

Paper Number: 1256

## **International standardisation of geoscience terminology via CGI-IUGS vocabularies**

Rattenbury, M.S.<sup>1</sup>, Cipolloni, C.<sup>2</sup>, Heaven, R.<sup>3</sup>, Raymond, O.L.<sup>4</sup>, Richard, S.M.<sup>5</sup>, Vuollo, J.<sup>6</sup>

<sup>1</sup>GNS Science, Lower Hutt, New Zealand, ([m.rattenbury@gns.cri.nz](mailto:m.rattenbury@gns.cri.nz))

<sup>2</sup>Geological Survey of Italy, ISPRA, Rome, Italy

<sup>3</sup>British Geological Survey, Nottingham, UK

<sup>4</sup>Geoscience Australia, Canberra, Australia

<sup>5</sup>Arizona Geological Survey, Tucson, USA

<sup>6</sup>Geological Survey of Finland, Rovaniemi, Finland

---

The standardisation of geoscience terminology through compilation of controlled vocabularies has proceeded in parallel with the development of geoscience data models such as GeoSciML and EarthResourceML. These vocabularies contain terms that describe geoscience features, ideally enough to encompass feature variability without unnecessary overlap and duplication. Vocabularies facilitate searches of multiple distributed and appropriately structured databases. These can include searching different languages as demonstrated with the OneGeology Europe web portal application <http://geoportal.onegeology-europe.org/geoportal>. The compilation, review and adoption of controlled vocabularies are the primary functions of the Geoscience Terminology Working Group (GTWG) that operates under the auspices of the Commission for the Management and Application of Geoscience Information (CGI). The GTWG has a membership of 28 geoscience information specialists from 14 countries well spread around the world [1].

The GTWG and predecessor groups have so far compiled and published 65 vocabularies covering geological, geomorphological, geophysical, geochemical, exploration and mining related terms. These account for approximately half of those required for the GeoSciML and EarthResourceML geoscience data models. Recent work has resulted in 21 vocabularies published on mining and mineral exploration themes, including 'commodity code', and geological themes, including the 'regional lithological unit synthesis' list of mapped rock associations at a continental scale for use in OneGeology. These vocabularies are accessible through a server administered by CGI, currently hosted in Australia.

Initial compilation of the vocabularies is undertaken by a nominated 'shepherd', usually from within the GTWG membership, who will typically assemble relevant lists of terms for a specific vocabulary theme from as many organisations as possible. A structured spreadsheet is compiled that hierarchically organises the terms from general to specific (parent-child) with synonyms, definitions and source information. The draft list is reviewed by the GTWG membership and their specialist colleagues and once all comments have been addressed, adoption of the list is voted on by the group. An adopted vocabulary is then processed from human-readable spreadsheet to a computer-readable, XML-based resource descriptor framework (SKOS-RDF) file format which in turn is published through the vocabulary server (available through <http://resource.geosciml.org>).

The initial working language for vocabulary development is English but a goal of the GTWG is to enable translation of the lists into as many languages as possible, making use of geoscience term labels developed by a predecessor CGI group (Multi Lingual Thesaurus of Geosciences) where available. Multi-lingual versions of the stratigraphic age list are currently available.

*Reference:*

[1] [http://cgi-iugs.org/tech\\_collaboration/geoscience\\_terminology\\_working\\_group.html](http://cgi-iugs.org/tech_collaboration/geoscience_terminology_working_group.html)

