# BETA TESTING THE PDS4 ARCHIVE: MARS PHOENIX REVISITED

Lynn D. V. Neakrase[1], Lyle Huber[1], Shannon Rees[1], Dylan White[1], Ernesto Gonzalez[1], Reta Beebe[1], Daniel J. Crichton[2], J. Steven Hughes[2], Sean Hardman[2]
[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

**Introduction:** The NASA Planetary Data System (PDS) is the distributed system of discipline nodes responsible for the archive of all planetary data acquired by robotic missions, manned missions, and observational campaigns through ground/space-based observation systems. Beginning late in 2013, the PDS will be publicly releasing version 4 (PDS4) of its archival system. Of greatest concern moving forward is the preservation of the integrity of older data sets, while achieving improved accessibility and streamlined processes for new data entering the archive. Migration of the older data is of the utmost importance while maintaining seamless usability during the transition. The first datasets to be completely migrated to PDS4 include the Atmospheres Node's Mars Phoenix Lander holdings, which include complete, straightforward tabular data from five instruments (Meteorology, Light Detection and Ranging, Atmospheric Structure Experiment, Atmospheric Opacity, and the Telltale Anemometer). Late in 2012, the Atmospheres Node redesigned the access websites for the Phoenix data to implement changes that would be available for the PDS4 migrated data. These access websites as well as basic structures within PDS4 were 'Beta-Tested' with a diverse group of end-users to introduce PDS4 and test a new organizational layout. The goals of the redesigned web pages were 1) to attempt to give end users a 'one-stop-shop' for newly migrated PDS4 data, 2) to improve access to the data and provide users with feedback on availability different types of data, and 3) to expand our usability to show interconnectivity between the Phoenix datasets and other Mars collections throughout the PDS archive.

**The Past (PDS3):** The PDS3 system implemented a label/product duet for each item within the archive [1]. The labels were implemented in the Object Description Language (ODL) developed and maintained by JPL/Caltech (used only by the PDS), which could be used as attached or detached labels. ODL allowed for a human-readable, "KEYWORD = VALUE" structure, that was not always the easiest for software to parse or use efficiently. Directory structures were organized specifically with physical media in mind (e.g., Magnetic Tape, CDs, DVDs, etc.) [1]. Data sets were organized into *volumes* that could easily be written onto physical media, which were designed specifically for transfer from archive to user (or user to user). Retrieval of data was based on search routines maintained by the individual discipline nodes and was not well-suited for overlapping datasets present within the archive in possibly piecemeal fashion between different nodes.

**The Future (PDS4):** PDS4 is the latest incarnation of the PDS archiving system. With PDS4, many of the perceived problems and shortcomings of the PDS3 system have been addressed. Organization and implementation are designed around the modern idea of all data being delivered to users across internet-based systems. PDS4 is an object-oriented system based on a central core Information Model, from which everything within the system is defined explicitly [2]. This differs greatly from past incarnations and provides continuity across discipline nodes, which has not been present in the past. The catalog system has been replaced by the new central registry, which allows more information to be ingested and tracked across the system. PDS4 is product-centric. A "product" is defined as a label file and the object (data, document, etc.) it describes. The registry allows metadata to be registered across the PDS, allowing better cross-referencing between various data products and between other discipline nodes. This approach also facilitates search and retrieval at the individual product level. The new system replaces the use of ODL (managed by JPL/Caltech, used only by PDS) with the commercially available eXtensible Markup Language (XML)[2,3]. XML is widely accepted as a modern standard for encoding data for use on the Internet by providing enhanced machine readability, focused on simplicity and generality [3]. XML allows commercially available software to use the PDS archive without extensive modification, which should allow for better, more wide-spread usage of PDS data. The first public release of the PDS4 system should be expected in Q4 of 2012.

**XML and the PDS:** As in many web-based markup languages, XML employs a tag-based system of designating meta-attributes used to distinguish differences in the data [3]. Tags simply bracket the assigned value:

<title>This is XML for PDS4</title>
*tag* *value* *end-tag*

The tags can be arranged in a specific fashion into a blueprint or *schema* for the entire label. The structure of this schema is determined and generated by the Information Model ensuring that 'valid' labels are 'consistent' with respect to the framework of the Information Model. This approach is a more rigorous approach than in PDS3, in which validation was done by a tool and not against a model.

