**THE ASTROMATERIALS DATA SYSTEM: TRANSFORMING ACCESS TO PLANETARY SAMPLE DATA.** K. Lehnert[1], K. Markey[1], P. Ji[1], C. Evans[2], and R. Zeigler[2], [1]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA, lehnert@ldeo.columbia.edu, [2]NASA Johnson Space Center.

**Introduction:** Sample science forms a critical base for planetary science and understanding solar system evolution. Samples of extraterrestrial materials - meteorites, cosmic and interplanetary dust, solar wind, and lunar rocks - have been studied and analyzed for decades, generating important new knowledge and discoveries. The Astromaterials Acquisition and Curation Office of NASA's Johnson Space Center (JSC) curates the world's most extensive and precious collections of extraterrestrial materials and makes them available to the global research community for scientific research. Thousands of geochemical, petrological, and geochronological studies have been conducted over time, producing a treasure trove of data that hold a huge potential for future discoveries. JSC and the EarthChem group at the Lamont-Doherty Earth Observatory have embarked on a major effort to build a comprehensive data system for all analytical data generated from JSC's astromaterials collections, the Astromaterials Data System (AstroMat). AstroMat will not only ensure long-term preservation and access of analytical data for specimens curated at JSC, but provide a context-rich searchable database of the data offering for the first time the ability to easily compare analytical data across diverse extraterrestrial collections, thus facilitating the study of comparative planetology.

**Objectives:** The ultimate goal is for AstroMat to become an indispensable data resource for planetary scientists, offering unprecedented opportunities to mine and analyze the vast number of analytical measurements that have been made on extraterrestrial samples from NASA missions. AstroMat is modeled after the geochemical databases PetDB, NAVDAT, and GEOROC, which have provided this type of service for many years and have demonstrated that new discoveries and new scientific paradigms such as statistical geochemistry become possible through the availability of these large data syntheses [1,2,3].

**Scope:** AstroMat is envisioned to serve as a comprehensive data management solution that ensures persistent online access, archiving, and exploration of analytical data generated by the study of all JSC astromaterial collections for the benefit of future science. Currently no equivalent data system exists for the astromaterial collections curated by NASA-JSC.

AstroMat will integrate past, present, and future data. Data will be compiled from the past and present literature and from any investigator contributions we will be able to solicit. AstroMat will comprise the entire MoonDB database of lunar sample data, which have been compiled over the past 4 years with funding from two NASA Awards of the PDART program (Planetary Data Archiving, Restoration, and Tools). All historical data from the Antarctic Meteorite, Cosmic Dust, Genesis, Hayabusa, Stardust, and Microparticle Impact collections will be added. Focus will initially be on geochemical, mineralogical, and geochronological data, but content will be extended later to include additional data types such as experimental petrology, images, and spectrography.

**System Components:** AstroMat is developed as an ecosystem of databases and software applications that help researchers, data curators, and developers track, manage, access, search and explore data and will comprise the following components:

*AstroDB:* Storage database (PostgreSQL) with rich and quality-controlled metadata content (analytical methods, sample lithologies, etc.) that integrates (fuses) data that were previously scattered across the scientific literature and in online PDF documents, making the data easily searchable and minable as a single data product with human and machine-readable interfaces. AstroDB uses a modified version of the Observation Data Model ODM2 [4].

*AstroAPIs:* A suite of Advanced Programming Interfaces (APIs) serves as a bridge between the storage database AstroDB and all applications that retrieve data from this database. Users can develop their own applications to access data via AstroAPIs.

*AstroSearch* and *AstroPages* provide interfaces for users to explore the content of AstroDB, select samples and data that they are interested in, view and browse them, and download them in a useful format. Sample type, sample lithology, sample composition, and analytical method are some of the criteria for selecting data. In the later phase of the AstroMat development we will add *AstroVis*, a suite of tools for interactive visualization of analytical data. Also planned is the capability to visualize in-situ measurements on thin section or microprobe mount images with direct links to the individual spot analyses.

*AstroRef* allows users to browse, search, filter, and sort a comprehensive bibliography of all references related to the JSC collections. This bibliography will include references both with data ingested into the database, and papers without data, in hopes of providing a valuable reference library to the community.

*AstroRepo* will provide repository services for user-submitted digital content from astromaterials research (analytical data, data synthesis, images, models, etc), supporting investigators in sharing their

data and complying with funders' and publishers' policies for FAIR data (Findable, Accessible, Interoperable, Reusable) [5]. Submitted data will be registered with Digital Object Identifiers (DOIs) and long-term archived, following the established procedures and policies of the EarthChem Library (ECL) (http://www.earthchem.org/library) that comply with international standards for Trusted Repositories.

*AstroAdmin* provides tools for AstroMat data curators to create, curate, organize, annotate, and manage highest quality data content including metadata and analytical data.

**Community Engagement:** In order for AstroMat to achieve its goal of empowering future research in planetary science, it is critical to establish a user community of researchers, who are aware of the data system, familiar with its capabilities and use, provide feedback on usability and advise on data content and functionality, recommend links and interoperability with other data resources and modeling tools, and, most importantly, who will contribute their data and encourage other investigators to do the same. AstroMat is planning to provide regular updates about the system development through a newsletter and on its web site (http://www.astromat.org), conduct regular webinars that will be recorded and made available online as tutorials, and organize regular user workshops at LPSC and other conferences as needed. A major focus will be engagement of Early Career Scientists (ECS) through workshops on data analytics, data management planning, and data publication. AstroMat is planning to support internships of ECS with planetary scientists as training opportunities in data management.

**Project Timeline:** The initial design and development of the Astromaterials Data System has been underway since late 2018. The AstroMat web site will be released by the end of February 2019 and will provide access to a catalog of references for all JSC astromaterials collections via a basic version of AstroRef that also allows users to monitor progress in data entry. The first version of AstroDB will be made available with a basic search interface in late 2019. The database will initially contain all lunar sample data from MoonDB. Data for the Antarctic Meteorite collection will be compiled and continuously added next, with a targeted date of completion in 2022. Data for cosmic dust, solar wind, and microparticle impact samples will be added last and completed by 2024. The AstroRepo is planned to become available in 2020. Users who want to deposit their data in a trusted repository before that date can use the EarthChem Library.

**References:** [1] Class C. and Goldstein S. L. (2005) *Nature, 436,* 1107-1112. [2] Gale A. et al. (2013) *Geochem. Geophys. Geosyst., 14,* 489-518. [3] Keller B. C. and Schoene B. (2012) *Nature, 485,* 490-493. [4] Horsburgh, J. et al. (2016) *Environ. Modelling & Software, 79,* 55–74. [5] Wilkinson M. D. et al. (2016) *Scientific Data, 3.*