PDS4 Training: Key Concepts and Vocabulary

7089

M. K. Gordon[1], E. A Guinness[2], L. D. V. Neakrase[3], J. Padams[4], A. C. Raugh[5]
[1] Carl Sagan Center, SETI Institute, Mountain View, CA (mgordon@seti.org), [2] Department of Earth and Planetary Sciences, Washington University, St Louis, MO. [3]Department of Astronomy, New Mexico State University, Las Cruces, NM, [4]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, [5]University of Maryland, College Park, MD

The PDS has moved to a new archiving standard, PDS4. This new standard is a model-driven system that simplifies data formats and stores more extensive metadata in XML labels. The resulting archive has improved connections between data products, supports more complex, faster data searches, and delivers an improved user experience.

1

PDS4 Provides a Hierarchical Organization for Data Archiving:

Bundle Product – A list of all related collections.

Collection Product – A list of related basic products of similar type (all raw images from a single instrument, all documents from a mission).

Basic Product – The smallest unit of data registered and tracked in the PDS (images, tables, documents, SPICE files).

2

PDS4 uses Four Fundamental Data Structures:

Array – A homogeneous n-dimensional array of scalars (images, spectral cubes).

Table – A set of repeating heterogeneous records of scalars (binary and/or character fixed-length tables).

Parsable Byte Stream – Don't need special software to read (text files, XML files, CSV tables).

Encoded Byte Stream – Need special software to read (PDF files, JPEG images). Not used for science observations.

3

Unique Identifiers and Cross References:

- Every PDS4 product – bundle, collection, or basic product, has an XML label and is registered in the PDS Central Registry.
- The XML label of every product contains that product's Logical Identifier (LID) and Version Identifier (VID).
- The combination of LID & VID, a LIDVID, provides a single, globally unique identifier for each product.
- PDS uses LIDs and LIDVIDs to cross-reference products in the PDS registry system.

4

First Steps for Data Providers:

- Review the PDS Data Providers Handbook and other information on the PDS website (https://pds.nasa.gov/pds4/about/).
- Talk to your PDS Discipline Node (DN).
- Outline your bundle – determine what products you will deliver and which collections you will need. Discuss with the DN.
- Develop your LID formation algorithms and get them approved by the DN.
- As soon as practical produce sample products and submit for DN review.

5

LID Formation:

- Constructed using four (bundle), five (collection), or six (product) fields.
- Fields are separated by colons. Colons may not be used within a field.
- For submissions to PDS, the first three fields are urn:nasa:pds. Similar construction is used for submissions to other planetary archives (e.g., urn:esa:psa, urn:jaxa:jaxa).
- Basic product IDs are constructed by appending to the parent collection's ID, and collection IDs are constructed by appending to the parent bundle's ID:

  urn:nasa:pds:<bundle_id>

  urn:nasa:pds:<bundle_id>:<collection_id>

  urn:nasa:pds:<bundle_id>:<collection_id>:<product_id>

LIDVID Formation:

- Append the Version ID (VID) to a LID separated by two colons
- For example: urn:nasa:pds:<bundle_id>::<version_id>

6

Resources for this meeting's PDS4 Training Workshop

- Downloads and instructions at:

  https://pds.nasa.gov/pds4/training

Other Key Resources

- The PDS4 wiki, hosted at the Small Bodies Node

  http://sbndev.astro.umd.edu/wiki/SBN_PDS4_Wiki

- The PDS4 home page, hosted at the Engineering Node

  https://pds.nasa.gov/pds4/about

- The PDS Tools Registry

  https://pds.nasa.gov/tools/tool-registry

Contact: Mitch Gordon:
mgordon@seti.org