OneGeology-Europe – an INSPIRE testbed for semantic harmonisation of „geology“ data across Europe (WP 3)

Kristine Asch and John Laxton
Project deliverables

- Interoperable on-shore geology spatial dataset
- with "progress towards harmonisation"
- Geological vocabulary and data specifications for Europe
- Use case studies
- Mutililingual metadata for discovery
- View services
- Forerunner and "guinea pig" for the implementation of INSPIRE Directive
Fact

Vast amount of data hidden in the archives and hard disks in governmental organisations across Europe ...
Edge matching at national boundaries?

- National boundary
- Geological terms and classifications (age, lithology, tectonics ..)
- Age of data (mapping campaign)
- Choice of units to be mapped
- Level of detail / scale
- Topographical base (projection, spheroid, drainage system, ...)
- Portrayal (colours and symbols)
- Mapped border of the units

And they are all different.
Interoperability and harmonisation

• **Interoperability** – when the data model/structure and properties to describe its parts (what GeoSciML does) is agreed
  – E.g. agreeing a data model will have the feature of “GeologicUnit” with properties of “age” and “lithology”

• **Semantic harmonisation** – when the use of the same definitions and classifications to describe a concept/term is agreed
  – E.g. ‘clay’. The same concept can be labelled with several terms (“argilla” in Italian, “Ton” in German), but needs to have the same definition, in this case of “clay/Ton/argilla, …”:
    > 50% particles < 0,004 mm (Wentworth grade scale),
    or:
    > 50% particles < 0,002 mm (ISO 14688)

• **Geometric harmonisation** – edge-matching at national boundaries
Harmonized?

- **Geometric harmonisation**
- **Semantic harmonisation**

Harmonized!

national boundary
• 516 agreed defined terms, definitions and parent/child relations:
  – lithology (sedimentary, magmatic, metamorphic)
  – geological age
  – genesis
  – faults and structures

• Taken on board: 532 review comments from 20 national data + 2 global NGOs

• > 100 new terms and definitions fed into the global geoscience vocabulary of the IUGS Commission of Geoscience Information

• Portrayal rules for age, lithology and structures

• Explanatory Notes how to encode in GeoScienceML

► THE base for semantic harmonisation
► Enables comparability of the information
► Basis for cross-boundary planning
Portrayal Rules
visualizing the content –
and its lack of harmonisation
Vocabulary Building

1G-E-core Team

1G-Europe scientific requirements

Particularly European (e.g. orogenies)

Global CGI Vocabulary

Extract

DRAFT

Review

1G-E Vocabulary
• Agreement: 10 individual country representatives defining the vocabulary (terms and definitions)
• 20 national representatives reviewing
• The (English) project Language – a neverending source for misunderstanding
• Integration with the global CGI vocabulary group – many heated discussions via e-mail, phone and in meetings
• Short time available
• Acceptance …. The need for a common denominator was not seen by all project participants from the start
WP 3 Geology
Harmonisation Workshop, Slovenia

- 26 experts
- 18 nations (EC)
- 2 days

Tasks:
- Define generic harmonisation issues
- Solve these individually
- Develop a workflow/method how to tackle this

Base: use of the 1G-E vocabulary
Our 1st obstacle

National boundary mismatches

Data gaps – no content
Harmonisation Workshop issues

- Hardly any country has harmonised its boundaries semantically and/or geometrically.
- Political issues… who provides Gibraltar, Northern Ireland?
- Base problem: the difference of scales; target scale 1/1000.000 includes 1/250.000, 1/300.000, 1/400.000, 1/500 000, 1/625 000, 1/1 000 000 → our advise: define the scale if there is a strong necessity for data harmonisation.
- Many mismatches solvable by bilateral negotiation, by use of more detailed data sets/maps and publications.
- However, there will mismatches remain that can only be solved with additional field work!
Geological harmonisation:
Draft of a general workflow process*

Condition: parties use same conceptual model and vocabulary

Define properties to be harmonized

Define level of harmonisation

Identify and classify mismatches

Solvable by negotiation?

Mark feature

Modify general unit

Modify just identified border polygons

Inform neighbours and start negotiating process

Generalize semantic /snap boundaries

Tools: Mismatch matrix

*Draft of workflow inspired by eSDIN project coordinator Natalie de Lattre, merci Natalie!
Summary and Conclusions

- The work on the OneG-E data vocabulary enriched and improved the global CGI vocabulary
- It has been reviewed and agreed by 20+4 countries (subsidiarity principle)
- Provides a solid base, easy-to-use vocabulary to describe the geology of Europe
- Explanation notes are written to help the implementation
- Provides a geology vocabulary and basis for semantic harmonisation of geology in Europe
- Reference material for the INSPIRE Geology and Mineral Resources Data TWG GE-MR Specification TWG GE-MR….
- “progress-towards-harmonisation” report provides guidance for future cross-border consistency and harmonisation
Thanks to:

- Ian Jackson
- Steve Richard, CGI
- Garry Baker
- Robert Tomas
- Jean-Jacques Serrano
- Agnès Tellez-Arenas
- Marco Klicker
- Horst Troppenhagen
- Mikko Nironen
- Urzula Stepień
- Piotr Czupek
- Francois Robida
- Luca DeMicheli
- Claudia Delfini
- Aleksandra Kuczerawy
- Alexander Tschistiakow
- Stephan Gruijters
- Mary Carter
- Marco Pantaloni
- Fernando Perez
- Stefan Bergman
- Stefan Kacer
- Pavla Guertlerova
- Jolanta Cylene
- Chris Schubert
- Sybille Hennings
- Kathryn Bull
- Dominique Janjou
- Milos Bavec
- Alan Smith
- Pierre-Yves Declerq
- Monika
- and many more of the 1G-E Team...
IUGS-CGI & OneGeology-Europe
Berlin International Geoscience Language Workshop
BEAAR 2010
+ Open GeoSciML Day
25.-27. August 2010

One Science – One Language:
Ontologies, vocabularies and terms:
Developments and implementation

Patron: Federal Minister Rainer Brüderle