Geophysical Research Abstracts
Vol. 15, EGU2013-10873, 2013
EGU General Assembly 2013

ODISEES: An Ontology-Driven Interactive Search Environment for Earth Science

Elisabeth Huffer (1), Michael Little (2), John Kusterer (2), and Mark McInerney (3)

(1) Lingua Logica, Baltimore, MD, United States (beth@lingualogica.net), (2) Atmospheric Science Data Center, NASA Langley Research Center, (3) Computational and Informational Sciences and Technology Office, NASA Goddard Space Flight Center

The ability to compare climate model outputs to other models and to satellite observations has emerged as an issue of growing importance to the Earth science community. Making such comparisons, however, currently presents a significant technical challenge, given that the relevant data products and model outputs are created by very different communities, working with very different instruments, algorithms, processing techniques, etc. Users face two significant challenges: 1) discovery of suitable data; and 2) understanding the different data products and/or model outputs available.

Data resources are often difficult to find and use because much of the data rely on idiomatic encoding systems that require significant expertise and familiarity to decode. Semantic technologies – ontologies, triple stores, reasoners, linked data – offer functionality for addressing this issue. Ontologies can provide robust, high-fidelity, high-precision domain models that illuminate the different contexts within which the numeric data values were derived, and can help explain potential differences in values. By making such contexts transparent to end-users, ontologies can serve as a framework for discovering, evaluating, comparing and integrating data from disparate products, providing support for intelligent search applications that allow users to discover, query, retrieve, and easily reformat data from a broad spectrum of sources.

As part of an on-going effort at NASA Langley's Atmospheric Science Data Center, and in cooperation with the Computational & Information Sciences & Technology Office at the Goddard Space Flight Center, we have developed a semi-automated method for finding and comparing equivalent data variables across disparate datasets. We will demonstrate a prototype variable matching service that is supported by an ontology that models a subset of variables from several data and model output products, including the Coupled Model Inter-comparison Project (CMIP5), the Modern-Era Retrospective Analysis for Research and Applications (MERRA) and the Clouds and Earth's Radiant Energy System Experiment (CERES).

An automated mapping among comparable variables from each of the three programs was accomplished by creating a queriable ontological model ("ontology") of the essential characteristics of both the sample variables and the parameters they represent. Each parameter was modeled in detail in the ontology and mapped to appropriate variables. Queries of the ontology and triple store are used to match comparable variables by searching for those variables that share a specified set of essential characteristics. This approach allows a user to rapidly transform a search for a parameter, such as aerosol optical depth, into an understanding of the various choices and to select the one which most closely matches their interest.