

METAFOR CIM Portal Deployed METAFOR Deliverable 4.3 M33

PROJECT	
Project acronym	METAFOR
Project full title	Common <u>Meta</u> data for <u>Climate</u> 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	D4.3 Month 33
Deliverable Title	Web Service Implementation
Document Identifier	METAFOR-D4.3 M33
Date	January 31 st 2011
Work Package	WP4 CIM Web Services
Authors	IPSL
Document Status	Draft
Document Link	http://metaforclimate.eu/

Nature 1		
R	Report	X
P	Prototype	
D	Demonstrator	
O	Other	

Document History			
Version	Date	Comment	Author/Partner
0.1	Feb 08 th 2011	Draft	M. Morgan/IPSL
1.0	Feb 15 th 2011	Final	M. Morgan/IPSL P. Slavin/University of Manchester

Table of Contents

1. SUMMARY.....	3
2. ACTORS.....	4
3. SITE PLAN.....	5
4. TECHNICAL IMPLEMENTATION.....	8
5. DEPLOYMENT & HOSTING.....	10
6. BIBLIOGRAPHY (LINKS VALID AS OF 01ST AUGUST 2010).....	11

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 eco-system, i.e. a set of useful CIM compliant tools, services & resources. Such an eco-system is essential to encouraging adoption of the CIM by the global climate modelling community

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.

This document details the Metafor CIM Web Portal Deployment by outlining the following:

1. Actors using the portal;
2. Site plan;
3. Technical implementation;
4. Deployment & hosting;
5. Bibliography of links to various web-sites and pages of interest.

2. Actors

The CIM Web Portal must support a wide array of actors, these actors can be sorted by their relative expertise with climate related meta-data. The level of expertise will determine the functionality available to the actor when they connect to the portal.

Member of Public

Arrives at the 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.

Policy Maker

Arrives at the portal as a result of a generalised outreach program. Is curious to gain a deeper understanding of the rationale behind the Metafor project. Wants to integrate information held on the portal into their policy making process. May wish to experiment with simple search functions.

Scientific User

Arrives at the portal as a result of either a scientific collaboration or a Metafor publication/presentation. Is curious to gain a deeper understanding of the rationale behind the Metafor project. Wants to understand some of the science encoded in the CIM ontology.

Climate Scientist

Arrives at the portal as a result of a focussed community outreach program. Requires a deep understanding of the rationale behind the Metafor project. Will experiment with the search engine.

Climate Modeller

Arrives at the portal as a result of a focussed community outreach program. Requires a deep understanding of the rationale behind the Metafor project & the CIM. Will leverage the search engine in their daily working life. Needs to be notified of changes in the portal. May provide input into the portal evolution.

Data Analyst (end user)

Arrives at the portal as a result of a focussed community outreach program. Requires a deep understanding of the rationale behind the Metafor project & the CIM. Will leverage the search engine in their daily working life. Needs to be notified of changes in the portal. May provide input into the portal evolution.

Data Expert

Arrives at the portal as a result of a focussed community outreach program. Requires a deep understanding of the rationale behind the Metafor project. Will experiment with the search engine.

Meta-Data Expert

Arrives at the portal as a result of a focussed community outreach program. Requires a deep understanding of the rationale behind the Metafor project & the CIM. Will leverage the search engine in their daily working life. Needs to be notified of changes in the portal. Provides input into the portal evolution.

