

CONVERTING PDS3 SCIENCE ARCHIVES FROM ACTIVE PLANETARY MISSIONS TO PDS4. Susan Slavney¹, Raymond E. Arvidson¹, Joy A. Crisp², ¹Planetary Data System Geosciences Node, Earth and Planetary Sciences Dept., Washington University in St. Louis, St. Louis, Missouri, 63130, Susan.Slavney@wustl.edu, arvidson@wunder.wustl.edu, ²JPL/Caltech, 4800 Oak Grove Drive, Pasadena, CA 91109, joy.a.crisp@jpl.nasa.gov.

Introduction: The NASA Planetary Data System (PDS) has worked with planetary missions to help them design and deliver quality science data archives since the 1980s. Some currently active missions have been accumulating archives for more than a decade. During this time the PDS archive standards have evolved, but these older missions continue to use the standards that were in place when they began. The missions have now been asked to plan the conversion of their archives to the new PDS4 standards to ensure the long-term value of the archives. PDS is helping the missions create and implement the conversion plans.

What is PDS4?: PDS4 is a revision to the PDS archive standards, replacing PDS3, which had been in place for many years. PDS4 is designed to produce higher quality, more accessible archives using modern software capabilities [1]. Like PDS3, PDS4 specifies the contents of a quality archive, including data, documentation, and ancillary material, with metadata for each product stored in a PDS4 label. PDS4 data types are fewer, simpler, and more rigorously defined compared to PDS3. While PDS3 labels used a keyword=value format, PDS4 labels are written in XML (eXtensible Markup Language), enabling more complete and consistent metadata, and allowing the use of existing XML software tools. Planetary missions confirmed for flight after November 1, 2011, are required to archive in PDS4. Active missions that were archiving in PDS3 before that date have continued to use the PDS3 standard [2].

Request To Convert PDS3 to PDS4: Active planetary missions with PDS3 science archives, including Mars Odyssey, Mars Exploration Rovers, Mars Reconnaissance Orbiter, Lunar Reconnaissance Orbiter, Mars Science Laboratory, Juno, and Dawn, along with the recently completed MESSENGER mission, have been asked by NASA HQ to develop plans for converting (migrating) their science archives to PDS4. The plans are to be developed in consultation with the PDS discipline nodes that host the archives.

PDS Policy for Scheduling Migrations to PDS4: PDS considers the following points when determining the priority and schedule for migrating PDS3 data sets to PDS4 [3].

- The uniqueness of the archives, i.e., from missions to bodies for which data sets will not be reacquired within a foreseeable period of time,

- Level of user demand for PDS4 versions of the archives,
- Availability of data providers to help in migration, as they are the experts on their data sets, and PDS node personnel, who are PDS4 experts, to help in migration planning and/or implementation,
- Availability of archiving and analysis tools for use by the community that are specific to PDS3, and the difficulty and timing of developing tools that work with PDS4,
- Evaluation of impact on provider production line (will PDS4 work put the current delivery schedule at risk?),
- Availability of funds for both PDS node personnel and data providers, as appropriate.

How PDS Is Helping Missions Plan: The PDS nodes who archive the data from a mission create a preliminary assessment of the level of effort required to convert each data set. The PDS nodes and the science teams' data providers then meet to scope the effort in detail and estimate the costs. The mission compiles the results in a report to deliver to NASA HQ.

Scoping the Conversion Effort: To estimate the effort required to convert a data set to PDS4, the following questions should be considered.

What data sets should be converted to PDS4? In addition to the standard science data products, are there ancillary data such as browse images and calibration data that should also be converted?

For a given data set, which category does it fit?

- A. PDS4-compliant data format, with PDS3 keywords that map easily into a PDS4 label,
- B. PDS4-compliant data format, with some data dictionary work needed to create PDS4 labels,
- C. Data that are not in a PDS4-compliant format.

If a data set is in category B, what dictionary work is needed? If it is in category C, what needs to be done to make the data PDS4-compliant?

What software will need to be acquired, revised, or created? What documentation will need to be revised or created? Who is best positioned to do the work; i.e., the science team, the PDS node, or a third party?

In most cases, the conversion effort will require the combined expertise of the PDS node, the science team, and mission ground data system engineers. If a PDS4 data dictionary specific to the mission is required, which is likely, then the PDS nodes must coordinate

this work across all the mission's data sets. For these reasons a science team should not attempt to convert their data in isolation from the rest of the mission and PDS.

Answers to Common Questions: The process of planning the conversion to PDS4 is still in the early stages for all the missions mentioned above. The following questions have come up frequently during the initial discussions.

Should science teams on active missions start converting to PDS4 right away? No. At present the focus should be on planning and costing the work, not actually doing it. PDS recommends postponing the conversion effort until the end of operations, so as not to detract from the critical work the science teams are doing currently. That said, some teams may have specific reasons for wanting to switch to PDS4 mid-mission. The advantages and disadvantages of doing so should be discussed during the planning meetings described above.

Should all of a team's PDS3 archives be converted to PDS4? Yes, unless there are good reasons not to do so. The cost of converting a data set in a non-PDS4-compliant format may be prohibitive, for example. A data set that gets little use may not be worth even a moderate amount of effort. Rather than assigning each data set a status of yes or no, it may be more useful to assign it a priority as well as a level of effort, and then schedule the conversion accordingly.

It is not necessarily true that the highest priority data sets should be scheduled for conversion first. For example, a popular data set with a widely used set of PDS3-based software may be a high priority candidate, but the level of effort may be reduced by waiting until better PDS4 tools for users become available, or until similar new PDS4 data sets are available to use as a model.

What about data users who prefer to continue using the PDS3 data? The existing PDS3 data archives are not going away. There is nothing to prevent users from continuing to use the data already in the PDS3 archives.

Conclusion: Working together, PDS and the missions are planning the efficient and cost-effective conversion of data sets to PDS4 to ensure their longevity and usefulness to the science community.

References: [1] PDS4 Information, <https://pds.nasa.gov/pds4/about/>. [2] Policy on Transition from PDS3 to PDS4, <https://pds.nasa.gov/policy/PolicyOnTransitionFromPDS3toPDS411162010.pdf>. [3] Policy for Scheduling Migrations of Data from PDS3 to PDS4 Archival Formats, <https://pds.nasa.gov/policy/PolicyOnPDS3to4MigrationSchedule08302017.pdf>.