

The HIKE European Fault Data Base: The interplay of structured data with Linked Data

Hintersberger, Esther; HIKE Team

Department of Hard Rock Geology, Geological Survey of Austria, Neulinggasse 38, A-1030 Vienna, Austria.

Knowledge of the movement along major fault systems is the key for understanding the tectonic evolution in such a complex environment as Europe. Besides being a typical product of geological mapping and a necessary part of kinematic models, faults are also prominent features defining resources (e.g. minerals, thermal conduits) and/or inducing potential hazard to subsurface drilling, injection and extraction activities (e.g. conventional hydrocarbon extraction in Groningen, Netherlands). Furthermore, the knowledge on faults, their subsurface geometry and deformation history has also increased in complexity, eventually showing the limits of printed maps for adequately representing the current state of knowledge. Therefore, collecting and structuring the available information on faults and presenting it in a harmonized and generally accessible way across national borders is a necessary challenge for geologists and geodata providers. Previous efforts focused on seismogenic fault (e.g., GEM Global Active Faults Database, SHARE European Database of Seismogenic Faults, USGS Quaternary faults database). Here, we present the HIKE European fault database (EFDB), which had the objective to provide a consistent and uniform repository of all types of faults and their characteristics across Europe. In order to capture the variety of geological environments and the highly varying levels of available data, the EFDB combines three different aspects of a fault object: fault geometry, kinematic attributes and a linked semantic vocabulary where non-structured information can be stored. In addition to structured information stored in attribute tables, the EFDB contains a hierarchical classification scheme, which sorts faults and shear zones into groups of local, regional or transregional relevance through a semantic vocabulary of named faults. The vocabulary that has been generated under the principles of Linked Data, which allows storing unstructured information such as geographic description, detailed investigation history, debated theories etc., but also creates a network beyond the actual fault database by including links to other existing fault databases and additional information, e.g. Wikipedia or other semantic vocabularies. We show that a balanced mix of structured information stored in attribute tables and an associated semantic vocabulary provides geologists the opportunity to share complex geodynamic and kinematic information. The European fault database was developed during the Horizon 2020 GeoERA projekt HIKE and contains data from Geological Survey Organizations in Austria (GBA), the Netherlands (TNO), Germany (BGR, LfU, LAGB, LBGR), Belgium (RBINS-GSB), Iceland (ISOR), Denmark (GEUS), Poland (PIG-PIB), Lithuania (LGT), Italy (ISPRA), France (BRGM), Ukraine (GEOINFORM), Portugal (LNEG), Slovenia (GeoZS), Albania (AGS) and 43 the Pannonian Basin Area (MBFSZ). The GeoERA HIKE project has received funding from the European Union's Horizon 2020 research and innovation programme under agreement No. 731166. The HIKE European Fault and the related fault inventory vocabulary is openly accessible via <https://geoera.eu/projects/hike10/faultdatabase/>.