

# **Walking the Line: How Can EarthChem Ensure Data Quality While Minimizing Burden on Researchers for Open & FAIR Data Publishing in Geochemistry?**

LUCIA PROFETA<sup>1\*</sup>, KERSTIN LEHNERT<sup>1</sup>

<sup>1</sup> Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

Researchers increasingly have to comply with policies for Open and FAIR data: Funders and publishers aim to ensure that the results of scholarly research are not only openly accessible, but that they are Findable, Accessible, Interoperable, and Reusable (FAIR) [1]. Researchers are asked by funding agencies or publishers to deposit data in trusted domain repositories that follow international best practices for data curation. With growing adoption and implementation of open and FAIR data policies, the need for clear and globally consistent procedures, practices, and guidelines is paramount. Many questions in this emerging landscape of FAIR data need to be solved: Who develops and maintains data standards? Who has the authority to approve and govern them at a global scale? How granular can, or should, domain-specific standard definitions be?

Domain specific repositories, such as the EarthChem Library in the case of geochemical data, are committed to maintaining the highest quality of data possible, but also need to make it as easy as possible for researchers to deposit data. This raises a number of important questions: Is data quality primarily the onus of the researcher who submits it? Should data submissions be reviewed in a process similar to article peer-review? How far should repositories and their data curators institute and enforce data quality practices? A clear vision needs to be established moving forward in order to help both researchers and repositories have a clear pathway to making geochemical data FAIR.

The EarthChem Library is exploring a tiered approach for streamlining metadata intake, based on researcher contributions. This would entail a FAIRness indicator, to reflect how comprehensive the data documentation is.

We encourage community input and feedback to improve data management practices and help guide the discussion for establishing standards.

[1] Wilkinson et al. (2016) *Sci Data*. 2016;3:160018.