



Semantic harmonization of geoscientific data using Linked Data and project specific vocabularies

M. Schiegl

GBA - Geological Survey of Austria

Chat time: Monday, 4 May 2020, 08:30–10:15



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



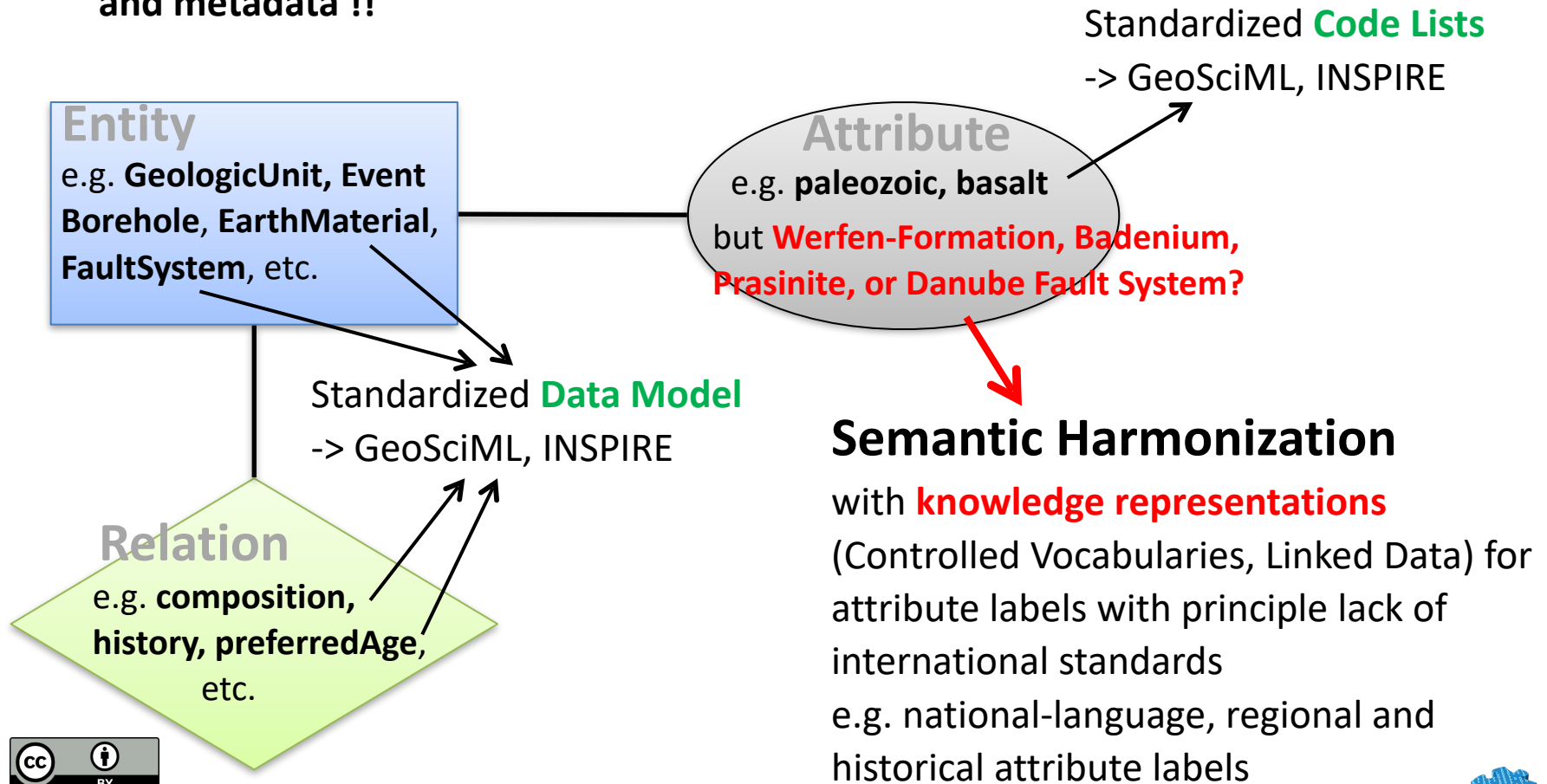
What is Semantic Harmonization?

