
<b>Search</b>	Supports discovery of CIM documents within CIM repository
<b>Validation</b>	Supports validation of potential CIM compliant documents.
<b>Viewer</b>	Supports html rendering of CIM documents.

### Service Operations

Below are outlined the set of CIM web service operations delivered as of September 2011. For technical details see METAFOR EU deliverable 5.5 *CIM Web Service Implementation* [<http://bit.ly/qbPa61>].



### 3. Web Applications

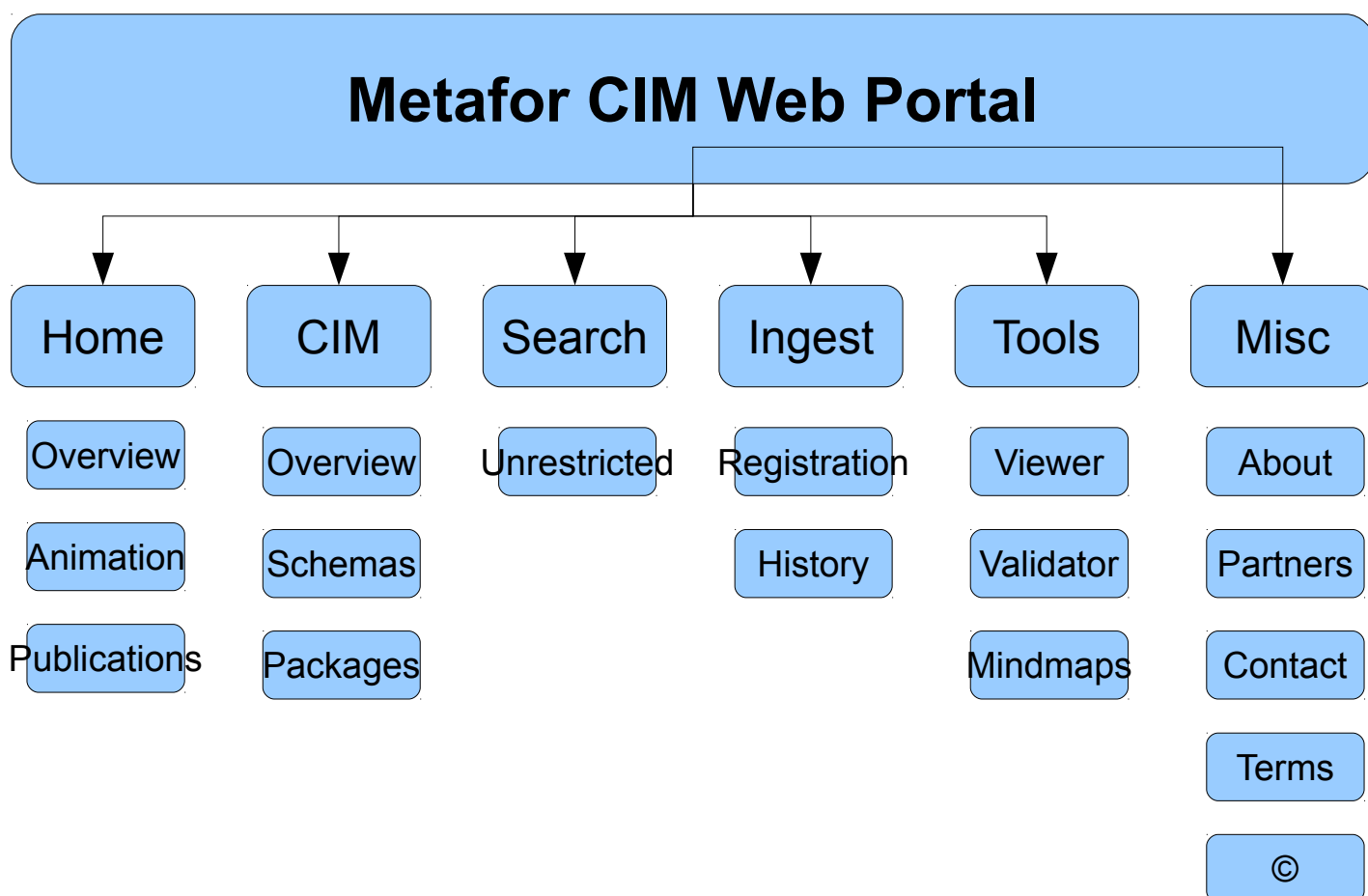
#### CMIP5 Questionnaire

The CMIP5 Questionnaire is in production and is being used by climate modelling institutes worldwide. It collects meta-data from institutes participating in the CMIP5 process. Users are guided through a series of webforms and asked to enter information describing their climate models, simulations, and data. Entered content is being ingested by both the ESG Curator portal (see <http://www.earthsystemcurator.org/>) and the CIM Web Portal (see below) as CIM instances. Available instances can be published as ATOM feeds. The CMIP5 Questionnaire can be found online at <http://q.cmip5.ceda.ac.uk/>.

