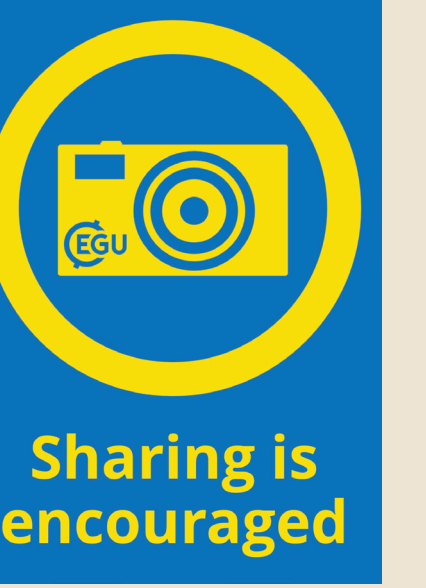


# Improving the findability of legacy laboratory data: enrich metadata using controlled vocabularies

Laurens Samshuijzen<sup>1</sup>, Otto Lange<sup>1</sup>, Ronald Pijnenburg<sup>2</sup>, Richard Wessels<sup>2</sup>, Maik Nothbaum<sup>2</sup>

<sup>1</sup>Utrecht University Library, Utrecht University, Utrecht, the Netherlands

<sup>2</sup>Faculty of Geosciences, Utrecht University, Utrecht, the Netherlands



## 1. Keyword enrichment by detecting