- making datasets and their attribute **data consistent** and **compatible**, relating to the **meaning in language** and logic
- **Examples:** *Werfen Formation (Italy) – Werfener Schichten (Austria) (=quasi synonym), or Zementmergelserie (=polysem), silty sandy limestone (=combination term), concept definitions of Ice Ages or the Paratethys Stage System, named groundwater bodies or fault systems*
- **Clarification of the meaning** -> bibliography, relations, links,..
- Linked context (knowledge base) -> semantic interoperability
- RDF triple store (Linked Data) + GIS/SQL database (relational)
- But use international standards like **INSPIRE codelists** when there is no need to clarify scientific terms!



harmonizing project data

and metadata !!



GeoERA Project, part GIP-P WP4

GeoERA **project vocabularies** are collections of (linguistic labeled) scientific concepts, stored in named graphs and concept schemes according to different modeling approaches. They are described in SKOS/RDF plus common metadata and properties with focus on scientific reusability.

GeoERA vocabulary concepts can be applied to following types of concepts:

1. Terms describing geoscientific feature types or properties (schema-level)
2. Terms named by classifications or prototype theory's
3. Combination terms for geologic features (like map legend items)
4. Located and named occurrences of geoscientific types or properties (instance-level)

These kind of vocabularies can be used both to harmonize and consequentially connect datasets and to extend standards like INSPIRE codelists.

GeoERA Project: <https://geoera.eu>

WP4 documentation: <https://github.com/GeoEra-GIP/WP4-Semantics>

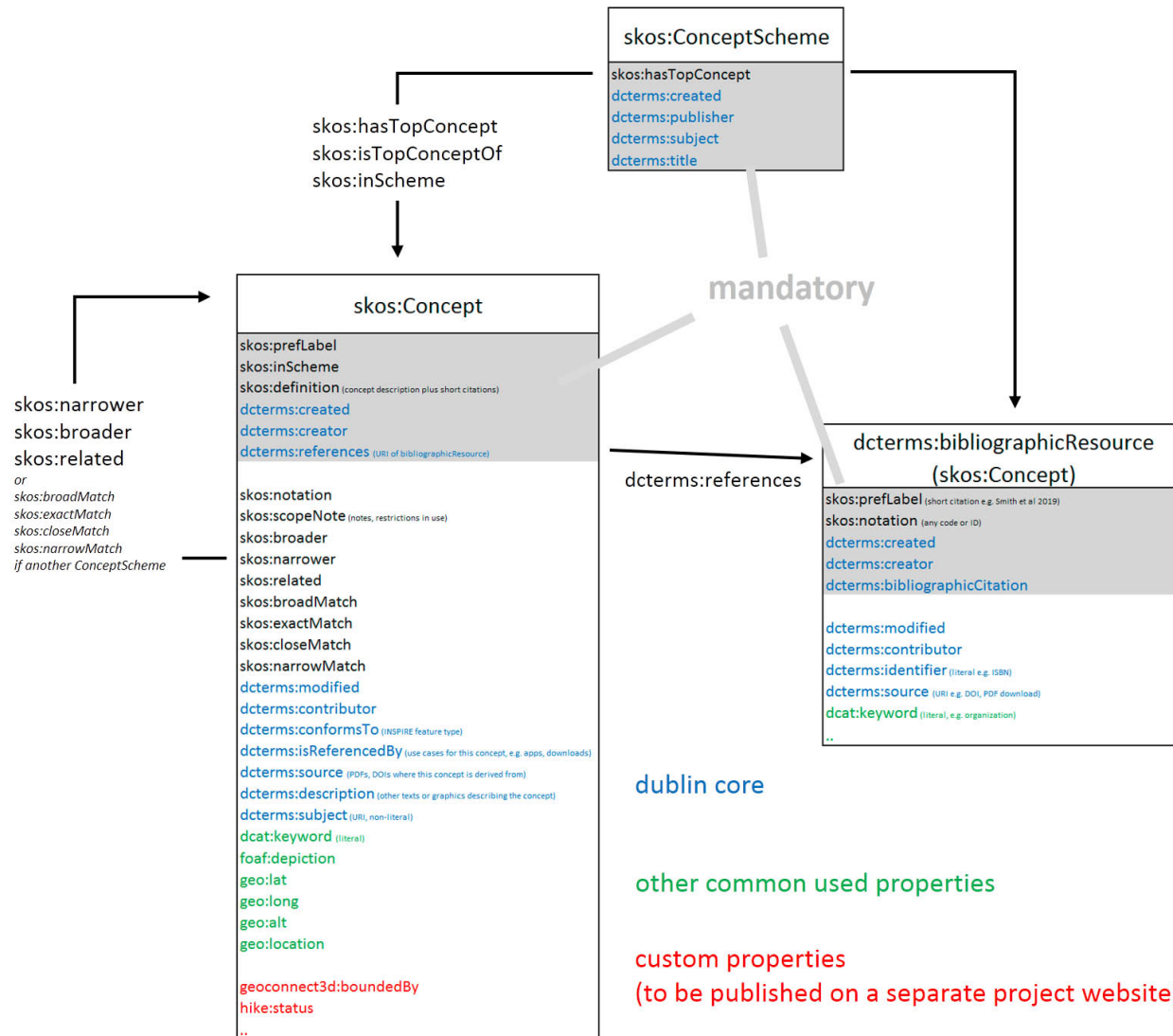
WP4 report: <https://geoera.eu/wp-content/uploads/2019/11/D4.3-GeoERA-Project-Vocabularies.pdf>



