

# 2016 SWWG Meetings

January 14, 2016

## Using Parliament for a GeoSPARQL query

On January 14, we continued to experiment with an internal installation of Parliament that provides a GeoSPARQL endpoint.

---

February 11, 2016

## Using a Vocabulary Service for Secondary Metadata Validation

Peter Schweitzer has modified an online metadata validator to use the vocabulary services to check the accuracy of keywords in metadata fields. We will get a tour and discussed this outcome of the SWWG vocabulary services project.

---

March 10, 2016

## The NERC Vocabulary Services at British Oceanographic Data Centre

Rob Thomas and Alexandra Kokkinaki gave us an overview of the vocabulary services at NERC ([http://www.bodc.ac.uk/products/web\\_services/vocab/](http://www.bodc.ac.uk/products/web_services/vocab/)) and answered questions.

---

June 9, 2016

## Discussion about progress on linked data and controlled vocabularies

Dalia reported that her group is moving toward a dedicated server and services for linked data, which she will be able to demonstrate soon. They will be using Fuseki, which does not support GeoSPARQL, but the triples they publish can be processed if a GeoSPARQL server links to them.

Unfortunately, word from Dave is that our experimental Parliament instance can no longer be maintained.

Discussion continued about the value of a bureau-wide effort to publish linked data. No solution was found, but we will pursue options within our offices and programs.

Leslie asked how the USGS Thesaurus is managed, and Peter gave us a tour of [internal.usgs.gov/thesaurus](http://internal.usgs.gov/thesaurus) and [www2.usgs.gov/science](http://www2.usgs.gov/science).

Fran, Peter, and Alan will present last year's controlled vocabulary project to the July 13 CDI WebEx.

---

July 14, 2016

## Discussion about progress on linked data and controlled vocabularies

Dalia reported that her group is making progress on a triple store, a mapping interface, and a demonstration of the advantages of semantic technology, but there is nothing to show us yet.

We had a long discussion about where, organizationally and on the network, USGS can (1) build a linked open data service and also (2) provide a long-term home for our controlled vocabulary server. There is not a clear process for providing central organizational services that can be used by the whole bureau.

We agree that we need to write and analyze use cases for a USGS controlled vocabulary registry.

---

Sept. 8, 2016 (note that we skipped August because of multiple scheduling complications)

**Check-in: individual members discussed semantic web efforts in our day jobs.**

On-going themes: desire for semantic versions of gazetteers that provide geospatial relationships between places. Value of controlled vocabulary keywords in scientific metadata that identify observation or measurements that are included in the data. How can we effectively publish linked data in the USGS this year?

---

**Oct. 13, 2016**

**Discussion of Vocabulary Server Governance**

After clarifying what we mean by **vocabulary server governance**, as opposed to **vocabulary governance** and **other good features and functions** of servers and vocabularies, we focused on the challenges and concerns with vocabulary server governance in USGS. We were in agreement that USGS vocabularies need a home where they will be a recognized bureau-wide resource, well-known and persistent. CSAS-L seems to be the right part of USGS to host vocabulary servers, combining technical capabilities with the mission of the library and librarians' understanding of the value of controlled vocabularies. CSAS-L also already has custody of two large USGS controlled vocabularies: ITIS and the Biocomplexity Thesaurus. We draft some notes that clarify our thinking about governance requirements and invite someone at CSAS-L to help us think this through.

Notes were taken on the google spreadsheet at <https://docs.google.com/a/doi.gov/spreadsheets/d/1qf80O0gPGJYVg2rQJtdTEDZK9OmiJUz8931vbYCledU/edit?usp=sharing>

---

**Nov. 10, 2016**

**Brainstorming about Contributions to CDI Workshop**

Our skit was a hit last time. This time we have a great title that lends itself to a nice photo-shop image: Trouble with Triples. Brainstorming for the May 2017 workshop stayed pretty close to some kind of game involving choice of controlled vocabulary index terms, perhaps two teams competing like in Family Feud. The details are still undetermined.

---

**Dec. 8, 2016**

**ITIS has vocabulary services. How do they handle vocabulary governance?**

Our guest was David Mitchell of USGS and the Smithsonian Institution.

**ITIS**, the Integrated Taxonomic Information System, provides a web interface and web services for a database of authoritative taxonomic information on plants, animals, fungi, and microbes of North America and the world. (Fossils are not covered unless they fit into the modern classification system.) Much of the governance of the terms is provided by the nomenclature codes that cover taxon names. Within ITIS, the valid name of a taxon, and the associated taxonomic serial number (TSN) are permanent, and the TSN is unique, which is useful for governance. The ITIS online system is a partnership between the Smithsonian and USGS.

David's slides will be attached here.



SemanticWebWork...\_ITISIntro.pptx

#### ITIS Links

Here are some links to services and features discussed at the meeting:

Traditional Web Services [https://www.itis.gov/ws\\_description.html](https://www.itis.gov/ws_description.html)

Solr Services [https://www.itis.gov/solr\\_documentation.html](https://www.itis.gov/solr_documentation.html)

Taxmatch <https://www.itis.gov/taxmatch.html>

TSN FAQ [https://www.itis.gov/pdf/faq\\_itis\\_tsn.pdf](https://www.itis.gov/pdf/faq_itis_tsn.pdf)