terms in

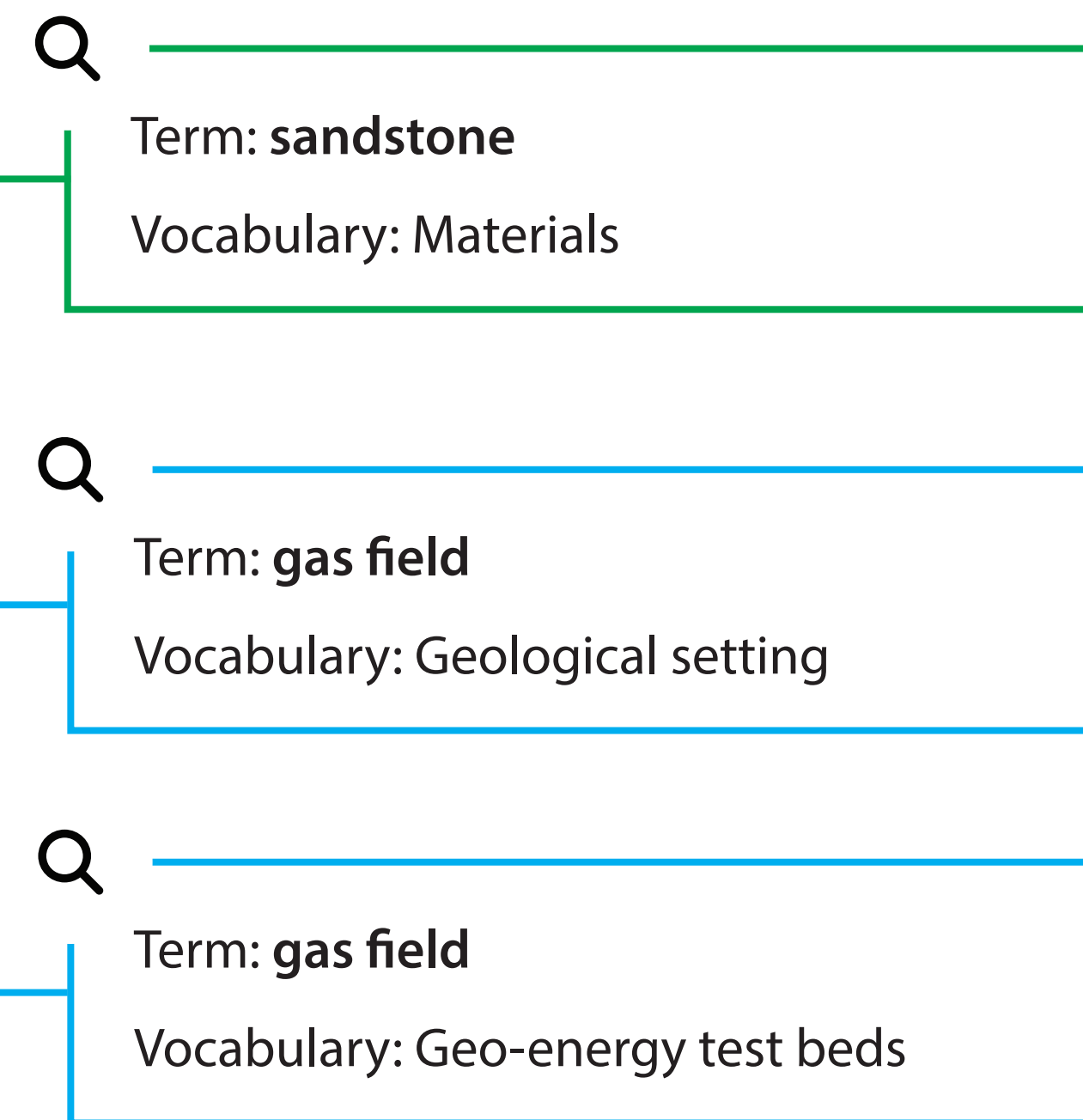
Title

Mechanical and microstructural data used in the article

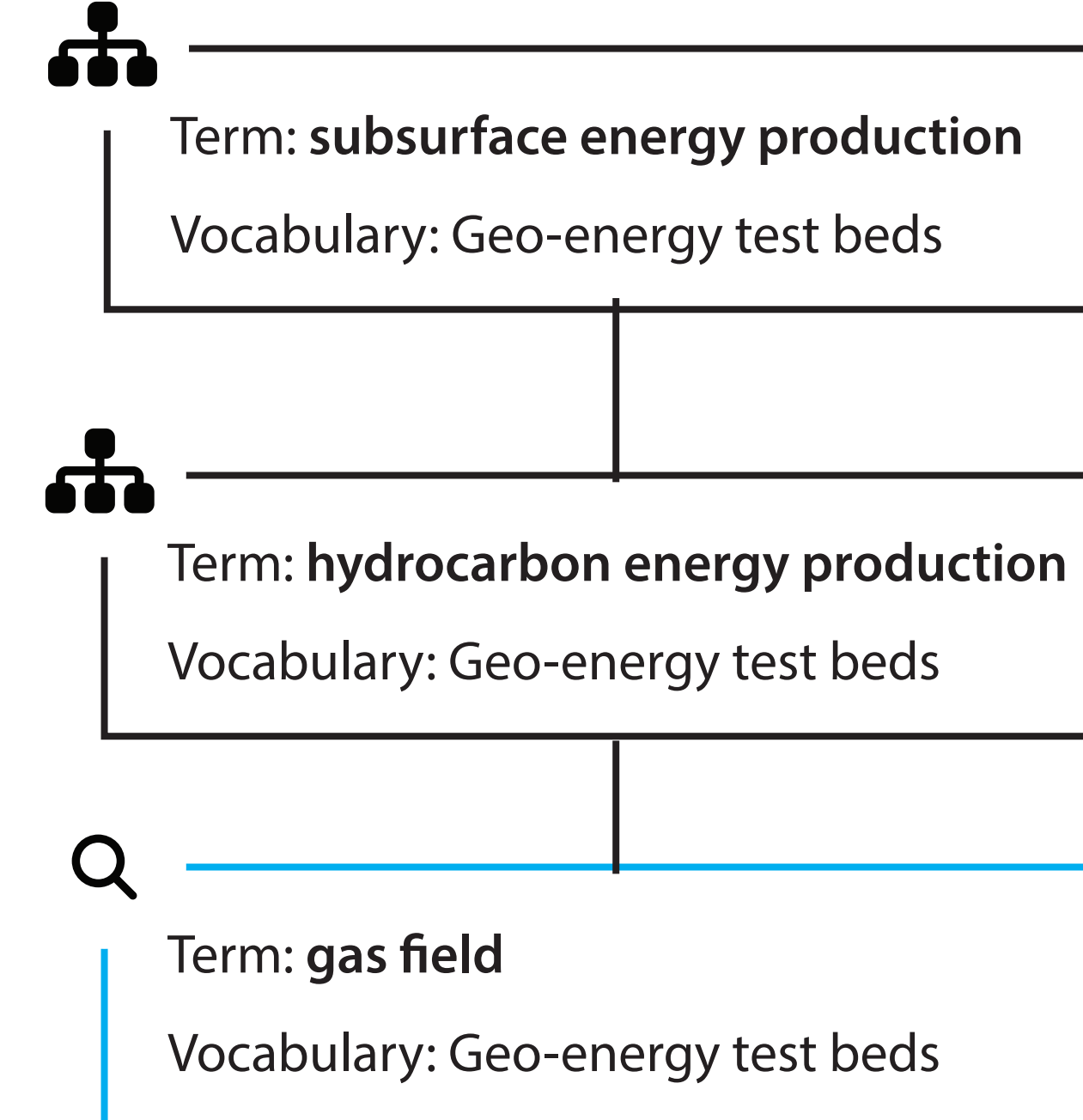
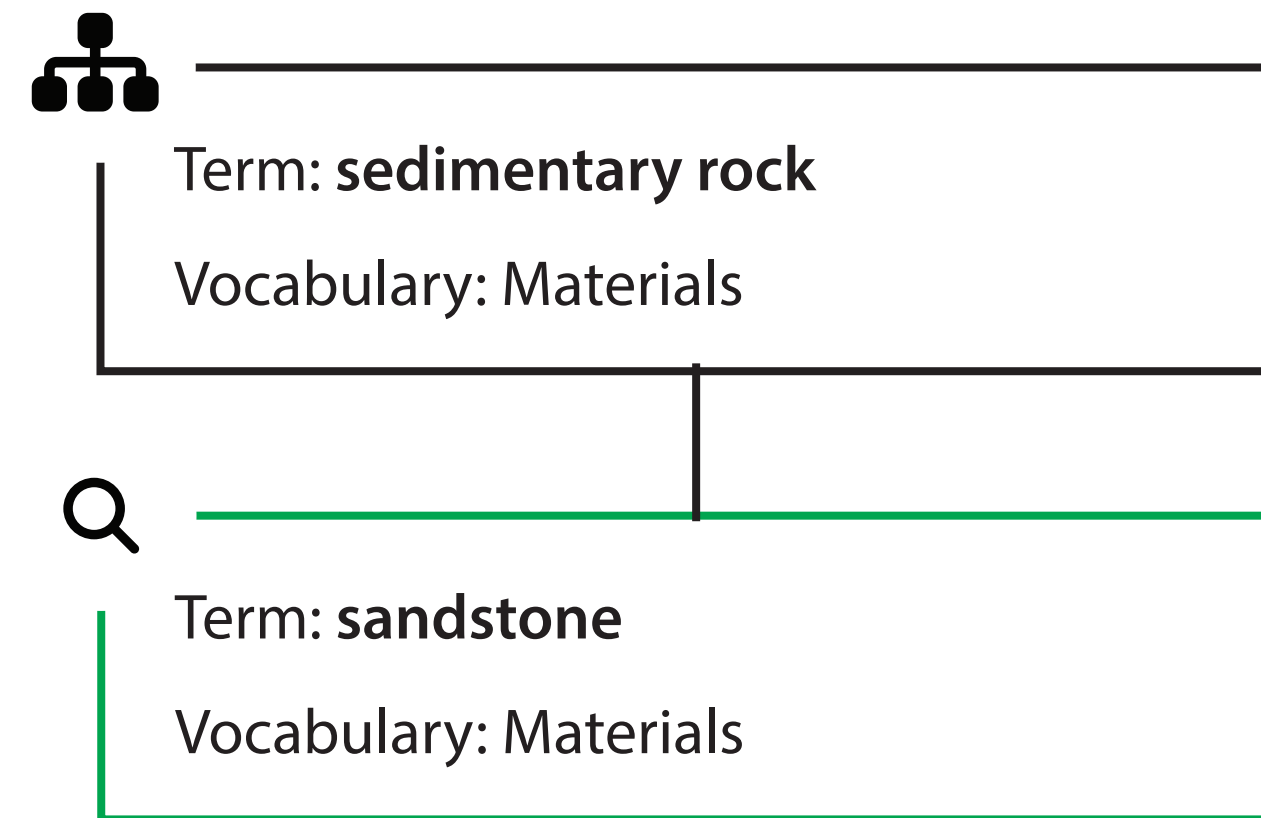
Pijnenburg et al., Deformation behaviour of sandstones