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



Project Vocabularies – data model (GBA/SGU)



Project Vocabularies – Excel template (GBA)

scheme	concept	concept	concept	concept	concept
Tectonic Boundaries Classification Scheme					
	Tectonic Boundary				
		Fault Set			
		Nappe Boundary			
		Large-scale Fault System			
			Fault System		
				Fault	
					Subfault
				Shear zone	
		Subfault System			

↑
Concept scheme label

↑
Top concept - Uppermost concept hierarchy (level 1)

↑
Narrower concept label – next hierarchy (level 2)

↑
Narrower concept label – next hierarchy (level 3)

↑
Narrower concept label – next hierarchy (level 4)

↑
Narrower concept label – next hierarchy (level 5)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



Editing system

The screenshot displays the GeoERA editing system interface. The top navigation bar includes 'PROJECT', 'CORPORA', 'TOOLS', 'OPTIONS', and 'ADVANCED'. A search bar shows 'de' and a notification indicates 'Last automatic snapshot at 10:07 - 06.03.2019'. The left sidebar shows a hierarchical tree of 'Geologische Strukturen' with categories like 'Klassifikation Tektonischer Grenzflächen (1)', 'Objekte Tektonischer Grenzflächen (1)', and 'Tektonische Grenzfläche in Österreich (60)'. The main content area shows the 'Engadin-Inntal-Störungssystem' concept, with a URL <http://resource.eoolba.ac.at/structure/170>. The interface includes tabs for 'Details', 'Notes', 'Documents', 'Linked Data', 'Triples', 'Visualization', 'Quality Report', and 'History'. Below these are sections for 'SKOS', 'Relations', 'Preferred Label', 'Alternative Labels', 'Hidden Labels', and 'Notation'. The 'Preferred Label' section shows 'Engadin-Inntal-Störungssystem' (de) and 'Engadin-Inntal Fault System' (en). The 'Alternative Labels' section shows 'Engadin-Teilstörungssystem' (de) and 'Inntal-Teilstörungssystem' (en). The 'Hidden Labels' section shows 'Engadin-Teilstörungssystem' (de) and 'Inntal-Teilstörungssystem' (en). The 'Notation' section shows 'Engadin-Teilstörungssystem' (de) and 'Inntal-Teilstörungssystem' (en). There are also buttons for 'Add to Collection', 'Delete Concept', 'Add to Blacklist', and 'Linguistics'.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166

Multifunctional representation

WMS Data:

OBJECTID	8
NummerderStörung	10049
SHAPE	Polyline
FEATURE_ID	101
Klassifizierung	Störung
Bezeichnung	Ebner-Störung
synonymeBezeichnung	Ebner Thrust
Großstörungenssystem	Null
Störungssystem	Null
Teilstörungssystem	NKA-Störungset (E-W)
Störungszw.Scherzone	Ebner-Störung
Kurzreferenz	Eisbacher und Brandtner, 1995, 1996
THESURL	http://resource.geolba.ac.at/structure/118

