

EGU22-10726

<https://doi.org/10.5194/egusphere-egu22-10726>

EGU General Assembly 2022

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



Global Data Standards for Geochemistry: Not the ‘One Ring to Rule Them All’, but a set of ‘Olympic Rings’ that Link and Integrate across Continents

Kerstin Lehnert¹ and Lesley Wyborn²

¹Columbia University, Lamont-Doherty Earth Observatory, Palisades, NY, 10964, United States of America

(lehnert@ldeo.columbia.edu)

²Australian National University, Canberra, Australia

As volumes of geoscientific data grow, research in geochemistry, volcanology, petrology, and other disciplines working with geoscientific data is evolving to data-driven and computational approaches that have enormous potential to lead to new scientific discoveries. Application of advanced methods for data mining and analysis including Machine Learning, and Artificial Intelligence, as well as the generation of models for simulating natural processes all require seamless machine-readable access to large interoperable stores of consistently structured and documented geochemical data. Standard protocols, formats, and vocabularies are also critical in order to process, manage, and publish these growing data volumes efficiently with seamless workflows that are supported by interoperable tools.

Today, easy integration of data into Analysis Ready Data stores and the successful and efficient application of new research methodologies to these data stores is hindered by the fragmentation of the international geochemical data landscape that lacks the technical and semantic standards for interoperability; organizational structures to guide and govern these standards; and a scientific culture that supports and prioritizes a global sustainable data infrastructure. In order to harness the scientific treasures hidden in BIG volumes of geochemical data, the science community, geochemistry data providers, publishers, funders, and other stakeholders need to come together to develop, implement, and maintain standards and best practices for geochemical data, and commit to changing the current data culture in geochemistry. The benefits will be wide-ranging and increase the relevance of the discipline.

Although many research data initiatives today focus on the implementation of the FAIR principles for Findable, Accessible, Interoperable, and Reusable data, most data is only human-readable, even though the original purpose of the FAIR principles has been to make data machine-actionable. The development of standards today should not focus on spreadsheet templates used to format and compile project-centric databases that are hard to re-purpose. These methods are not scalable. The focus should be on global solutions where any digital data are born connected to agreed machine readable standards so that researchers can utilize the latest AI and ML techniques.

Global standards for geochemical data should not be perceived as 'one ring to rule them all', but rather as a series of interoperable 'rings' of data, which like the Olympic rings will integrate data from the all continents and nations.