OneGeochemistry: creating a global FAIR-way to access and share geochemical data.

LESLEY WYBORN[1] AND KERSTIN LEHNERT[2]

[1]Australian National University, Canberra, ACT, 2600 Australia, lesley.wyborn@anu.edu.au.

[2] Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, 10964, USA, lehnert@ldeo.columbia.edu.

Geochemistry has the capacity to contribute to our u understanding of many natural systems, from the interior of the Earth to its surface environments, to the entire solar system. All geochemical research is underpinned by data of some type, and the capacity of new generation analytical facilities to generate that data at ever increasing rates is rapidly leading to the development of multiple geochemical database systems at national, thematic, and project levels to simply enable storage and management of data generated. At the same time, OpenAccess policies now require that data, particularly that which supports scholarly publications/ research grants meet the FAIR principles (Findable, Accessible, Interoperable, and Reusable) and necessitates datasets being housed in trusted repositories that can provide persistent unique identifiers and landing pages. Unlike some other geoscience domains which have built global data sharing systems (e.g., seismology, geodesy, marine geoscience, OneGeology), best efforts in Geochemistry are currently only making data Findable and Accessible: it is not possible to create Interoperable networks of Reusable geochemical data mainly because the global Geochemistry community has yet to define the required, internationally-agreed standards and best practices. OneGeochemistry is a new proposal to build a global sharing network of geochemical data. The highest priority is on defining the required standards, vocabularies, ontologies and documenting best practices as a first step towards enabling linking of all geochemical data globally. Due to the ever increasing volumes of data being generated, these standards and practices will also have to conform to best scientific data practices that enable machine actionable processing, Artificial Intelligence and Machine Learning.