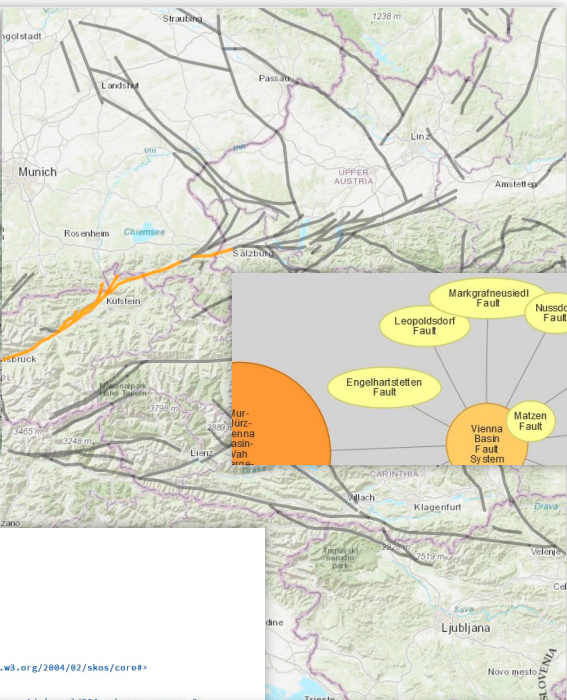
LinkedData:
Ebner-Störung

ist eine E-W-streichende spröde Störung mit überschneidender Kinematik in den westlichen Nordosten Kalkalpen (NKA). (Eisbacher & Brandtner, 1995, 1996)

in **NKA-E-W-Störungset**

Eisbacher, G. & Brandtner, R. (1995). Role of high-angle faults during tectonic contraction, Innal Thrust Sheet, Northern Calcareous Alps, Western Austria. - *Geologisch-Paläontologische Mitteilungen Innsbruck* 25, 385-406

Eisbacher, G. & Brandtner, R. (1996). Superposed fold-thrust structures and high-angle faults, east Austrienne Calcareous Alps, Austria. - *Geologische Helvetiae* 86, 85-111



Geological Survey of Austria

Engadin-Inntal Fault System

URI: <http://resource.geolba.ac.at/structure/170> → RDF download

Engadin-Inntal-Störungssystem Engadin-Inntal Fault System

GBA Status: **official use**

This ca. 225 km long, approximately SW-NE trending Fault System extends from Bregaglia-Tal (north of the Bergell massive), following the Inn Valley to Nauders, Pfunds, Imst and Telfs. Here, it is displaced right-laterally by the Telfs Subfault System. The eastern part starts in Innsbruck to Wörgl and Kufstein. It includes Subfault Systems along the Engadin and Inn Valley and parallel trending structures such as the Schling, Gallo and Arlui shear zones. It shows a long-lasting history and was initiated as ductile nappe boundaries and shear zones with E-dipping top E-dipping thrusting and top ESE-directed normal faults during Cretaceous and Paleogene times. In the Oligocene it was reactivated along shear zones with top WNW kinematics (e.g. Engadin and Schling shear zones). Left-lateral brittle strike-slip faults are active from the upper Oligocene to Miocene, indicated by sediment deposits in the Inn Valley (Froitzheim et al., 1997; Ortner et al., 2006). Displacement rates can be summarized, reaching around 17 km normal faulting and 1 km km lateral displacement. (Froitzheim et al., 1997; Linzer et al., 2002 and references therein)

Ortner, H.-G., Decker, K., Persson, H., Dell'Mour, R. & Frisch, W. (2002): *Balancing lateral orogenic float of the Eastern Alps. In: Tectonophysics* 354, Nr. 3-4, S. 211-237. [Catalog]

Froitzheim, N., Corti, P. & Van Dalen, M. (1997): *Late Cretaceous, synorogenic, low-angle normal faulting along the Schling fault (Switzerland, Italy, Austria) and its significance for the tectonics of the Eastern Alps. In: Tectonophysics* 280, Nr. 3-4, S. 267-293. [PDF]

Ortner, H., Reiter, F. & Brandner, R. (2006): *Kinematics of the Inntal shear zone-sub-Taurn ramp fault system and the interpretation of the TRANSALP seismic section, Eastern Alps. In: Tectonophysics* 424, Nr. 1-2, S. 1-15. [PDF]

thesaurus

Search for...

Applications

- Network diagram
- Database tables
- Structure Viewer

Geologic Structures (subject)

The Theme Geologic Structures includes linear and planar predominantly deformation structures in geologic maps. Shear sense indicators and fold structures are also covered by this theme.

