PDS Support for the Democratization of Space. L. B. Dahl[1], R. G. Deen[1], J. H. Padams[1], P. M. Ramirez[1], [1]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, Luke.Dahl@jpl.nasa.gov

Introduction: Shifting trends in planetary exploration will require new approaches to engaging and supporting deliveries into PDS. Key changes are the proliferation of CubeSat and SmallSat missions and instruments, the Moon to Mars program emphasizing commercial and international partnerships, and an increase in PDARTs. Common characteristics of these changes are their limited budgets and a lack of experience delivering to PDS. Enabling adherence to the new Planetary Data System standard (PDS4)[1] is crucial to maximizing scientific value for these lower-cost activities.

This “white paper” poses the question: what steps are necessary to be prepared to support the democratization of space and these new providers in the two to five-year timeframe? How do we make seemingly esoteric processes and requirements for archiving data as easy to digest as it is to build and deploy a CubeSat at a University?

The Foundational Change: Lunar and interplanetary space exploration is no longer limited to a handful of countries with deep pockets and longstanding resources to support space exploration. Hardware miniaturization and the commoditization of key components, along with standards and shared use of launch vehicles has reduced entry costs from hundreds of millions to as low as 10s of millions for a Lunar or interplanetary mission. This puts them within the range of universities and other smaller organizations.

For example, the Mars Cube One (MarCO) CubeSats are a class of spacecraft based on a standardized small size using modular use of off-the-shelf technologies. Many CubeSats have been made by university students, and dozens have been launched into Earth orbit using extra payload mass available on launches of larger spacecraft.[2] Thirteen CubeSats will be sharing a ride on the next generation Space Launch System (SLS) launching next year, traveling to near earth asteroids, the moon, and various earth and solar orbits to support mission criteria. Many of these CubeSats are being developed by Universities and commercial entities.

Additionally, scientific data capture and analysis is undergoing a transformation from gigabytes to terabytes of science data generated on a typical day or observation. Combining the volume of data with advancements in machine learning and computational capabilities results in the generation of new products from historical datasets. This is evident across the Planetary Data Archive, Restoration, and Tools (PDART)[3] and Research Opportunities in Space and Earth Science (ROSES)[4] Discovery Data Analysis programs. These activities are creating value-added datasets from existing mission data on shoestring budgets.

The Challenge: Creating a science data processing pipeline that prepares, validates and delivers archival-ready products in a PDS-compliant bundle is not a trivial activity. The knowledge, processes and technology required to enable compliance can be costly and time-consuming. This isn’t problematic for projects and instruments from NASA centers and similar institutions that have amortized the startup and ongoing costs over decades of projects and familiarity with PDS nodes and processes. The challenge lies in enabling PDS compliance for the next generation of projects – likely a much higher volume of projects funded at much lower levels, with a compressed schedule and a lack of experience with PDS delivery requirements.

Potential Solutions: There is no perfect solution to this problem. There are, however, approaches that will address many requirements that would serve to mitigate the challenges described. Eliminating the need to fully understand the PDS4 standard to create a compliant label is the first step. The PDS Label Assistant for Interactive Design (PLAID)[5] is available for use now and does just that. PLAID distills hundreds of pages of reference guides and XML format knowledge down into an intuitive Web application to allow a layperson to create a compliant mission/project/product label schema. It also allows for reuse of previously generated and approved schemas to improve adherence and commonality to standards. PLAID is continually updated with the latest PDS4 information model to assure new labels are adhering to the latest standards.

The next key enabler would be to provide a reference data pipeline processing system. It would be pre-configured to consume the output from PLAID (the schema) as well as providing a drop-box like interface to upload and process products. It would include templates and predefined workflow tasks for most common activities, such as processing products, submitting samples to support peer reviews, and packaging and submitting bundle deliveries. The system could be delivered through a cloud offering and made available on an as-needed basis. The AMMOS PDS Pipeline Service (APPS)[6] offering could readily serve as the foundation for this capability.

This solutions described will not require a significant investment in new development or creating new capabilities. This is simply improving the packaging, coordination and delivery of existing products to the emerging community of data providers. The opportunity to transition from tool provider to software-as-a-service provider for these fundamental activities is enabled through the emergence of cloud-hosted environments coupled with automated packaging and deployment offerings.

The Benefits: Both data providers and PDS nodes reap the benefits of the proposed solutions. Data providers receive an efficient and easy to use approach to meeting PDS requirements. They are able to focus on the complexities of developing and fulfilling their mission objectives and expend far less effort on the mechanics of complying with PDS requirements. They are also able to rapidly iterate through label schema design and data processing needs throughout their project phases without being reliant upon dedicated infrastructure. This reduces uncertainty and improves cost estimation for archival requirements.

PDS nodes will benefit as the new data providers will be using sanctioned and integrated solutions that will eliminate typical support needs for a new provider. Though documentation exists that enables new providers to familiarize themselves with the standards and processes, nodes spend a disproportionate amount of time answering questions and assisting providers that simply won't read them. This approach also facilitates the implementation of evolving standards and processes to support PDS needs. Those standards and processes will only need to be updated within the PLAID and APPS systems in the future – effectively encapsulating PDS standards and requirements within the applications.

Conclusion: PDS needs to be prepared to support an increase in volume and diversity of emerging archival data providers. These providers will more likely be comprised of grad students and early career hires than science data system engineers with decades of experience working with archival products. We need to design solutions to support their needs and align with modern expectations for easy to use solutions.

References:

[1] https://pds.nasa.gov/datastandards/about/
[2] https://www.jpl.nasa.gov/cubesat/missions/marco.php
[3] https://nspires.nasaprs.com/external/solicitations/summary.do?solId={CE125B0A-E59C-218C-5B96-F3FA97F6A6FF}&path=&method=init
[4] https://nspires.nasaprs.com/external/solicitations/summary.do?solId=%7B334D0E7E-1D79-0E19-BEAA-CD9BA2A38A47%7D&path=&method=init
[5] S. S. Algermissen, J. H. Padams, C. Radulescu, *3rd Planetary Data Workshop (2017), Abstract #7030* https://plaid.jpl.nasa.gov/
[6] C. Radulescu, S. R. Levoe, S. S. Algermissen, E. D. Rye, S. H. Hardman, J. S. Hughes, M. D. Cayanan, E. M. Sayfi *Second Planetary Data Workshop (2015), Abstract #7050* https://ammos.nasa.gov/pdf/Brochure-IDS-APPS.pdf