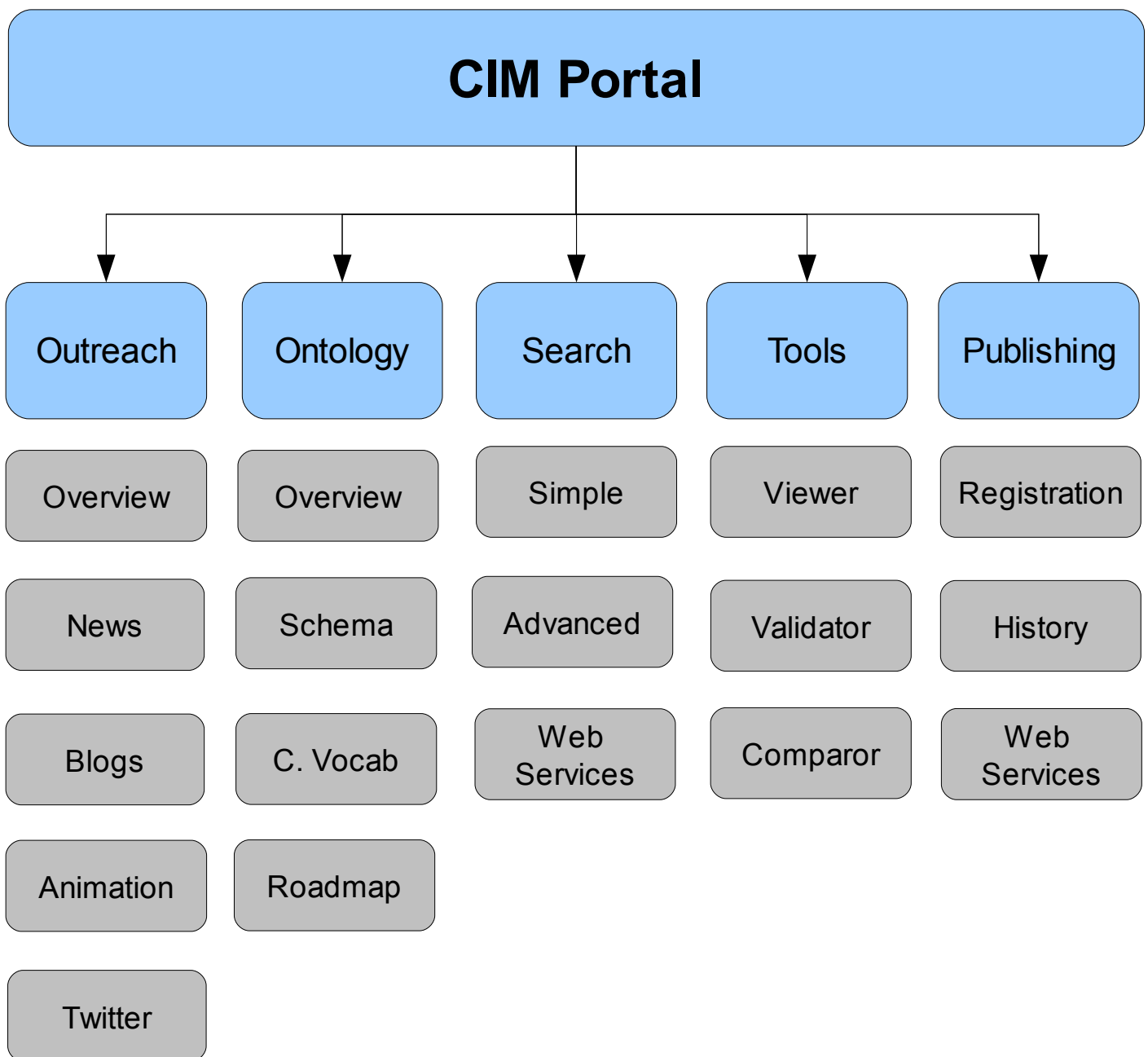
Meta-Data Administrator

Arrives at the portal as a result of a focussed community outreach program. Needs to be certain that meta-data generated by their organisation appears in and is accessible from the search engine results. May leverage the search engine in their daily working life. Will wish to be notified of changes in the portal. Provides input into the portal evolution.

3. Site Plan

The Metafor team formally drafted a set of use cases deemed of potential use to the community. These were reviewed and either rejected or accepted as being of real use to the community. Accepted use cases were ranked in terms of impact, importance and required developer resources to implement. A final list was agreed upon and acts as the basis for all development work.

These use cases have been deconstructed into a site plan (see below) that outlines how functionality is partitioned across the portal. The plan is essentially a set of nodes within a hierarchy, the root node is the portal itself. From a security perspective each node is mapped to a list of user roles (i.e. actors) and the portal enforces authorisation policy. Thus for example members of the public are not granted access to CIM tools.



Outreach

Communicating the CIM and it's evolution to the community

The community must be made aware of the ongoing work of Metafor. Gradually, as the CIM eco-system matures, the outreach task transitions to ensuring an active 2-way engagement with the community.

Overview

An overview of the Common Information Model and associated eco-system.

News

Dedicated blog feed for publishing news relevant to the CIM community.

Blogs

A set of ongoing articles by Metafor personnel covering various aspects of the CIM eco-system.

Animation

The official CIM animation developed by Team Digital Preservation and the Metafor team.

Twitter

Twitter feed to broadcast news such as events, publications ... etc.

Ontology

Formal technical resources concerning the CIM ontology

The CIM ontology is complex and lies outside the normal domain of the climate community. Thus in-depth technical & training resources are provided to allow users to get up to speed with the CIM ontology.

Overview

An overview of the Common Information Model ontology and it's evolution.

Schema

Documentation, downloads, samples of official CIM schemas.

C. Vocab

Documentation, downloads, samples of controlled vocabularies referenced by CIM schemas.

Roadmap

The future of the CIM ontology in the short, medium & long terms.