Concept relations

broader	Engadin-Inntal-Innsbruck-Sa
narrower	Inntal Subfault System Schling Subfault System Engadin Subfault System

S	L
http://resource.geolba.ac.at/structure/61	"Tectonic Boundary in Austria"@en
http://resource.geolba.ac.at/structure/102	"NCA NE-SW Conjugated Fault Set"@en
http://resource.geolba.ac.at/structure/116	"NCA NNE-SSW Conjugated Fault Set"@en
http://resource.geolba.ac.at/structure/111	"NCA E-W Conjugated Fault Set"@en
http://resource.geolba.ac.at/structure/120	"Königsee-Lammertal-Trunsee Subfault System"@en
http://resource.geolba.ac.at/structure/121	"Tauern Window Subfault System"@en
http://resource.geolba.ac.at/structure/124	"Telfs Subfault System"@en
http://resource.geolba.ac.at/structure/125	"Trojeralm Conjugated Fault Set"@en
http://resource.geolba.ac.at/structure/126	"Landshut-Neutötting Fault System"@en
http://resource.geolba.ac.at/structure/127	"Deferegggen-Antholz-Vals Fault System"@en
http://resource.geolba.ac.at/structure/128	"Danube Fault System"@en
http://resource.geolba.ac.at/structure/130	"Drautal-Zwischenbergen-Wollatratn Fault System"@en
http://resource.geolba.ac.at/structure/133	"Görschtal Fault System"@en
http://resource.geolba.ac.at/structure/137	"Hochstihl-Gegendtal Fault System"@en
http://resource.geolba.ac.at/structure/139	"Iseltal Fault System"@en

```

1 <!DOCTYPE html>
2 <html>
3 <head>
4 <meta charset="UTF-8">
5 <title>SPARQL queries with Javascript</title>
6 </head>
7 <body>
8 <script>
9 // Sparql query
10 let query = encodeURIComponent('PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
11 SELECT *
12 WHERE {
13 <http://resource.geolba.ac.at/mineral/284> skos:narrower* ?s .
14 ?s skos:prefLabel ?L .
15 }');
16 fetch('http://resource.geolba.ac.at/PoolP...
17 .then(res => res.json())
18 .then(jsonData => {
19   jsonData.results.bindings.forEach(function(binding) {
20     // do something
21   });
22 });
23 </script>
24 </body>
25 </html>

```

SPARQL Endpoint

```

PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
SELECT *
WHERE {
<http://resource.geolba.ac.at/structure/61> skos:narrower* ?s .
?s skos:prefLabel ?L . FILTER (lang(?L)="en")
}

```



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



GeoERA vocabularies workflow in comparison

	Vocabulary draft	GeoERA Project Vocabularies	European Geoscience Register	INSPIRE register federation	INSPIRE codelists
Governance	project team	GeoERA project team	EGS expert groups	JRC INSPIRE registry team	JRC INSPIRE registry team
Publisher	project team	GIP-P WP4	BRGM	Relevant organization (BRGM)	JRC INSPIRE registry team
basic ontology	none	SKOS (https://www.w3.org/TR/skos-reference/)	REG (purl.org/linked-data/registry)	none	none
Owner	project team	GeoERA project team	EGS, EPOS?	Relevant organization	European Union
Entities	Scientific terms, codes, texts, links, citations..	concepts, concept schemes, links, codes, metadata, .	codes, ontology concepts, complete ontologies, coordinate reference systems, units of measure, spatial objects, organizations, licenses, metadata etc.	Codelists, enumerations	Codelists, enumerations
Data transfer	Excel	trig, rdf/xml, ttl, Sparql (test API)	ttl, rdf/xml, json-ld, csv, Sparql	Links to registers	rdf/xml, json, atom, csv
Data standards	none	Semantic web standards	Semantic web standards	INSPIRE standards	INSPIRE standards
Linked Data (https://5stardata.info/en/)	☆☆	☆☆☆☆☆☆	☆☆☆☆☆☆(☆)	☆☆☆☆☆☆	☆☆☆☆☆☆
Semantic relations	draft	pivotal	partially	partially	only parent/narrower
cross-linked vocabularies	draft	pivotal	partially	uncommon	no
scope	elaborate scientific terms, or assemble codelists	Clarification of the meaning of scientific terms, providing context, knowledge base, bibliographic references	Register codes, concepts, data and other entities	Extending INSPIRE codelists	Providing standardized codes
explanations	https://github.com/GeoEra-GIP/WP4-Semantics/tree/master/Project%20Vocabularies/templates	https://geoera.eu/wp-content/uploads/2019/11/D4.3-GeoERA-Project-Vocabularies.pdf	https://github.com/UKGovLD/registry-core/wiki/Principles-and-concepts	https://inspire.ec.europa.eu/id/document/tg/registers-and-register-federation	http://inspire.ec.europa.eu/codelist
Project Vocabularies	create concepts >> Project Vocabulary				
codelists	create codes	>> skip PV >>	publish codes		
INSPIRE codelists			extend INSPIRE codelists		
				use INSPIRE codelists plus official extensions	

