CDI Virtual Training in Using Controlled Vocabularies

Did you miss the Controlled Vocabularies Presentation and Demo? Here's a summary! (Caution: CDI Coordinator and author of this post is not a metadata /semantic web expert.)

Summary

The CDI Semantic Web Working Group has been working to improve the recall and precision of searches in USGS data catalogs by making controlled vocabularies conveniently useful. Work began with a CDI project in 2014 and 2015 that developed use cases and prototypes. On July 13 the team demonstrated vocabulary services and several metadata tools that use vocabulary services. They also talked about the vision, objectives, and next steps of the Working Group's Controlled Vocabulary Manifesto. (Fran Lightsom, Peter Schweitzer, and Alan Allwardt)

The Controlled Vocabulary Manifesto

The **Controlled Vocabulary Manifesto** expresses a vision for the future where

> *People using USGS data catalogs will be confident that their search results are both comprehensive and focused, with good recall (nothing relevant missed) and good precision (nothing irrelevant included).*
>
> *USGS will efficiently use controlled vocabularies available through services to achieve this result.*

The **manifesto strategy** achieves these objectives.

1. Select controlled vocabularies and make them available through services.
2. Include controlled vocabulary terms in all USGS metadata.
3. Update USGS data catalog interfaces to take advantage of controlled vocabularies.
4. Educate and motivate the USGS workforce.

Five points to take away

1. A **controlled vocabulary is** a set of terms that a community agrees to use, with clear definitions and consistent spellings.
2. Different vocabularies have been developed with different audiences and foci in mind, therefore we (users and web applications) **may need to use multiple vocabularies** to adequately annotate a resource. Browse and search the vocabularies available through USGS at the USGS Science Topics page: https://www2.usgs.gov/science/tab-term.html
3. The Semantic Web Working Group has developed **new web services for these vocabularies**. (Web services "support interoperable machine-to-machine interaction over a network" according to the W3C and wikipedia.) Documentation of the services is at https://www2.usgs.gov/science /services.html.
4. Applications that USGS users employ, such as Metadata Wizard, use the web services, making consistent keywords for search, and bringing us closer to **the vision of the Controlled Vocabulary Manifesto** where USGS data catalog searches are comprehensive, focused, have good recall (nothing relevant missed) and good precision (nothing irrelevant included).
5. In general, a normal user will interact with the USGS Controlled Vocabularies through the **interface of an application** (e.g., one for creating metadata), which will use the web services to provide a way to choose terms and related terms.

More things you can do

- See the slides and recording here.
- Read further Q&A coming from this session.
- Join the Semantic Web Working Group if you are interested in helping the cause.
- Register for the Reviewing Metadata Virtual Training.

--
See more CDI Blog posts