

Stratigraphy in Web 3.0 – Advantages of a controlled vocabulary for stratigraphic units

Hörfarter, Christine

Department of Geoinformation, Geological Survey of Austria, Neulinggasse 38, A-1030 Vienna, Austria.

The digital processing of scientific knowledge on geological units and entities can be a challenge. A relational database system quickly becomes complex and poorly performing if poly-hierarchies, synonyms, homonyms, errata, alternative spellings, or content vagueness are to be considered. A more feasible option to address these challenges is to create and use a controlled vocabulary in the form of an online thesaurus. Controlled vocabularies provide the potential to clarify expert knowledge and terminology in the form of thematic vocabulary concepts (terms) at a scientific level and to use them in datasets to code data. By using persistent identifiers instead of labels, which are often impermanent and ambiguous, these concepts are assigned to a URI and can be used to encode data and serve as a unique resource in data visualization tools, for example. In the Thesaurus of the Geological Survey (GBA Thesaurus – <https://thesaurus.geolba.ac.at/>), concepts from dataset publications on geological units have been processed in such a controlled vocabulary according to the SKOS standard (Simple Knowledge Organization System) for more than ten years now. The GBA Thesaurus results in a knowledge database of bibliographically referenced concepts that have been developed through scientific publications. Within the GBA Thesaurus it is already possible to establish sharp and less sharp associative and (poly)hierarchical relations between different terms and vocabulary schemas. In addition, linking can be made to existing content and information systems on the web, such as e.g. the CGI vocabularies (<http://geosciml.org/resource/>) e.g. and soon the Lithostratigraphic Lexicon of the German Stratigraphic Society. With the GBA Thesaurus and the linking of concepts, we are thus providing a cornerstone for the cross-border harmonization of geological units. With this use of Linked Data technology, we are already part of the world of the Semantic Web – the Web 3.0.