GeoERA project vocabularies

- **Harmonization** effects by using a **common project vocabulary**
- **Online availability** for web applications, EU **multilingualism**, cross-linked bibliographic references
- **meaning of concepts defined** by provided context (eg bibliography)
- **scientific** traceability of information (**provenance**)
- Even **weakly structured** data/knowledge can be processed
- **Sustainable** knowledge management through **reusability** and expandability in the next project
- **interdisciplinary** usable, because of crosslinking
- **extend**/complete the harmonization efforts of **INSPIRE and GeoSciML**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



Mur-Mürz-Vienna Basin-Vah Large-scale Fault System

preferred Label

URI: <http://resource.geolba.ac.at/structure/186> ⇒ PDF download **URI**

Mur-Mürz-Vienna Basin-Vah Large-scale Fault System **en**
 Mur-Mürz-Wien-Becken-Vah-Großstörungssystem **de**
 GBA Status: **official term**

other Labels

Descriptions

From S to N, this ca. 400 km long Large-scale Fault System includes the Mur-Mürz Fault System, Vienna Basin Fault System, and the Vah Fault System (Decker et al., 2005) and seems to be kinematically linked to thrust faults in the outer Carpathians in Polish Galicia. This large-scale fault system consists mostly of ENE-WSW trending left-lateral strike-slip and NNE-SSW trending normal faulting. Onset of faulting along the entire large-scale fault system started in the Miocene as consequence of eastward extrusion of the Central Eastern Alps (Ratschbacher et al. 1991, Linzer et al., 2002). In total, Miocene offset in the Mur-Mürz-Vienna Basin area has been estimated to up to 40 km (Linzer et al., 2002). Earthquakes are observed for the first time along the so-called Mur-Mürz-Vienna Basin-Zilina seismogenic zone by Gutdeusch & Aric (1988) proving recent activity of the large-scale fault system.

— Gutdeusch, R. & Aric, K. (1988): *Seismicity and neotectonics of the East Alpine-Carpathian and Pannonian area. The Pannonian Basin. AAPG memoir 45, 183-194.* - [PDF]
 — Linzer, H.-G., Decker, K., Persson, H., Dell'Mour, R. & Frisch, W. (2002): *Balancing lateral orogenic float of the Eastern Alps.* - In: *Tectonophysics 354, Nr. 3-4, S. 211-237.* - [Catalog]
 — Ratschbacher, L., Behrmann, J.H. & Pahr, A. (1990): *Penninic windows at the eastern end of the Alps and their relation to the intra-Carpathian basins.* - *Tectonophysics 172, 1-2, 91-105.* - [PDF] - [Catalog]
 — Decker, K., Persson, H. & Hinsch, R. (2005): *Active tectonics and Quaternary basin formation along the Vienna Basin Transform fault.* - In: *Quaternary Science Reviews 24, Nr. 3-4, S. 307-322.* - [Catalog]

short citation

references

Concept relations

broader	Tectonic Boundaries in Austria
narrower	Mur-Mürz Fault System Vienna Basin Fault System

relations

plus links to:
 media, images, libraries
 data-downloads, geonames
 applications, wikipedia articles

