The Prod-Trees Project
Extending netCDF and CF conventions to support enhanced Earth Observation Ontology services

Paolo Mazzetti[1], Bernard Valentin[2], Manolis Koubarakis[3], Stefano Nativi[1] e Sabina Di Franco[1]

[1] National Research Council - Institute for Atmospheric Pollution Research (CNR-IIA) Monterotondo (RM), Italy, [2] Space Applications Services N.V., Zaventem, Belgium, [3] University of Athens, Athens, Greece

Project Presentation

Prod-Trees (Enriching Earth Observation Ontology Services using Product Trees) is a project funded by ESA/ESRIN and aiming to enhance the CF-netCDF product format and vocabulary to allow storing metadata that better describe the products, and in particular EO products.

The project started in December 2012 and will end in May 2014. The project is coordinated by Space Applications Services (Belgium) and carried out with CNR (Italy) and University of Athens (Greece).

Rationale

Access to Earth Observation products remains not straightforward for end users in most domains. Semantically enabled search engines, generally accessible through Web portals, have been developed. They allow searching for products by selecting application-specific terms and specifying basic geographical and temporal filtering criteria. This mostly suits the needs of the general public, but the scientific communities require more advanced and controlled means to find products. Ranges of validity, traceability, accuracy, uncertainty, are concepts that must be taken into account in research activities. RARE and SMAAD are two such semantically-enabled search engines: RARE permits to search for EO Products by entering terms used by the users in their application domains. In SMAAD, the focus is on the interoperability between the available ontologies: the Web interface allows the users to navigate ontologies and terms that characterize EO products. RARE and SMAAD rely on a preparatory work (ontology mapping, EO product characterization, reasoning rules) that is barely re-usable in other systems without integrating the original software components. Also, they require the EO products to be annotated with controlled vocabularies that have not been standardized. These limitations prevent other organizations (scientific communities, commercial companies) from implementing their own compatible solutions, adapted to their own requirements and application domains.

Objectives

The project will bring a standardized solution that permits annotating EO products, in netCDF and related conventions, in a manner that official and third-party software libraries and tools will be able to search for products using advanced tags and controlled parameter names (e.g. Ontology Browser and Reasoner). Annotated EO products will be automatically supported by all the compatible software. The entire product information will come from the annotations and the standards and there will be no need for integrating extra components and data structures that have not been standardized. In the course of the project, the most important and popular open-source software libraries and tools will be extended to support the proposed extensions of the proposed CF-netCDF. The result will be provided back to the respective owners and maintainers for ensuring the best dissemination and adoption of the extended format.

Expected Outcomes

- Definition of EO Conventions for netCDF (EO-netCDF) compliant with other relevant conventions including CF Conventions and Uncertainty Convention (netCDF-U)
- Definition of the EO Vocabulary and its publication as SKOS/RDF accessible through a SPARQL interface
- Mapping of the EOP O&M onto the EO-netCDF
- Mapping of ISO 19115-1/2 onto the EO-netCDF basing on previous works (ncISO, and GI-cat data model)
- Extension of netCDF Java APIs to fully support the EO-netCDF
- Extension of GI-cat, GI-axe and GI-sem framework to support discovery and access of EO-netCDF datasets
- Enhancement of the RARE system to allow semantic queries based on several ontologies to discovery EO-netCDF datasets
- Ontology Browser and Reasoner for EO-netCDF

Requirement Analysis

The first activity in the project was the elicitation of requirements to describe the needs, wants, desires, expectations and perceived constraints of users and stakeholders. For this activity a "traditional" means of eliciting requirements has been adopted, to this purpose a questionnaire has been used

The Questionnaire

A questionnaire has been prepared in order to collect requirements and suggestions for guiding future work in the Prod-Trees project. The questionnaire was intentionally kept short in order to encourage people to provide answers. In the end the questionnaire included:

- 5 questions/fields for respondent's profiling
- 5 questions on core Prod-Trees issues
- 1 question on use-cases' proposal

The questionnaire has been published online using Google Forms at the address http://goo.gl/CaEIT.

The questionnaire has been submitted to the identified contact persons and mailing list, asking to either fill the form on-line or the document off-line.
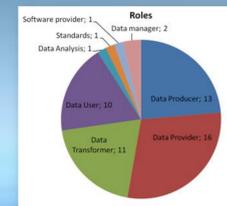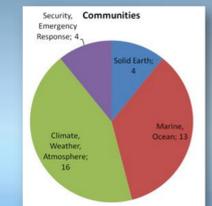
Validation Group

A Validation Group has been created including all the respondents to the requirements questionnaire. The Validation Group will be contacted on selected topics during the project to assess the progresses and validate the outcomes. The Validation Group is co-chaired by Stefano Nativi (CNR) and Ben Domenico (UCAR/UNIDATA).
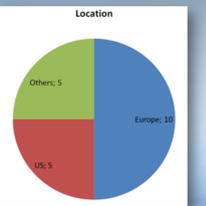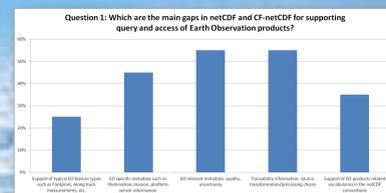
Questionnaire Responses
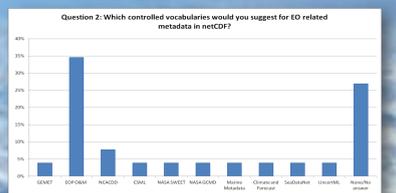

Respondents' roles

Represented communities

Respondents' location

Question 1: main gaps in netCDF

Question 2: controlled vocabularies

Question 3: ontologies for thematic queries

Question 4: libraries and tools

Question 4: other libraries and tools

Respondents also provided specific suggestions on vocabularies and ontologies, libraries/tools, products and possible case studies
For further information:
http://wiki.services.eoportal.org/tiki-index.php?page=PRODTREES

Contacts: paolo.mazzetti@cnr.it, difranco@iia.cnr.it, bernard.valentin@spaceapplications.com, koubarak@di.uoa.gr