Search

Accessing CIM meta-data held within the CIM repository

Users need to be able to search over meta-data stored within the official CIM repository. The CIM search engine supports simple and advanced search and is designed to be user friendly and as quick as possible.

Simple

Simple keyword search that returns collection of matching models, experiments, simulations & data.

Advanced

A set of advanced search that allows users to perform fine-grained queries against models, experiments, simulations & data.

Web Services

Technical documentation of the underlying CIM query web service – allows the direct integration of CIM queries in 3rd party portals.

Tools

User friendly CIM value add functions

Specialized tools will enrich the CIM eco-system by obviating the need for institutes to build their own. The tools support essential tasks such as CIM meta-data view, validation & comparison.

Viewer

Renders views of a CIM meta-data document derived either from a file on the user's local machine (standalone viewer) or from the search engine (integrated viewer).

Validator

Allows user to validate a CIM meta-data document against the latest schemas.

Comparator

Allows user to compare multiple CIM meta-data documents returned within a simple search.

Publishing

Publishing remotely held CIM meta-data to the CIM repository

Meta-data administrators need to publish CIM meta-data. Ideally they register meta-data servers for inclusion in overnight ingestion jobs. Alternatively they can publish to dedicated web-services.

Registration

Allows user to register a meta-data server for inclusion in automatic overnight batch ingestion jobs.

History

Displays to the user the history of overnight batch ingestion jobs.

Web Services

Technical documentation of the underlying CIM repository/validation web services. Allows developers to integrate such services into CIM meta-data publishing pipelines.

4. Technical Implementation

The CIM portal is a rich internet application (RIA), i.e. a web-based user interface (UI) that consumes data served by back-end web services. Separating the UI from such back-end services is a key tenet of service-oriented architecture and permits 3rd parties to integrate CIM web services into their own portals.

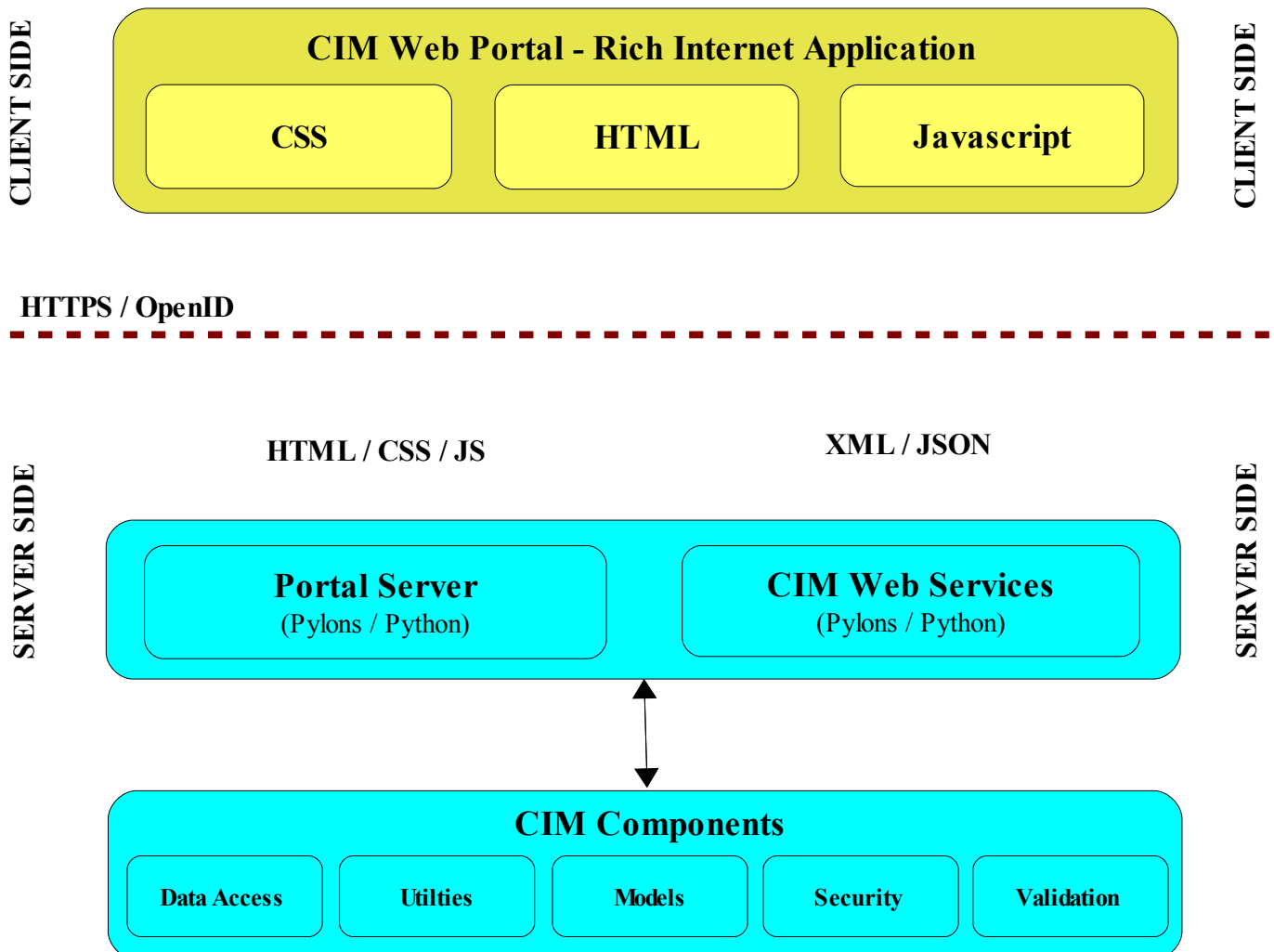
On the client side (i.e. the browser) the portal is developed using HTML, Cascading Stylesheets (CSS) & Javascript (JS): HTML manages UI layout; CSS manages UI styling; JS manages UI behaviour (including calling CIM web services). JQuery, a well-known JS & CSS framework, is extensively leveraged.