From **scientific term**
 to **scientific concept**

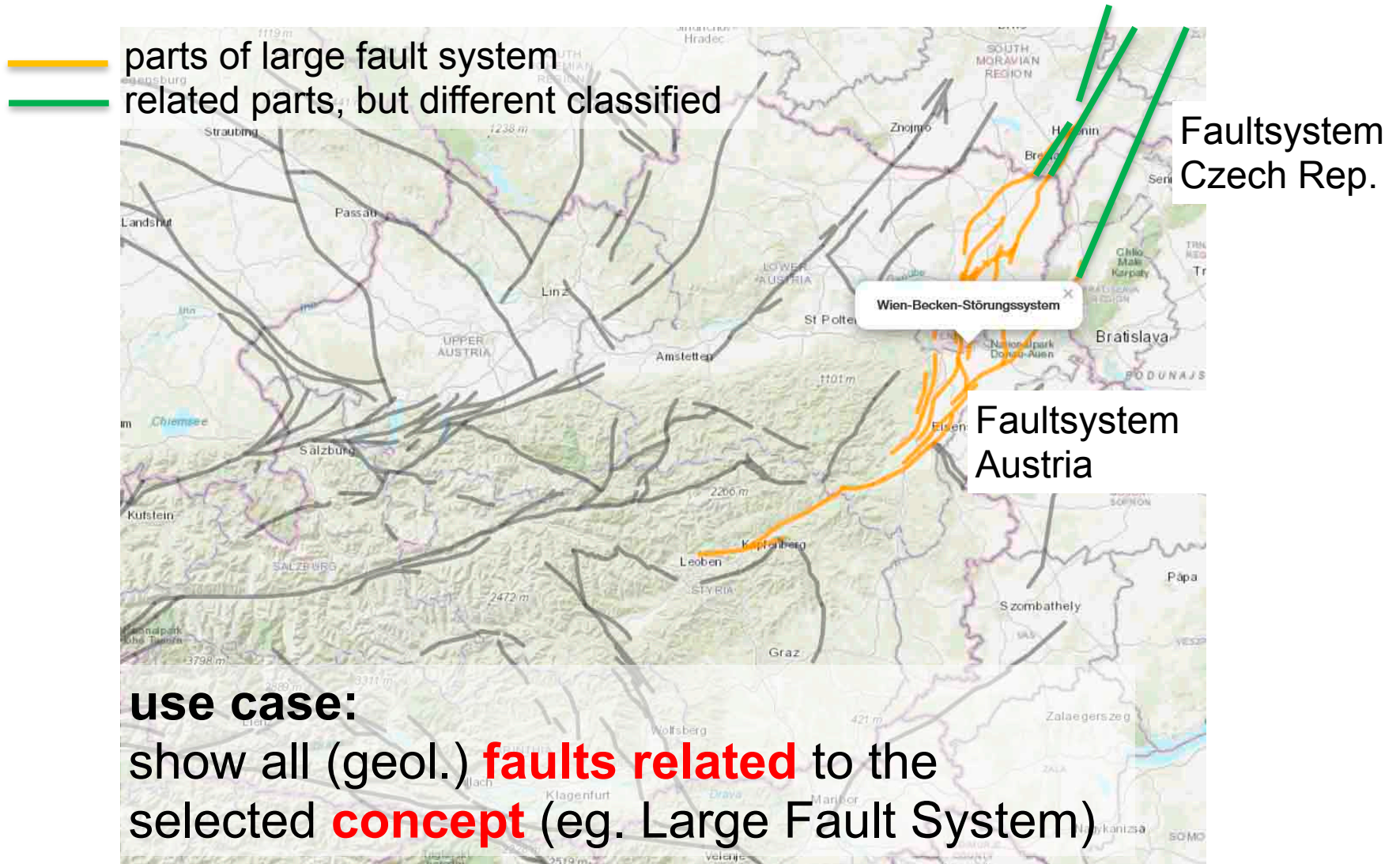
<http://resource.geolba.ac.at/structure/186>

- example GBA Thesaurus (HIKE fault systems)
- scientific confirmability
- loosely structured information
- multilinguality



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166





<https://thesaurus.geolba.ac.at/structureViewer.html?uri=http://resource.geolba.ac.at/structure/186&lang=en>



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



WMS Data:

OBJECTID	2
NummerderStörung	154
SHAPE	Polyline
FEATURE_ID	101
Klassifizierung	Störungssystem
Bezeichnung	Wien-Becken-Störungssystem
synonymeBezeichnung	Wiener-Becken Fault System
Großstörungssystem	Mur-Mürz-Wien-Becken-Vah-Großstörungssystem
Störungssystem	Wien-Becken-Störungssystem
Teilstörungssystem	Null
Störungsbzw.Scherzone	Null
Kurzreferenz	Beidinger and Decker, 2011; Decker, 1996; Decker et al., 2005 and references therein; Hölzel et al., 2010; Peresson and Decker, 1997
THESURL	http://resource.geolba.ac.at/structure/190

