A Strategy for Data Management: Data Models & Standards

Workshop
Digital Field Data Acquisition and Data Management
October 13th 2016

Sabine Brodhag
A Strategy for Data Management: Data Models & Standards

Digital Field Data Acquisition and Data Management


The smartphone and tablet revolution has changed how geologists work in the field, but now the community must come up with standards to tame the flood of data.
Data Management

Data management encompasses the entire data life cycle.

2. Data collection
   Field observation, measurements etc.
Data Management

Data management encompasses the entire data life cycle.

1. Planning / Concept
   What do I want to do with the data?

2. Data collection
   Field observation, measurements etc.
Data Management

Data management encompasses the entire data life cycle.

1. Planning / Concept
   What do I want to do with the data?

2. Data collection
   Field observation, measurements etc.
   Documentation
   Media, technical devices, formats, Metadata

---

**Local time:** 01:34:08
**Last update:** 01:34:08
**Coordinate:**
- **Latitude:** 46°25'40.96"N
- **Longitude:** 30°42'47.47"E

**Coordinate:**
- **Latitude:** 46.428045
- **Longitude:** 30.713185

**Altitude:** 54.9 meters
**Speed:** 0 m/s
**Speed (alt):** 0 km/h
**Max speed:** 0
**Accuracy:** within 7 meters
**Bearing:** 0°
**Cur. distance:** 19.93 meters (0.01)
**Full distance:** 19.93 meters
**Error**
Data Management

Data management encompasses the entire data life cycle.

1. Planning / Concept
   What do I want to do with the data?

2. Data collection
   Field observation, measurements etc.
   Documentation
   Media, technical devices, formats, Metadata

3. Analysis, Processing
   evaluation
   interpretation
   production
Data Management

Data management encompasses the entire data life cycle.

4. Storage
   Additional metadata
   confidentiality
   archiving
Data Management

Data management encompasses the entire data life cycle.

4. Storage
   Additional metadata
   confidentiality
   archiving

5. Retrieval
   indexing
   related data
Data Management

Data management encompasses the entire data life cycle.

4. Storage
   Additional metadata
   confidentiality
   archiving

5. Retrieval
   indexing
   related data
Data Management

Data management encompasses the entire data life cycle.

- Exchange
- Sharing
- Distribution
- Collaboration

Terms of use
citation
packaging info
and…

… common data models & standards
Are we already using common standards when acquiring our field data?

- Coordinate reference system
- Units (m, mm, kg...)
- Rock unit names
- Pdf, jpg, txt, ascii
- English!

- Map legend (colours)
Without well defined standards
With well defined standards
With well defined standards

RGB colour codes are provided on www.strati.ch:
swisstopo Data Model Suite

Data Model Geology
Description in UML Format and Object Catalogue, Version 2.1

Data model geology

Datenmodelle der nicht-Geobasisdaten der Landesgeologie, geophysikalischen und geotechnischen Karten (Pixelkarten)
Objekt-Katalog und Beschreibung und INTERLIS, Version 1.5

Data model boreholes

www.strati.ch
What is a data model?

Data Model

Geology

Theme: Rock Bodies

UML diagram (Unified Modelling Language)
The backbone: structure & relations

Theme: Rock Bodies
Class: Bedrock

Sedimentary Rock
Igneous Rock
Metamorphic Rock

Lithostratigraphic Units:
- Lithology
- Lithostratigraphy
- Chronostratigraphy
- Tectonostratigraphy

Description:
- Main Component
- Second Component
- Bonding Material
- Bedding
- ….

Description:
- Texture
- Grain Size
- ….

Description:
- Minerals
- Metam. Structure
- ….

→ Lookup tables
Muscles & fat: the lookup tables

Table 6 — Examples of terms which may be used in the description of rock mass structure

<table>
<thead>
<tr>
<th>Sedimentary</th>
<th>Metamorphic</th>
<th>Igneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedded</td>
<td>Cleaved</td>
<td>Massive</td>
</tr>
<tr>
<td>Interbedded</td>
<td>Foliated</td>
<td>Flowbanded</td>
</tr>
<tr>
<td>Laminated</td>
<td>Schistose</td>
<td>Folded</td>
</tr>
<tr>
<td>Folded</td>
<td>Banded</td>
<td>Lineated</td>
</tr>
<tr>
<td>Massive</td>
<td>Lineated</td>
<td></td>
</tr>
<tr>
<td>Graded</td>
<td>Gneissose</td>
<td>Folded</td>
</tr>
</tbody>
</table>
Muscles & fat: the lookup tables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>012</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>013</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>014</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Pliocene</td>
<td></td>
</tr>
<tr>
<td>015</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Pliocene</td>
<td>Placenzian</td>
</tr>
<tr>
<td>016</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Pliocene</td>
<td>Zanclean</td>
</tr>
<tr>
<td>017</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Late Miocene</td>
</tr>
<tr>
<td>018</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Late Miocene</td>
</tr>
<tr>
<td>019</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Late Miocene</td>
</tr>
<tr>
<td>020</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Messinian</td>
</tr>
<tr>
<td>021</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Tortonian</td>
</tr>
<tr>
<td>022</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Middle Miocene</td>
</tr>
<tr>
<td>023</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Serravallian</td>
</tr>
<tr>
<td>024</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Langhian</td>
</tr>
<tr>
<td>025</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Early Miocene</td>
</tr>
<tr>
<td>026</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Early Burdigallian</td>
</tr>
<tr>
<td>027</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Late Burdigallian</td>
</tr>
<tr>
<td>028</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Neogene</td>
<td></td>
<td></td>
<td></td>
<td>Miocene</td>
<td>Early Burdigallian</td>
</tr>
<tr>
<td>029</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>030</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>031</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>032</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>033</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>034</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>035</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>036</td>
<td>Phanerozoic</td>
<td>Cenozoic</td>
<td>Tertiary</td>
<td>Paleogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Federal Office of Topography swisstopo
Swiss Geological Service
The backbone: structure & relations

Lithostratigraphic Units:
- Lithology: sandstone
- Lithostratigraphy: Sense Schichten
- Chronostratigraphy: Burdigalien
- Tectonostratigraphy: no alpine deformation

Description:
- Main Component: quartz
- Second Component: feldspar
- Bonding Material: calcareaus cement
- Bedding: bedded
- ….

Theme: Rock Bodies
Class: Bedrock

Sedimentary Rock
Why use data models?
What’s the benefit?
Why use data models?
What’s the benefit?

1. Common “language”

2. Structuring and harmonisation of data

3. Structuring of complete data life cycle

4. System independent

→ Effective exchange of data
→ Access to data!
Why use data models?  
What’s the benefit?

1. Combination of data across projects
2. Access to temporally (and spatially) unique data
3. Independent analysis of original data
4. Uncovering of statistical or coding errors

→ Transparency and reproducibility
→ Integrity of results
→ More & more progressive data-sharing policies
Data Management at swisstopo?

Next talk: Roland Baumberger
Questions?