LADEE & MAVEN: ACTIVE MISSION PIPELINE DEVELOPMENT USING PDS4

Lyle Huber[1], Lynn D. V. Neakrase[1], Shannon Rees[1], Matias Roybal[1], Reta Beebe[1], Daniel J. Crichton[2], Gregory T. Delory[3], Alexandria DeWolfe[4], J. Steven Hughes[2], Joseph Mafi[5]
[1]Planetary Data System Atmospheres Node, 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]MAVEN Science Data Center Manager, Laboratory for Atmospheric and Space Physics, University of Colorado at Boulder, Boulder, CO 80303
[5]Planetary Data System Planetary Plasma Interactions Node, University of California, Los Angeles, CA 90095-1567

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. Beginning late in 2012, the PDS will be publicly moving from version 3 to version 4 of its archival system. The first two missions to archive under PDS4 standards will be the Lunar Atmosphere & Dust Environment Explorer (LADEE) and the Mars Atmosphere & Volatile EvolutioN (MAVEN) spacecraft. The former is scheduled for launch in May 2013 with the latter scheduled for November of that same year. Both missions are currently beginning pipeline development and will serve as test cases for the new PDS system. The instrument teams will be 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), Planetary Plasma Interactions (PPI), Small Bodies (SBN), and the Navigation and Ancillary Information Facility (NAIF).

The Missions and PDS4: These two missions provide the perfect opportunity to test out the end-to-end process of archiving data with an active mission into the new architecture of the PDS. Each mission has a limited number of instruments producing data in a few data structures, that are already included in the initial release of PDS4, making them perfect 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.

The *Mars Atmosphere and Volatile EvolutioN* mission will be part of the NASA Mars Scout program scheduled to explore Mars’ upper atmosphere focusing on the ionosphere and interactions with the Sun and the solar wind. Of utmost importance is determining the role played by the loss of volatile compounds in the history of martian climate, liquid water, and potential habitability through time [3]. MAVEN will include 8 instruments providing data to the PDS. Of these, 6 instruments comprise the Particles and Fields package including the following experiments: Magnetometer (MAG), Langmuir Probe and Waves (LPW), Solar Wind Electron Analyzer (SWEA), Solar Wind Ion Analyzer (SWIA), Solar Energetic Particles (SEP), and SupraThermal And Thermal Ion Composition (STATIC). The other two instruments are the Neutral Gas and Ion Mass Spectrometer (NGIMS) and the Imaging UltraViolet Spectrometer (IUVS). This set of instruments will expand the range of data types being archived under the initial PDS4 mission effort, and will represent a more complex mission interface than the LADEE example. All MAVEN data will be in ASCII or binary tables (including a limited form of CDF), with the exception of IUVS, which will be in FITS format. Again much of the MAVEN mission is flying experiments with a rich heritage from previously flown instruments, presumably corresponding to data that are similar to sets already archived under PDS3.

Data Organization with PDS4: PDS4 will be implemented using eXtensible Markup Language (XML), which allows better 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 [5]. A Volume was a logical grouping of data and accompanying documentation specifically designed to be delivered via physical media, so volumes 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 [6]. 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 [6].

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 [4,5]. 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 should help in providing uniformity across the PDS and better data coverage should lead to better user confidence in the system.

Conclusions: Currently, LADEE and MAVEN 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. These first two missions 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 later this year (2012).

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] NASA – MAVEN website: http://www.nasa.gov/mission_pages/maven/overview/index.html; [4] PDS Standards Reference, version 3.8 (2009); [5] Data Preparers Handbook, version 4.0, *in prep.*;