LADEE PDS ARCHIVE: ACTIVE MISSION PIPELINE DEVELOPMENT USING PDS4

Lyle F. Huber[1], Lynn D. V. Neakrase[1], Shannon K. Rees[1], Jeffrey Jasek[1], Ernesto Gonzalez[1], Reta F. Beebe[1], Daniel J. Crichton[2], Gregory T. Delory[3], Carol Neese[4]
[1]Department of Astronomy, New Mexico State University, Las Cruces, NM 88003
[2]Planetary Data System Engineering Node, Jet Propulsion Laboratory, Pasadena, CA, 91109
[3]LADEE Deputy Project Scientist, NASA Ames Research Center, Moffett Field, CA 94035
[4]Planetary Science Institute, Tucson, AZ 85719

Introduction: The NASA Planetary Data System (PDS) is the distributed system of discipline nodes responsible for the archival of all planetary data acquired by robotic missions, manned missions, and observational campaigns through ground/space-based observation systems. The PDS has moved from version 3 to version 4 of its archival system. The first mission to archive under PDS4 standards will be the Lunar Atmosphere & Dust Environment Explorer (LADEE). LADEE is in the midst of pipeline development and will serve as a test case for the new PDS system. The instrument teams are working closely with their discipline nodes to adapt to the new changes and to help develop best practices for future data providers. The PDS discipline nodes involved in this effort include Atmospheres (ATM), Small Bodies (SBN), and the Navigation and Ancillary Information Facility (NAIF).

LADEE and PDS4: This mission provides the first opportunity to test out the end-to-end process of archiving data with an active mission into the new architecture of the PDS. The limited number of instruments producing data, all in simple data structures, makes for good candidates for testing.

The *Lunar Atmosphere & Dust Environment Explorer* (LADEE) [1,2] is a short mission that aims to study the exosphere and dust environment surrounding the Moon, investigating sources, sinks, and surface interactions as well as controls on the distribution and variability of the lunar atmosphere. The instrument package includes three main instruments providing data to the archive: Neutral Mass Spectrometer (NMS), UltraViolet Spectrometer (UVS), and the Lunar Dust Experiment (LDEX). Each of these instruments share heritage from past flown experiments and should provide "well-behaved" ASCII-Table data to exercise the new structure of the archive system.

Data Organization with PDS4: PDS4 will be implemented using eXtensible Markup Language (XML), which allows improved interfacing between users, the data, and the Internet. XML uses schema documents (analogous to blueprints) to determine the structure of the corresponding XML labels. In the case of PDS4, these schemas allow management of the labels and their content by forcing validation dictated by the underlying Information Model. The use of a central, underlying Information Model will be a vast improvement over PDS3 because of the uniformity it provides across all the discipline nodes.

Under PDS3, the organization structure revolved around the "Volume" for datasets [3]. A Volume was a logical grouping of data and accompanying documentation specifically designed to be delivered via physical media, so a volume became synonymous with the Tape, CD, or DVD it was written on. In PDS4, the motivation is to make data truly accessible across the Internet, with very little reliance on physical media [4]. The structure and organization of the data products allows for more user services as opposed to distributing data as "volumes". As a result, PDS4 will implement a product-centric approach for archiving data and supplemental documentation. Products can be organized into *Collections*, which are logical groupings of files. The Collections can be then organized into *Bundles*, in which all collections are logically related. An example of this structure would include an Instrument Bundle for a mission. Within that Bundle end users should expect to see various collections of documents and data products, organized into directory structures for the respective instrument and data types including all references for documenting the use and provenance of the data [4].

Another change under PDS4 will be how the archive is organized across discipline nodes. Replacing the PDS3 central catalog, will be a Central Registry, in which all products (including Bundles and Collections) will be registered and therefore accessible to search engines [4]. Because PDS4 is product-centric, and documents, data, cross-references, and other ancillary data are all products, everything will be registered within the system. Together with the XML implementation, the Central Registry will allow the search routines to be more complex and inclusive than they have been in the past. Better searching will help in providing uniformity across the PDS and better data coverage will lead to better user confidence in the system.

Process: The development of the LADEE archive under PDS4 has generally followed the development of archives for other recent missions. PDS discipline nodes and LADEE instrument team representatives worked together to identify the data products that LADEE would produce. Documentation describing the instruments and data products were produced by the instrument teams and reviewed by the PDS peer review process. XML templates for labeling the data products were developed by the PDS nodes and provided to the instrument teams to begin their pipeline work.

The first delivery to PDS will be in late March with a full peer review scheduled. The results of this peer review will be passed back to the instrument teams to make any corrections to their data products. Final delivery of the entire mission data is planned for June 2014.

Conclusions: Currently, LADEE instrument teams are in the process of developing pipelines for PDS4 data archiving. This process is fundamentally unchanged from PDS3, with a few notable exceptions. The XML process compared to past ODL work is quite a bit different, although the philosophy behind label development is nearly identical. The discipline nodes are working closely with

the teams to help with PDS4/XML expertise, and to document and construct future pipeline development protocols that ensure PDS4 compliance and ease of use for both the data providers and the end users. Implementation through XML Schema version 1.1 should allow streamlining of the schema-to-label process, and most of the discipline nodes are developing techniques to allow data providers not well-versed in XML to edit and design their labels with help from the nodes. This mission should be a good shakedown test of the new PDS4 system from start to finish providing a vital resource to the final development steps bringing PDS4 to the public.

PDS4 is the first step toward modernization of the archive and when fully implemented should make the archive a more usable tool for the data providers and end-users alike. Most importantly as the PDS makes the move to PDS4, the integrity and usability of all the data in its holdings will be assured, continuing the long tradition of making planetary data accessible to the public.

References: [1] Delory, G.T. et al. (2009) LPSC Abstract #2025; [2] NASA – LADEE Website: http://www.nasa.gov/mission_pages/LADEE/main/index.html; [3] PDS Standards Reference, version 3.8 (2009); [4] PDS4 Standards Reference, version 1.0.0, (2013);