



Enabling Scalable FAIR Data Archiving for High- Cadence Commercial Planetary Missions

Madison N. Hughes, Ph.D.
J. G. Ward, T. C. Stein

PDS Geosciences Node • Washington University in St. Louis

mnhughes@wustl.edu



The CLPS Landscape

4

Lunar missions
launched since 2024

17+

Instrument bundles
archived or in development

5

PDS nodes
coordinating

NASA's Commercial Lunar Payload Services (CLPS) program has shifted lunar exploration from infrequent flagship missions to frequent, commercially-operated deliveries, each carrying multiple NASA science instruments to the Moon.

This pace demands more scalable, streamlined processes, with less hands-on oversight.

Supporting FAIR Data Principles by Archiving in PDS



Findable

PDS4 XML labels with structured metadata, persistent Logical Identifiers (LIDs), and DOIs assigned.



Accessible

Open public distribution through PDS. No access restrictions, no paywalls. Data nominally online within ~6 months of mission end.



Interoperable

Standardized PDS4 data formats and dictionaries. Common label structures and schemas across CLPS missions.



Reusable

Detailed documentation, external peer review validating scientific usability, and clear provenance metadata.

Our Approach: A Scalable Framework

Beginning ~6 months before launch, PDS engages with project scientists and instrument teams to establish a structured, repeatable archive development process. **CLPS Data Archive Support Service (DASS)**



Standardized Resources

- Task rubrics & checklists
- Example products and labels
- Documentation templates
- Suggested timelines



Proactive Engagement

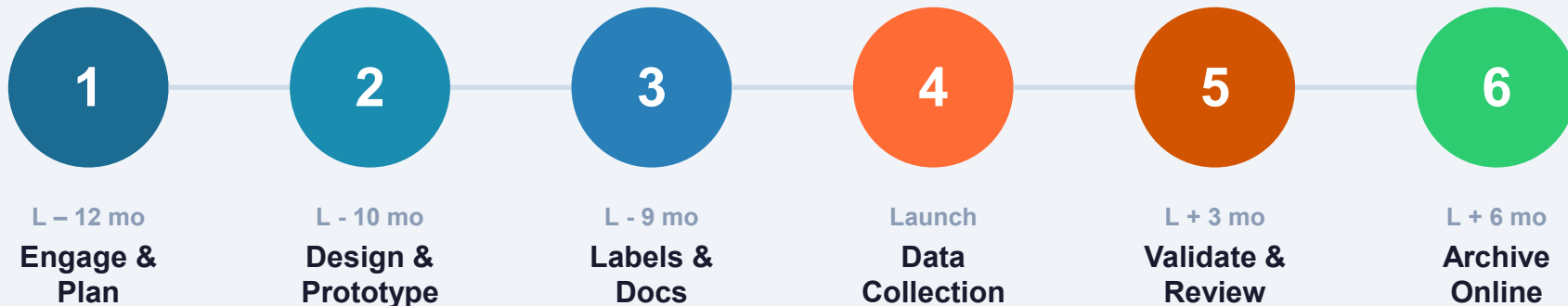
- Regular tag-ups with teams
- Cross-node coordination
- Project scientist liaison
- Iterative feedback loops



Defined Milestones

- LID & DOI reservation
- Archive design review
- Draft product upload
- Peer review scheduling

Nominal Archive Development Lifecycle



Pre-Launch

- Assign LIDs & reserve DOIs
- Context product creation
- Create label templates
- Update data dictionaries

Pre- and Active Mission

- Teams upload draft products
- Iterative PDS feedback
- SIS development
- Create label pipeline

Post-Mission

- Validation with PDS4 tools
- External peer review
- Final corrections & release
- Public distribution







The Collaboration Environment

Task Rubric System

Each instrument team receives a standardized task rubric with defined steps, deadlines, and status tracking. Rubrics are shared documents that PDS nodes and project scientists can monitor.

Register bundle LIDs	Done
Reserve DOIs	Done
Create label templates	In Progress
Submit draft bundle	Pending
Complete peer review	Pending

Centralized Resources

-  PDS4 label & data product examples
-  Archive development schedule
-  Validation tools & instructions
-  References to relevant sections of the DPH
-  Documentation templates
-  Context product templates

Coordination Across PDS Nodes

Each CLPS mission distributes instruments across multiple PDS nodes. Regular cross-node meetings (CLPS DASS meetings) with project scientists keep all parties aligned and accountable.

Example: TO_19D — Firefly Blue Ghost 1 (Landed March 2025)

GEO Node

LISTER, LMS (shared),
EDS, Misc. Data Buy

IMG Node

SCALPSS, RAC,
LPV, Lander Camera Data
Buy

PPI Node

LMS (shared),
RadPC

NAIF

SPICE Kernels

Why This Is Challenging



Time Management

6 months from landing to first (and usually only) data release. Multiple missions on overlapping schedules. Every bundle needs to be peer reviewed.



Delayed Engagement

Most instrument teams and project scientists don't start engaging with the CLPS DASS system until 6 months before launch.



Variable Team Experience

Some teams have pipelines ready for label generation and prior PDS experience. Others need significant guidance, and engagement levels vary widely.



Limited Team Resources

Instrument teams can sometimes be quite small, with only a few people in charge of both instrument operations and developing the archive.

The Main Hurdles

Teams Not Taking Advantage of the Resources and Task Rubrics

Most teams seem to not use to the task rubrics or be aware of the resources available to them.

Metadata Unknowns with Data Buys

Commercial operators may not be equipped to provide complete metadata. PDS end up needing to make adequate documentation for a large csv with *many* unknowns.

Variable Experience levels with PDS Archives

Some instruments have very little experience with creating a PDS4 archive and creating XML labels for products.

Deadline Slippage

Draft bundles arrive late. Peer reviews don't start on schedule. Personnel changes, leave, and various external factors create delays.

What's Working



Early Engagement Pays Off

Teams that engaged early with their PDS node delivered on or near schedule. The framework front-loads coordination to avoid late surprises.



Project Scientist as Lever

When PDS can't reach a team, the project scientist can. Regular DASS meetings create accountability that email alone cannot.



Iterative Feedback

Draft product uploads catch label and structural issues months before peer review. Teams that use the feedback loop need fewer major bundle corrections.

Looking Ahead

Scaling Up

- More CLPS missions in the pipeline (TO_CT3, CP-11, TO-20A, and beyond)
- Monthly CLPS launches in the future?
- Additional commercial providers with different metadata practices
- Commercial data buys may become more routine. Standardized processes needed

What Would Help

Revamping the task rubrics to make them better fit realistic CLPS timelines

More engagement of the instrument teams with the task rubrics

Developing more templates to make it as easy as possible for teams to create PDS4 compliant archives

Minimum metadata standards for commercial data buys

Key Takeaways

- 1 High-cadence commercial missions demand proactive, standardized archiving approaches.
- 2 Structured frameworks with clear milestones and regular coordination meetings keep teams on track.
- 3 Challenges remain. Making the task rubric timeline more accurate and developing more user-friendly resources would help.
- 4 No one can use the data if they don't understand it. Sufficient documentation and metadata is VERY IMPORTANT!

Questions?

mnhughes@wustl.edu

[1] Wilkinson et al. (2016) Scientific Data, 3(1).

[2] PDS Standards Reference, v. 1.25.0