

MANAGING ANALYTICAL DATA FROM PRISTINE RETURNED SAMPLES IN COMPLIANCE WITH NASA'S DATA STRATEGY: THE ASTROMATERIALS DATA SYSTEM



80NSSC19K1102

K. Lehnert, L. Profeta, J. Mays

Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

- Analytical studies of astromaterials samples returned by NASA space missions generate unique and highly valuable data that contribute fundamentally to our knowledge and understanding of the origin and evolution of Earth, our solar system, and the universe.
- These data need to be openly accessible and curated in a manner that maximizes their reuse and utility for future science and that ensures their quality and long-term preservation.

Open Data

In 2013, the Obama Administration released an Executive Order that required federal government information to be open and machine-readable by default.

Obama issues Executive Order in support of open data

Timothy W. Vander, May 19th, 2013

SCIENCE MISSION DIRECTORATE POLICY
Scientific Information policy for the Science Mission Directorate

SMD Policy Document SPD-41

August 4, 2021

Executive Order on Open Data

In several recent strategic documents and reports, NASA recognizes the benefits of open data and is adjusting its science information policies. In 2021, the new Science Mission Directorate Policy SPD-41 states: "Results of federally funded research and development need to be shared openly in order to maximize the benefit and reach of the information. Data need not only to be archived but also to be curated – that is, the data are assured to have continued accessibility and usability for multiple decades."



- Applies to Publications – Data – Software
- ✓ produced by SMD-funded missions.
- ✓ produced by investigations funded via research awards.

FAIR Data

- SMD-funded data shall be made publicly available without fee or restriction of use.
- Data formats shall be machine-readable (i.e., data are reasonably structured to allow automated processing).
- SMD-funded data shall include robust, standards-compliant metadata that clearly and explicitly describe the data.
- SMD-funded data shall be reusable with a clear, open, and accessible data license.



FAIR Data

- ✓ should be accessible and understandable to humans & machines.
- ✓ should be deposited in certified trusted repositories to ensure persistent access & preservation.
- ✓ should be citable (registered with unique persistent identifier systems) and should have clear usage licenses.
- ✓ should be documented by rich metadata that support discovery and reuse.

AstroMat aims to facilitate discovery, access, mining, publication, and preservation of laboratory data acquired on astromaterials samples.

The Astromaterials Data System (Astromat) is a data infrastructure that curates, archives, and provides access to analytical data that are generated from astromaterials samples. Astromat

- preserves astromaterials data and ensures their long-term access and reusability for new science;
- restores legacy data of astromaterials samples acquired in the past;
- provides human- and machine-actionable interfaces to access comprehensive, analysis-ready data.

AstroMat Repository
Data Publication & Archiving

AstroMat Synthesis
Data Mining, Visualization, & Analytics

Users deposit data → Metadata Catalog + File system → Curators compile data → Fusion database (data & metadata) → AstroMat Repository → AstroMat Synthesis

AstroMat's systems, services, policies, procedures, and management have been designed to comply with the FAIR and TRUST* principles for repositories. AstroMat also follows best practices for laboratory analytical data established in the geochemistry community and FAIR physical samples that are emerging through various initiatives. (* TRUST = Transparent, Responsible, User Focused, Sustainable, Technology)

Apr 23, 2021

Planetary Data Ecosystem Independent Review Board Delivers Final Report to NASA

NASA's Planetary Data Ecosystem (PDE) Independent Review Board (IRB) delivered its final report on the information collected about planetary data. The IRB provided prioritized actionable recommendations for long-term planning and support of the Planetary Data Ecosystem (PDE).

In Nov 2020, the PSD commissioned the IRB to conduct a thorough review of the Planetary Data Ecosystem, which includes the user community and the organizations and facilities involved in planning, obtaining, analyzing, preserving, and sharing data from planetary space missions.

<https://science.nasa.gov/researchers/science-data>



Important IRB recommendations are already fulfilled by the Astromaterials Data System:

Recommendations of the IRB	Astromat Compliance
R9: NASA should seek CoreTrustSeal certification, and thereby WDS membership, for the PDS data nodes. NASA should encourage CoreTrustSeal certification for other PDE elements that serve as data repositories	Astromat has already completed its application for CoreTrustSeal Certification.
R33: NASA should establish a requirement for the preservation of mission-supported laboratory analyses of returned sample material that makes the information accessible to the planetary science community	Astromat provides the infrastructure and services that support the preservation and access of laboratory analyses of returned sample material.
R47: NASA should support and encourage expanded use of DOI-like identifiers for data, thereby connecting data at various levels of processing to assist users in locating the best version of a data set for their needs.	The AstroMat Repository registers data with DataCite to assign DOI. Astromat also is working to assign persistent, globally unique identifiers (IGSN) to the samples.
R49: NASA should fund the development of more analysis-ready data (ARD) products derived from the lower-level products created by NASA missions.	Astromat's Synthesis Database delivers astromaterials data analysis-ready via machine-actionable interfaces.