

Maximizing the Return on Investment in Astromaterials Samples Collection and Analysis with FAIR & Analysis-Ready Data at the Astromaterials Data System

KERSTIN A LEHNERT¹, JENNIFER MAYS¹, ROBERT DOWNS², LUCY PROFETA¹, GRIFFIN DANNINGER¹, PENG JI¹, MOLLIE CELNICK¹, JUAN DAVID FIGUEROA¹, ANNIKA JOHANSSON¹ AND STEPHEN RICHARD¹

¹Columbia University

²Center for International Earth Science Information Network (CIESIN), Columbia Climate School, Columbia University

Presenting Author: lehnert@ldeo.columbia.edu

Immense public investment supports the collection of astromaterials samples and their subsequent analysis in laboratories to produce data that are of fundamental significance for our understanding of planetary processes, the evolution of the solar system and the origin of life on Earth. These investments continue with recent and planned sample-return missions including OSIRIS-REx, Mars2020, and Artemis. It is imperative that the return on these investments (ROI) is maximized, and one of the opportunities to do so is to ensure that the data generated by laboratory analysis of astromaterials samples is acquired, curated, archived, and disseminated in a way that enables members of the research community and the general public to use the data for further analysis effectively and efficiently. The Astromaterials Data System (Astromat) is NASA's primary archive for laboratory analytical data of returned samples, and its mission is to maximize the ROI that were contributed to collect and analyze astromaterials samples by improving the findability, accessibility, interoperability, and reusability (FAIR) of the astromaterials samples data. Astromat is now creating new capabilities beyond FAIR that facilitate the reuse of the data as analysis-ready data (ARD), providing access to large-scale, fully harmonized and integrated collections of data via machine-actionable interfaces. Analysis-ready data enable the scientific community to apply computational methods including Machine Learning and Artificial Intelligence, overcoming the huge burden of 'data wrangling', i.e. consistently formatting, harmonizing, and integrating data, that researchers still have to perform in order to apply computational methods. In collaboration with researchers, Astromat is developing and improving its API (Application Programming Interface) and the structure and format of the data. We invite researchers to participate in the testing and improvement of the API and the delivery of ARD.

By providing FAIR and ARD for subsequent use, Astromat is fostering opportunities for the astromaterials samples data community and the general public to realize the ROI and the benefits of the extensive contributions by those who have developed and pursued capabilities for collecting and analyzing astromaterials samples to produce data products and services.