On the server side the CIM portal server is built using Pylons, a python based web development framework. Valid page requests are routed to Pylons controllers for processing. These controllers typically orchestrate calls to backend components & return html responses for rendering by the client browser.

Data rendered in the UI is generally served by back-end web services, i.e. CIM web services. JS packages invoke these services by issuing AJAX (Asynchronous Javascript and XML) calls to the relevant service. The web services typically return data in one of two formats: XML or JSON. The CIM web services orchestrate calls into the same backend components as invoked by the CIM portal server.

All browser to server communications are encrypted. All users are authenticated via the open source security standard known as OpenID. An OpenID filter is provided by the Earth System Grid Federation. Each user will be expected to have an OpenID account that will be mapped to the relevant user role(s).

The portal is thus a set of web pages served up by the portal web server. Each page delivers functionality in support of one or more use-cases. Pages render data served up by CIM web services. All calls to the back end servers are secured in a consistent fashion via the HTTPS & OpenID standards.



5. Deployment & hosting

The British Atmospheric Data Centre (BADC) has taken responsible for hosting the Metafor CIM Portal. Both the CIM portal server and CIM web services run upon Linux machines running the well-known Apache web server. The `mod_wsgi` Apache module is used to host the the Pylons web framework upon which both the CIM portal server and CIM web services are built.

The CIM Portal server runs under a dedicated Linux user account with highly restricted access in order to minimise the attack vector. A security audit has been performed using Open Web Application Security Project (OWASP) guidelines in order to ensure that the hosting & deployment process is as secure as possible.

The CIM Web Portal at the time of writing (08th February 2011) is in development mode. Development snapshots are being released to a test server at a regular basis in order to elicit user feedback as early as possible in the development lifecycle (as per Agile guidelines). The first production release of the CIM Web Portal is scheduled for 1st May 2011.

6. Bibliography (links valid as of 01st August 2010)

Metafor	Overview	http://metaforclimate.eu
	Deliverables	http://metaforclimate.eu/trac/wiki/deliverables
	CIM	http://metaforclimate.eu/trac/browser/CIM/trunk
SOA	Overview	http://en.wikipedia.org/wiki/Service-oriented_architecture
	HTTP 1.1	http://www.w3.org/Protocols/rfc2616/rfc2616.html
	Web Services	http://en.wikipedia.org/wiki/Web_service
REST	Overview	http://en.wikipedia.org/wiki/Representational_State_Transfer
	Atom Publishing	http://bitworking.org/projects/atom/rfc5023.html
	XML encoding	http://en.wikipedia.org/wiki/XML
	JSON encoding	http://www.json.org/
	Base64 encoding	http://en.wikipedia.org/wiki/Base64
Security	PKI	http://en.wikipedia.org/wiki/Public_key_infrastructure
	X509 Certificates	http://en.wikipedia.org/wiki/X.509
	OpenID	http://openid.net/
Hosting	Linux	http://www.linux.org/
	Apache	http://www.apache.org/
Development	Pylons	http://pylonshq.com/
	Python	http://www.python.org/
	Schematron	http://www.schematron.com
	SKOS	http://www.w3.org/2004/02/skos/
Databases	eXist	http://exist.sourceforge.net/
	XQuery	http://www.w3.org/TR/xquery/
	PostGres	http://www.postgresql.org/
	T-SQL	http://www.tsql.de/transact-sql