



A SKOS based thesaurus of the Geological Survey of Austria exposed through an Open Linked Data Web-Service

Marcus Ebner, Martin Schiegl, Werner Stöckl, Ralf Schuster, and Christoph Janda
Geological Survey of Austria, Vienna, Austria (marcus.ebner@geologie.ac.at)

The Geological Survey of Austria is a subordinate agency of the Austrian Federal Ministry for Science and Research and as such the premier advisory body for the Austrian Government for geosciences. As a public authority that maintains spatial data sets it is legally obligated by the INSPIRE directive to provide data that fall under this directive (geology, mineral resources and natural risk zones) to the European commission in an harmonized and technically interoperable way.

The foundation of such a semantic harmonization is the development of a controlled vocabulary that covers the main thematic aspects of the datasets. The controlled vocabularies will be provided in the form of specialized thesauri in terms of information science (i.e. a lightweight ontology for information retrieval). For the encoding of the thesauri we build upon the W3C standard SKOS (=Simple Knowledge Organisation System), a thesaurus specification for the semantic web, which is itself based on the Resource Description Framework (RDF) and RDF Schema. For the development of these thesauri we use the commercial software PoolParty, which is a tool specially build to manage multilingual thesaurus systems and a SKOS editor. For the concepts in the thesauri preferred labels and definitions will be provided in both German and English. In addition to SKOS and RDF(S) encoding standards PoolParty makes use of Dublin Core (DC) for resource description.

The corporate thesauri of the Austrian Geological Survey are exposed via a web-service that is conformant with the linked data principles (<http://www.w3.org/DesignIssues/LinkedData.html>). This web-service (here called a “linked data frontend”) gives access to a (1) RDF/HTML representation of the concepts/resources via its http URIs (2) a WIKI-page for simple browsing and navigation within a thesaurus and (3) a full-fledged SPARQ-Endpoint to query the thesaurus.

Using the linked open data principles suggested by Tim Berners-Lee (on the web, open licenses, non proprietary formats, RDF standards etc.) when exposing data on the web, especially for controlled vocabulary allows knowledge retrieval and (simple) inferencing in an interoperable way for both humans and machines.