

[ODISEES: An Ontology-Driven Interactive Search Environment for Earth Science](#) [1]

Submitted by Beth Huffer on Fri, 2012-12-21 18:54 **Event:** [Winter Meeting 2013](#) [2]

Abstract:

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. Intercomparison among data products from different observational sources continues to be important as well. 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. Researchers must be able to find those data products and model outputs that are both relevant to their projects and suitable for comparison. In addition, to use the data appropriately they must have a strong understanding of the precise meaning of the numbers in the files and how they were created.

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 domain models that highlight 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, an ontology can serve as a framework for discovering, evaluating, comparing and integrating data from disparate products. Reasoning engines and triple stores can leverage this type of ontology to support intelligent search applications that allow users to discover, query, retrieve, and easily reformat data from a broad spectrum of sources, without losing track of the distinctions among them

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 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[1] ("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. Variables, in turn, are represented as a set of specifications describing the data owner's intended interpretation of the variable with sufficient detail to allow a scientist to assess, with reasonable precision, the nature of the information encoded by such variable. 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.

In addition to describing characteristics of the variables themselves, the ontology associates each data variable with the program from which it is derived and the data products in which it occurs, and provides additional information about how the associated parameter is measured within the selected product. This allows users to not only find parameters that satisfy their search criteria, but also to find data products that contain those parameters. Scientists will be able to discover data products that exactly meet their particular criteria, link to information about the instruments and processing methods that generated the data; and compare and contrast related products to improve their understanding.

[1] The model is *ontological* rather than *linguistic* because the terms that comprise the model are meant to stand in for the actual objects and processes that make up the Earth science domain. This allows us to assign terms from data vocabularies and standards to the *objects* or *processes* that they refer to, rather than just mapping synonymous terms to each other, such as is done in a thesaurus.

Collaboration Area: [Semantic Web](#) [3]

Author(s):

Name: [Beth Huffer](#) [4]

Organization(s): [Lingua Logica](#) [5]

Email: beth@lingualogica.net [6]

Name: [Mike Little](#) [7]

Organization(s): [Atmospheric Science Data Center](#) [8], [NASA Langley Research Center](#) [9]

Email: m.m.little@nasa.gov [10]

Name: [John Kusterer](#) [11]

Organization(s): [Atmospheric Science Data Center](#) [8], [NASA Langley Research Center](#) [9]

Email: john.m.kusterer@nasa.gov [12]

Name: [Ryan Kohl](#) [13]

Organization(s): [Inovex](#) [14]

Email: ryanckohl@gmail.com [15]

Name: [Mark McInerney](#) [16]

Organization(s): [NASA Goddard Space Flight Center](#) [17], [Computational and Information Sciences & Technology Office](#) [18]

Email: mark.mcinerney@nasa.gov [19]

Keywords: [ontologies](#) [20]

[faceted search](#) [21]

[data discovery](#) [22]

[data mapping](#) [23]

Source URL: <https://commons.esipfed.org/node/1130>

Links

[1] <https://commons.esipfed.org/node/1130>

[2] <https://commons.esipfed.org/taxonomy/term/464>

[3] <https://commons.esipfed.org/collaboration-area/semantic-web>

[4] <https://commons.esipfed.org/node/1115>

- [5] <https://commons.esipfed.org/taxonomy/term/602>
- [6] <mailto:beth@lingualogica.net>
- [7] <https://commons.esipfed.org/node/1117>
- [8] <https://commons.esipfed.org/taxonomy/term/603>
- [9] <https://commons.esipfed.org/taxonomy/term/604>
- [10] <mailto:m.m.little@nasa.gov>
- [11] <https://commons.esipfed.org/node/1127>
- [12] <mailto:john.m.kusterer@nasa.gov>
- [13] <https://commons.esipfed.org/node/1128>
- [14] <https://commons.esipfed.org/taxonomy/term/605>
- [15] <mailto:ryanckohl@gmail.com>
- [16] <https://commons.esipfed.org/node/1237>
- [17] <https://commons.esipfed.org/taxonomy/term/246>
- [18] <https://commons.esipfed.org/taxonomy/term/639>
- [19] <mailto:mark.mcinerney@nasa.gov>
- [20] <https://commons.esipfed.org/taxonomy/term/416>
- [21] <https://commons.esipfed.org/taxonomy/term/606>
- [22] <https://commons.esipfed.org/taxonomy/term/489>
- [23] <https://commons.esipfed.org/taxonomy/term/607>