from the seismogenic Groningen gas field

Role of inelastic versus elastic mechanisms



## 2. Providing context by adding overarching



## 3. Usage in data catalogue

Detailed keyword information

Originally assigned keyword

Harmonized keywords

The enriched keywords contain on average many more keywords than the originally assigned keywords. Combined with keyword harmonization this greatly improves the findability of data publications.

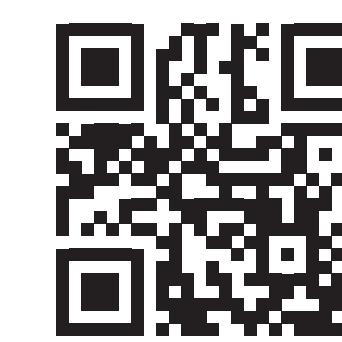
Filtering data publications

Using controlled vocabularies as filters greatly improves the findability combined with keyword detection.

Filters datasets with increasing granularity

- material
  - minerals
    - carbonate minerals
    - ...

Visit the data catalogue at:



epos-msl.uu.nl



- Vocabularies:
- Material
  - Geological age
  - Pore fluid
  - Geological setting
  - (sub)surface utilization setting
  - Analogue modelling of geological processes
  - Geochemistry
  - Geo-energy test beds
  - Microscopy and tomography
  - Paleomagnetism
  - Rock and melt physics

Containing a total of 2963 terms

**Challenge: How to populate the vocabularies with descriptions for all terms?**



Developed and harmonized by the Earth scientific laboratory community in Europe.

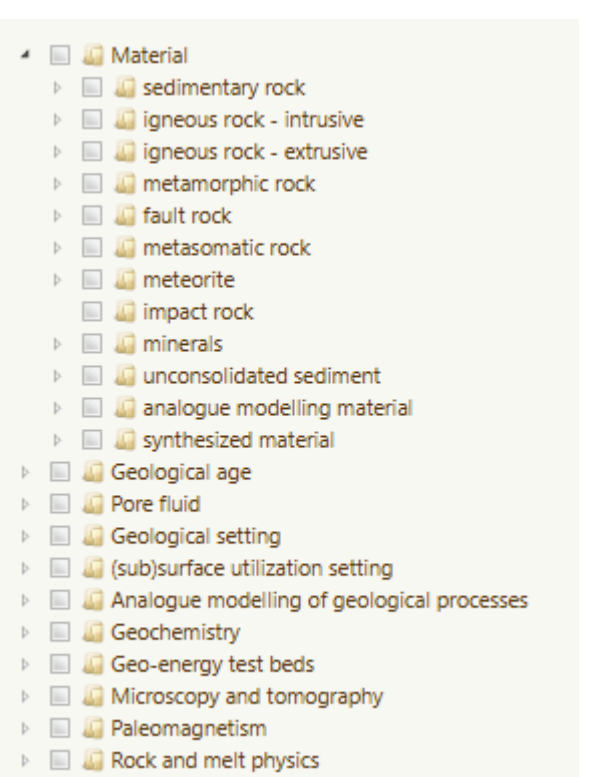
Published in several open formats on GitHub:

- JSON
- Excel
- Turtle
- RDFXML



[github.com/UtrechtUniversity/msl\\_vocabularies](https://github.com/UtrechtUniversity/msl_vocabularies)

Reusable web component available to help integrate the vocabularies within other applications.



Contact us:

epos.msl.data@uu.nl