**Contents Organization:** PDS4 products are organized into focused "Collections" with separate collections for observational data products, calibration products, browse products, document products, etc.

PDS currently identifies nine broad collection types. Related collections are then organized into logical Bundles. A Bundle, for example, might consist of the collections, which together contain all products from an instrument on a specific mission. Within the Bundle, there could be one collection for raw, calibrated, and reduced data products, or there might be three separate collections (one each for raw, calibrated, and reduced data products), a collection for calibration products, a collection for document products, and a collection for browse products.

**Web Design for Phoenix:** Web design for PDS4 data access is a delicate interplay of considering veteran and novice users that may or may not be directly familiar with the data for which they are searching. Contrary to previous iterations of the web interface for accessing the Phoenix atmospheric data, the revised design goals included expansion of the mission-related information with a focus on retaining both a simple organization of access points to navigate directly to the data products and a historical sense of the purpose and goals of the Phoenix mission as a whole and how that applies to the overlying context of the scientific investigation of Mars. Key web design enhancements included the mission logo and title bars as a 'wrapper' for the Phoenix mission section of the Atmosphere web pages, informational text and graphics for mission goals and instrument descriptions, sample visual plots of the tabular data for each instrument, and links to individual useful documents for use of PDS4 and the data themselves.

For more of a sense of these free to visit the Beta site at: http://atmos.nmsu.edu/PDS4BETA/phoenix.html. Sample exercises and score cards are available from the main site, should you have questions or comments for the developers.

**Beta Testing:** In late 2012, all five of the Phoenix atmospheric instrument bundles were migrated to the PDS4 standards and registered within the system. The web redesign was completed and two reviews were conducted. First an internal review allowed a few NMSU researchers not involved with PDS4 development to review the new design and time the proposed introduction/review exercises. They tested the system for obvious broken links and also determined the review should take between 45-90 min to complete. The Beta-site was then opened up to a diverse group of users for more general review and a realistic test of the new organization. This phase utilized around 35-40 reviewers who were asked to complete a worksheet and scorecard to provide feedback on their experience with the newly designed web access. The worksheet provided a brief introduction to the new organizational concepts and the XML labels, and then requested the users work through two brief exercises. The first focused on gaining familiarity with the new layout, PDS4 help documents, Bundle and Collection organization. The second allowed users to compare the new PDS4/XML labels with the PDS3/ODL counterpart as well as a simulated output from translation software that provides PDS3-looking text versions of the PDS4/XML labels mainly for human readability as users become more familiar with XML.

Reviews for the new system were received throughout January 2013, and have provided useful insight for the development of ways to use and access the data. In general the 'One-Stop-Shop' approach has been well-received, and most users have found the organization of the Phoenix pages themselves to be a vast improvement over the old system. However, the use of XML is of concern for some users mainly because the tag-system is perceived as being cluttered and less human-readable. This point is directing efforts to make the online presentation of the labels more appealing and useful for many users. Color-coding or other formatting of a translated text file version of the XML labels might accomplish this in the short term by highlighting the important metadata and de-emphasizing the XML structures.

PDS4 development moving forward might face community acceptance hurdles given the general user's lack of familiarity with XML. However, two approaches can help to mitigate problems with users regarding XML. PDS may have to provide more sophisticated web interfaces for displaying and using the XML labels. One side-effect may include more online help tutorials for users wanting to interface existing data reduction software or image processing routines to the new label structures. In this regard, clearly organized self-consistent web pages will be key to educating the users on both the new capabilities of PDS4 and how to use XML efficiently to their benefit. As PDS4 development moves forward and more of the PDS archive is migrated into the new format, more online services will become available, and the system will be further tested to ensure user are addressed to make PDS4 a useful tool for planetary science.

**References:** [1] PDS Standards Reference, version 3.8 (2009); [2] Data Preparers Handbook, version 4.0, *in prep.*; [3] Extensible Markup Language, http://www.w3.org/XML, (2011)