LinkedData:
[Wien-Becken-Störungssystem](#)

Dieses ca. 200 km lange, NNE-SSW-streichende Störungssystem beinhaltet alle Störungen und Strukturen des rhomboedrischen Wiener Pull-Apart-Beckens mit vorwiegend NE-SW-streichenden sinistralen Seitenverschiebungen an seinem Ostrand und NNE-SSW-streichende listrische Abschiebungen als Folge seiner E-W-Extension (Decker, 1996). Es zieht von Gloggnitz, seinem Südende, über die Donau zwischen Wien und Hainburg bis Zillna, an seinem Nordende. Es entstand als Pull-Apart-Becken während des Badenium (ca. 16 Ma) bis ins obere Miozän (9-8 Ma; Decker, 1996; Peresson & Decker, 1997; Hölzel et al., 2010). Quartäre Reaktivierungen der Störungen werden durch pleistozäne Terrassenbildungen und quartäre Beckenbildungen dokumentiert (Decker et al., 2005, Beidinger & Decker, 2011).

in ..Mur-Mürz-Wien-Becken-Vah-Großstörungssystem
 incl...Nussdorf-Störung, Steinberg-Störung, Pottendorf-Störung, Aderklaa-Bockfließ-Störung, Markgränesiedl-Störung, Engelhartstetten-Störung, Schrattnberg-Störung, Poysbrunn-Störung, Leopoldsdorf-Störung, Matzen-Störung, Lassees-Störung, Pirawarth-Hochleiten-Störung, Bisamberg-Störung,

Peresson, H. & Decker, K., (1997): Far-field effects of Late Miocene subduction in the Eastern Carpathians: E-W compression and inversion of structures in the Alpine-Carpathian-Pannonian region. - Tectonics 16, 1, 38-56.

use case:
 get additional information
 from **vocabulary** into a
WMS-identify (popup)

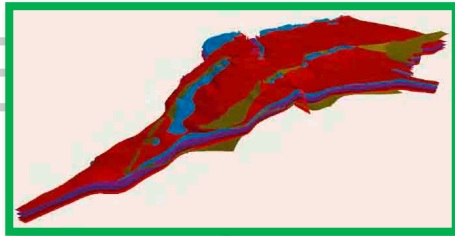
http://www.ce-gic.org/wms/GBA_structures_2.html



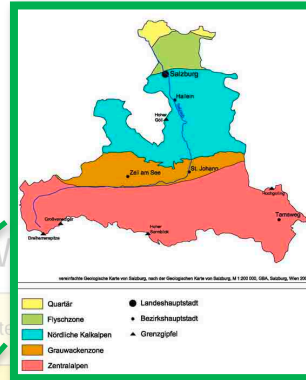
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



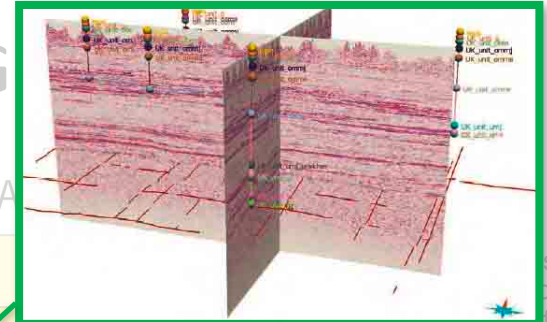
local 3D model



Map dataset



Borehole dataset



attributed with

attributed with

narrower terms

Formation B

Formation A

Formation C

related terms

Formation Z

narrower terms

Formation X

Formation Y

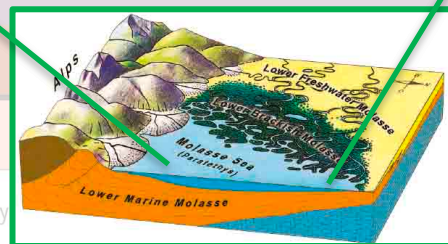
Swabian & Helvetic facies of Upper Jurassic

related terms

Germanic facies of Upper Jurassic

GeoERA project vocabulary

click&identify



3D model viewer

use case:

click on 3D view to **access** (bridge) **all source datasets** related to the selected 3D body



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166



Questions?



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731166

