THE ASTROMATERIALS DATA SYSTEM: ENABLING OPEN AND FAIR DATA FROM ASTROMATERIALS SAMPLES ANALYSIS

K. Lehnert[1], J. Mays[1], P. Ji[1], L. Profeta[1], J.D. Figueroa[1], A. Johansson[1], S. Morrison[2]
[1]Lamont-Doherty Earth Observatory, Columbia University, USA, lehnert@ldeo.columbia.edu
[2]Carnegie Science Earth & Planets Laboratory, USA

Introduction: In order to advance Open, transparent, and reproducible science and data-driven research, data generated in scientific studies need to be openly shared in a manner that makes them persistently Findable, Accessible, Interoperable, and Reusable (=FAIR) by both humans and machines as codified in the FAIR Principles [1]. Funders and publishers alike are aligning their data policies with these principles. For example, NASA's new Science Information Policy SPD-41 [2] lays out the specific requirements that the Science Mission Directorate expects for data to be approved as FAIR. Implementation of such policies requires infrastructure for FAIR-compliant data curation that allows mission scientists and researchers who generate post-mission data to readily comply with these data policies.

In this presentation we will describe how the Astromaterials Data System (https://www.astromat.org) ensures FAIR astromaterials sample data. We will also connect Astromat's FAIR data services to the recommendations of the Planetary Data Ecosystem Independent Review Board (PDE-IRB) released in 2021 [3].

Astromaterials Data System (Astromat). The Astromaterials Data System is a NASA-funded data facility that develops and maintains online systems and services for publishing, disseminating, and archiving newly acquired data as well as for restoring and preserving historic data generated by the analysis of astromaterials samples. Astromat's systems, services, policies, procedures, and management have been designed to comply with the FAIR principles for data and with the TRUST principles for Transparent, Responsible, User Focused, Sustainable, and Technologically stable repositories [4]. Among the services that make Astromat curated data FAIR are: 1) documentation of data with rich metadata for discovery and reuse, specifically extensive documentation of data provenance (e.g., methods of data and sample acquisition and processing) and data quality; 2) review of submitted content by data curators, who engage with users to augment and enrich, if necessary, documentation, 3) registration of data with DOI to make the data citable; 5) inclusion of persistent identifiers in the metadata (e.g., DOI for cited publications, IGSN for cited samples, funding award numbers, and mission identifiers), so that datasets can be linked to other digital resources; 6) clear licenses for access and reuse of curated data; 7) machine-readable interfaces (API) to access metadata and data; 8) use of community best practices and data standards, including vocabularies. In order to further enhance reusability and impact of Astromaterials sample data, Astromat data curators synthesize historic and new data into a comprehensive, analysis-ready data store that allows users to explore and mine astromaterials data in ways that have not been possible before, using new technologies such as Machine Learning and Artificial Intelligence. Astromat is also integrating the content of MetBase (https://metbase.org), the world's largest compilation of meteorite data, to make its data collection as comprehensive as possible.

Astromaterials Data Archiving Special Study. Additional demands for the management of astromaterials data have emerged driven by a new generation of sample-return missions and subsequent sample analysis that require an assessment of the current landscape of astromaterials data archiving capabilities and capacity. Astromat is conducting a Special Study for NASA, currently underway, that will result in a report with recommendations for a NASA Astromaterials Data Archive that provides a scalable and standards-compliant solution for archiving and disseminating astromaterials data as Open and FAIR data. We will report on outcomes of this Special Study.

References: Use the brief numbered style common in many abstracts, e.g., [1], [2], etc. References should then appear in numerical order in the reference list, journal names should be spelled in full. Please, use the following style:

[1] Wilkinson et al. (2016) *Sci Data. 2016;3:160018*. [2] Scientific Information Policy for the Science Mission Directorate, SMD Policy Document SPD-41 (August 2021). [3] Besse, S., et al. (2021). *LPI Contributions 2549, 7070*. [4] Lin, D., et al. (2020). *Sci Data, 7(1), pp.1-5*.