Goldschmidt 2026 Abstract
https://doi.org/10.46427/gold2026.38141

Making Returned Sample Analytical Data Reusable: Metadata Curation in Astromat

RUOLIN DENG[1], KERSTIN LEHNERT[1], JENNIFER MAYS[1], PENG JI[1], IRINA ISHAN[1] AND STEPHEN RICHARD[1,2]

[1]Lamont-Doherty Earth Observatory of Columbia University
[2]US Geoscience Information Network (USGIN)

Returned extraterrestrial samples from planetary bodies are crucial to shaping our understanding of Solar System history and natural processes beyond Earth. With an increasing number of global planetary missions, there is a growing need for robust curation of analytical data and metadata of returned samples. While scientific insights are typically drawn from data, metadata provides data context, supports data credibility, and makes data more findable, accessible, interoperable, and reusable (the FAIR principles; [1]). Metadata is therefore equally important for scientific research.

Astromat is the primary NASA-sponsored data archive for laboratory analyses of returned samples from planetary missions (*e.g.*, Bennu samples from the OSIRIS-REx mission) and astromaterials curated at the Johnson Space Center [2], hosting more than 4,000 datasets and metadata records in its repository and more than 2.5 million analytical data values in its synthesis database. Essential for supporting data reusability for future research, Astromat includes associated metadata such as instrument configuration, analytical settings, data processing, and quality control. Currently, reporting formats, vocabularies, and level of detail of archived metadata vary by author, method, mission, etc. This heterogeneity in metadata quality has become a practical challenge as Astromat continues to develop and improve its database infrastructure and design curation policies in preparation for increasing archiving demands from future planetary missions.

Here we present the progress of ongoing efforts by Astromat to develop an improved metadata curation pipeline for returned samples. The pipeline draws from existing planetary mission policies, archived data and metadata in Astromat, protocols established by other geo/cosmochemistry databases (*e.g.*, EarthChem, GEOROC, the Meteoritical Bulletin Database), and evaluation by research experts. The pipeline is expected to encompass recommended metadata reporting protocols, submission and registration workflows, review and curation policies, and enhanced metadata searchability. Ultimately, it aims to bring analytical data in Astromat into full compliance with the FAIR principles, improve overall curation standards, and assist data-driven research on returned samples.

[1] Wilkinson, Mark D., et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific data* 3.1 (2016): 1-9.

[2] NASA Planetary Science Division. "Planetary Science Division Information and Data Management Policy." *NASA Science Mission Directorate* (2023)