

A Semantically Enabled Metadata Repository for Solar Irradiance Data Products

Anne Wilson, Michael Cox, Douglas M Lindholm, Irfan Nadiadi, Tyler Traver

Laboratory for Atmospheric and Space Physics (LASP)

AGU Fall 2014 Meeting



The Problem

Information about LASP datasets was managed in an ad hoc fashion, at times resulting in inconsistent and dated information being provided to the public. The information was embedded in HTML, thus suitable only for use by a human reader.

The Need

A repository, a Single Source of Truth, was needed for current, vetted information about solar irradiance datasets. In order to participate in the semantic web, the information should be semantically enabled and machine readable, which would improve the discovery and understanding of the data.

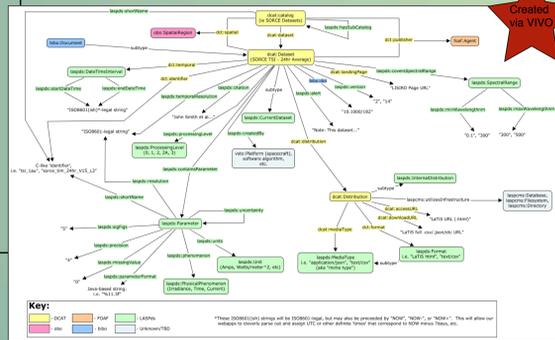
The Approach

Use VIVO to build a semantically enabled metadata repository. Start with use case of LASP Interactive Solar Irradiance Data Center (LISIRD). Also, be able to create metadata records in both SPASE and ISO 19115 format.

SPASE is a metadata model widely used in the heliophysics community. ISO 19115 is an international metadata schema.



The ontologies that come with VIVO.



The laspds ontology, integrated with other VIVO ontologies.

Definitions

ontology - a formal definition of concepts (classes), properties, and relationships for a particular domain

triple - a 3-tuple, interpreted as <subject, predicate, object>
- allows definition of relationships
- from relationships, new knowledge can be inferred, knowledge not explicit in the database

triplestore - a database of instances of triples

SPARQL - a query language designed for triplestores

An ontology was needed to describe the domain of solar irradiance data products. We considered SWEET and integrated the VSTO ontology, but neither fully met our needs. Thus, using VIVO, we created the laspds ontology, an extension of the W3C DCAT ontology, and integrated it into the VIVO suite.

SWEET - Semantic Web for Earth and Environmental Terminology
VSTO - Virtual Solar Terrestrial Observatory

Using VIVO, the database was populated with instances of relationships (triples) defined via the ontology. For example, **SORCE Total Solar Irradiance (TSI)**:

- is an instance of a DCAT Dataset,
- was 'createdBy' the SORCE spacecraft,
- has a 'temporal' property called 'dateTimeInterval', which
 - has instances of a 'startDateTime' with the value '10/31/2004' and 'endDateType' of '8/22/2013',
- has a concept called a 'spectralRange' with a minimum and maximum wavelength,
- has 2 parameters, TSI_1AU and TSI_trueearth,
 - each of which, in turn, have units, uncertainties, etc.

