

OneGeochemistry: moving beyond stand-alone databases to an online Open FAIR Global Geochemical Data Network

LESLEY A. I. WYBORN¹, KERSTIN A LEHNERT²,
MARTHE KLÖCKING³, KIRSTEN ELGER⁴, REBECCA
FARRINGTON⁵, DOMINIK C. HEZEL⁶, DR. ALEXANDER
PRENT^{5,7}, ANGUS NIXON⁸ AND LUCY PROFETA⁹

¹Australian National University

²Lamont-Doherty Earth Observatory, Columbia University

³University of Göttingen

⁴GFZ German Research Centre for Geosciences

⁵AuScope

⁶Goethe-Universität Frankfurt

⁷AuScope Ltd

⁸The University of Adelaide

⁹Columbia University

Presenting Author: lesley.wyborn@anu.edu.au

Since the 1980's, multiple individual geochemical databases have been created. As most were developed with opportunistic project-based funding, many now reside in the graveyard of unsustainable databases. PETDB, GEOROC and databases of the National Geological Surveys (e.g. USA, UK, Australia, Japan) have been the survivors.

Building and sustaining geochemical databases over the longer term is not easy. The field of geochemistry is constantly evolving, as technological advances, increased automation and computerisation enable creation of new analytical methods and data types. Data volumes continue to grow exponentially, particularly as the use of 2D and 3D imagery becomes pervasive in geochemistry.

Developers of machine-actionable geochemical databases also have to keep pace with new developments in online delivery systems, which enable the creation of 'virtual datasets' that facilitate aggregation of data from multiple sources and offer the potential to create a FAIR and Open Global Geochemical Data Network as a collaborative effort between distributed systems. To achieve this ambition, datasets need to be FAIR (Findable, Accessible, Interoperable, Reusable), machine-actionable and as Open as possible.

Developing machine-actionable FAIR data models for each geochemical data type is time consuming. A more efficient approach is to leverage the Observation, Measurement and Samples Standard (ISO 19156: 2023) and create generic modules for samples, instruments, analytical methods, properties, etc. Several of these modules are common to all data types and can be reused. What is then required is for expert communities to define minimum agreed variables and controlled vocabularies for each specialised data type. However, to then create a global geochemical network these community standards require international agreement.

OneGeochemistry is an international collaboration that seeks

to coordinate global efforts in geochemical data standardisations by building and maintaining consensus-driven standards that make geochemistry data FAIR to both humans and machines. OneGeochemistry is also one of 11 case studies of the CODATA-led EU-Horizon WorldFAIR project whose goal is to advance implementation of FAIR and improve interoperability and reusability of data across multiple disciplines. Combined, both projects will democratise geochemical data and facilitate creation of a global network that will enable online sharing and reuse for both disciplinary and interdisciplinary grand challenges.