

# SWWG 2014 Proposal Project

## Use of Controlled Vocabularies in USGS Information Applications (CDI14-LF336)

### Abstract:

Large online data catalogs use controlled vocabularies to categorize datasets in ways that allow end users to sort and select data matching their needs. Although the semantic web community has developed methods of expressing controlled vocabularies, we lack firm standards specifying how web services should provide vocabularies to end users and applications. The eventual goal of this project is to build functional services using the USGS Thesaurus and Biocomplexity Thesaurus as examples, but use cases must be developed and analyzed first to design services that work effectively with existing applications and processes. From prior experience we have learned that it is essential for the use case team to include a mix of skills and points of view, along with a skilled facilitator, and that an initial face-to-face meeting is necessary to reduce the risk of miscommunication and divergent goals. The initial product of this work will be a white paper detailing the requirements for implementing USGS vocabulary services or contributing USGS thesauri to externally managed vocabulary services. Following use case development and analysis, our goal is to expose the USGS Thesaurus and Biocomplexity Thesaurus for wider use as authoritative controlled vocabularies for describing and categorizing scientific information.

### Products:

[Use cases](#), analyzed to develop requirements for vocabulary services.

[Improvements](#) in USGS Vocabulary Services and the Metadata Wizard, a tool for creating metadata.

The [Controlled Vocabulary Manifesto](#), a plan for improving the discovery of USGS data by competent and consistent use of controlled vocabularies.