

ESGF Metadata Telco 4
Thursday 27 October 2011, 0900 MT

Attending:

- Mark Morgan
- Gerry Devine
- Allyn Treshansky
- Sylvia Murphy
- Luca Cinquini
- Eric Nienhouse
- Cecelia DeLuca
- Tobias Weigel

Apologies:

-

Agenda:

1. combining elements of CIM portal and Curator trackback in the context of ESGF nodes
 - 1a. review CIM Portal
 - 1b. review Curator TrackBack
 - 1c. discuss how to integrate functionality into ESGF nodes

Minutes:

Overview:

- The CIM Portal is a service-based web application built around the CIM
 - it provides a CIM Viewer
 - and CIM Search
- ESG Curator trackback is accessed via ESG gateway under the search engine
- Trackback is ingesting CIM metadata from CMIP5 questionnaire
- CIM Portal is ingesting CIM metadata from CMIP5 Questionnaire and other sources (such as CIM QC Tool and other projects such as ENSEMBLES and CORDEX)
- Curator Trackback uses RDF triple store in the backend; CIM Portal uses XML database in the backend.

review of CIM Portal:

- Python-based web application (Python on server-side, JavaScript on client-side)
- connects to an XML database; XML is (currently) the only representation of metadata.
- Ontology Section:
 - static knowledge-based pages
- Tools Section:
 - viewer - similar to Curator Trackback
 - validator (requires C Library patch) - checks XML file against XSD and Schematron rules (derived from mindmaps);
 - mindmap - checks mindmaps against a set of rules

- Repository Section:
 - provides search against an eXist XML database (populated by ATOM feed from DKRZ & CMIP5 Questionnaire)
 - can constrain by document type and document facet and keywords
 - results are returned in a table
 - each table row can be viewed in the viewer

Users have to go to these portals to get at the metadata; but they would prefer the metadata to appear integrated into their own sites.

Mark M. has prototyped a lightweight viewer that can be embedded in any portal; JavaScript plugin which communicates to web services running w/in the CIM Portal; returned metadata is JSON not XML

Use Cases for this viewer:

- 1) statistical downscaling portal (U. of Cantabria)
- 2) NCAS local model description pages

Curator TrackBack Review:

- each gateway harvests CIM documents from METAFOR ATOM feeds and builds up a triple store to drive the interface
- no webservice hooks in trackback; no public API to access services

Possible Areas of Work:

- LC: What about having CIM services that operate off of multiple distributed repositories instead of a single local database?
- LC: another way to create CIM instances apart from going to CMIP5 Questionnaire
 - work is underway on this w/in COG
- LC: incorporate a display of observational metadata to ESGF
- adding definitions of terms to CIM display
- merging the two display technologies
 - currently there are 2 metadata formats (RDF & XML) and 2 underlying technologies; we want the best of both worlds
 - note that there is effort underway to serialize the CIM into native Python objects
- differencing
 - AT: this technology has been started in XML, but it makes sense to compare Python objects instead of XML.
- other items from recent prioritization call

Curator Trackback interface is in-use; but metafor viewer is a better technical direction
MM: should be able to move from prototype to production code in January/February 2012

Working Procedures:

CD: Does Metafor want a collaboration w/ Curator to work on the CIM Portal?

MM: Yes

MM: ESGF should be the authoritative source; metadata should be available from ESGF nodes

CD: Is Metafor willing to have their priorities influenced by an external group?

CD: what metadata systems are we going forward with and who are the teams that will do the work?

CD: we want an EU-led project

MM: IS-ENES has scope for continued development of CIM technologies

MM: IS-ENES2 does not start funding until 2013

GD: All published questionnaire instances have a link to the Curator trackback page

LC: What is the URL to use to link to the data returned by search?

TODO: bring up working procedures at the next METAFOR meeting

TODO: write a strategic document