Geophysical Research Abstracts
Vol. 16, EGU2014-11212, 2014
EGU General Assembly 2014

Publishing NASA Metadata as Linked Open Data for Semantic Mashups

Brian Wilson, Gerald Manipon, and Hook Hua
JET PROPULSION LABORATORY, Atmospheric Remote Sensing, PASADENA, United States (Brian.Wilson@jpl.nasa.gov)

Data providers are now publishing more metadata in more interoperable forms, e.g. Atom or RSS 'casts', as Linked Open Data (LOD), or as ISO Metadata records. A major effort on the part of the NASA's Earth Science Data and Information System (ESDIS) project is the aggregation of metadata that enables greater data interoperability among scientific data sets regardless of source or application. Both the Earth Observing System (EOS) ClearingHOuse (ECHO) and the Global Change Master Directory (GCMD) repositories contain metadata records for NASA (and other) datasets and provided services. These records contain typical fields for each dataset (or software service) such as the source, creation date, cognizant institution, related access URL's, and domain and variable keywords to enable discovery.

Under a NASA ACCESS grant, we demonstrated how to publish the ECHO and GCMD dataset and services metadata as LOD in the RDF format. Both sets of metadata are now queryable at SPARQL endpoints and available for integration into "semantic mashups" in the browser. It is straightforward to reformat sets of XML metadata, including ISO, into simple RDF and then later refine and improve the RDF predicates by reusing known namespaces such as Dublin core, georss, etc. All scientific metadata should be part of the LOD world.

In addition, we developed an "instant" drill-down and browse interface that provides faceted navigation so that the user can discover and explore the 25,000 datasets and 3000 services. The available facets and the free-text search box appear in the left panel, and the instantly updated results for the dataset search appear in the right panel. The user can constrain the value of a metadata facet simply by clicking on a word (or phrase) in the "word cloud" of values for each facet. The display section for each dataset includes the important metadata fields, a full description of the dataset, potentially some related URL's, and a "search" button that points to an OpenSearch GUI that is pre-configured to search for granules within the dataset.

We will present our experiences with converting NASA metadata into LOD, discuss the challenges, illustrate some of the enabled mashups, and demonstrate the latest version of the "instant browse" interface for navigating multiple metadata collections.