#### CIM Web Portal

The CIM Portal is designed to deliver a diverse set of use cases placed into several categories: outreach, ontology, search, tools and publishing. The set of potential users can be placed upon a spectrum of familiarity with CIM meta-data ranging from members of the public to climate or metadata experts.

During the 6 month extension the portal was refined in response to community demand and now includes extra tooling support and a technical knowledge regarding the CIM. Plans for advanced search and differencing were dropped due to both resource & time constraint, however these are to be picked up within



the context of Earth System Grid Federation activities. CMIP5 questionnaire sourced data is being displayed via the search and view functions.

Below is outlined the CIM Portal site plan - for technical details see METAFOR EU deliverable 5.5 *CIM Web Portal Deployed* [[http://bit.ly/metafor\\_d43](http://bit.ly/metafor_d43)].

#### CMIP5 QC Tool

The CMIP5 QC (Quality Control) Tool is in production and has been developed under the auspices of METAFOR. It assists the community to document quality related issues connected with CMIP5 output. It

exposes its content, i.e. CMIP5 output quality control issues/reports, as CIM compliant XML documents embedded within an Atom feed. It can be found online at <http://qc.cmip5.ceda.ac.uk/qc>. For further information please refer to its wiki page [[http://bit.ly/cmip5\\_qc\\_tool](http://bit.ly/cmip5_qc_tool)].

### ***CIM Geonetwork***

GeoNetwork provides a GUI for the editing of XML documents. It is an Ajax based standardized and decentralized spatial information management environment. Any kind of data that is based on an XML schema can be made available inside the GeoNetwork framework. This XSD file describes everything about the structure of a record. The GeoNetwork software analyses the XSD schema during start up so that the appropriate structures and settings appear when a record is displayed in 'Show' or 'Edit' mode e.g. an enumeration list in the schema is shown as a selection list at the screen.

A GeoNetwork server in support of the CIM has been set up by the DKRZ group in Hamburg. CIM instances hosted by this server are eligible for ingestion into the CIM Web Portal (see above) and thus are discoverable.

## 4. Command Line Utilities

### ***CIMValidator***

Validates potentially CIM compliant XML documents by passing the document across the internet out to the validation web service within the CIM portal. The validation service returns a report documenting the validity/invalidity status of the document.

### ***Cmip5QIngestor***

Ingests CIM compliant information exposed by the CIMP5 Questionnaire atom feeds. Ingested information can be persisted to a back end XML database for use within the CIM Web Portal.

### ***Cmip5QCIngestor***

Ingests CIM compliant information exposed by the CIMP5 QC Tool atom feeds. Ingested information can be persisted to a back end XML database for use within the CIM Web Portal.

### ***CIMGeonetworkIngestor***

Ingests CIM compliant information exposed by the CIM GeoNetwork server. Ingested information can be persisted to a back end XML database for use within the CIM Web Portal.

### ***TDS2CIM***

Parses a THREDDS metadata server and converts discovered metadata into CIM DataObject documents. Can optionally publish the generated documents to CIM web services. This utility can potentially play a key role in aggregating output data.

### ***CIMfromCODE***

Under development by a group from the University of Manchester. This utility aims to support model code (i.e. Fortran) documentation by parsing Fortran files, extracting metadata, converting the extracted metadata to a CIM compliant structure, and optionally publishing via CIM Web Services (see above) to the CIM Web Portal (see above).

## 5. Software Catalog

Below is defined the list of software artefacts delivered during the course of the Metafor project.

<b>30/09/11</b>	Web Services	CIM Web Services	Complete
	Web Applications	CMIP5 Questionnaire	Complete
		CIM QC Tool	Complete
		CIM Geonetwork	Complete
		CIM Web Portal	Complete
	Command Line Utilities	CIMValidator	Complete
		Cmip5QIngestor	Complete
		Cmip5QIngestor	Complete
		CIMGeonetworkIngestor	Complete
		TDS2CIM	Complete
		CIMfromCODE	Complete
	Deployments	CMIP5 Questionnaire	Production
		CIM QC Tool	Production
		CIM Geonetwork	Production
		CIM Web Services	RC1-Test (awaiting production)
		CIM Web Portal	RC1-Test (awaiting production)