PLANETARY DATA ECOSYSTEM CHIEF SCIENTIST REPORT

Planetary Data Ecosystem Report

Presented by
Moses Milazzo (he/him)
PDE Chief Scientist

PLANETARY DATA ECOSYSTEM CHIEF SCIENTIST REPORT

Presented by
Moses Milazzo (he/him)
PDE Chief Scientist

Planetary Data Ecosystem Independent Review Board

NASA's Planetary Science Division chartered the Planetary Data Ecosystem (PDE) Independent Review Board (IRB) in October 2020.

- The IRB's charter directed the team to:
  - define the full environment
  - identify missing or overly redundant elements
  - provide findings and prioritized recommendations
- NASA Announced selection of PDE Chief Scientist (16 Dec 2021).
- CS work began late winter / early spring.
- CS is an Independent Contractor, not a NASA employee

The final report is available at:
https://science.nasa.gov/researchers/science-data
(scroll to the bottom on the page)

PLANETARY DATA ECOSYSTEM CHIEF SCIENTIST REPORT

(Incomplete) History of the Planetary Data Ecosystem

1963 – USGS Astrogeology Science Center established

1966 – NASA Space Science Data Coordinated Archive (NSSDCA) established

1977 – NASA Regional Planetary Image Facilities (RPIFs) established (sunset in 2020)

1982 – National Academy of Sciences Committee on Data Management and Computation (CODMAC) chartered

1985 – NASA's Advanced Multi-Mission Operations System (AMMOS) initially developed

1989 – NASA Planetary Data System (PDS) established

1998 – NASA Center for Near Earth Object Studies (CNEOS)* established

2005 - NASA Data Analysis Program (e.g. DDAP, MDAP) established

2014 – NASA Planetary Data Archiving, Restoration, and Tools (PDART) program created

2014 – Mapping and Planetary Spatial Infrastructure Team (MAPSIT) established

2016 – Idea of Planetary Data Environment/Ecosystem starts to take hold at NASA HQ

2021 – Planetary Data Ecosystem (PDE) Independent Review Board (IRB) report

The Planetary Data Ecosystem today

PLANETARY DATA ECOSYSTEM CHIEF SCIENTIST REPORT

Physical Facilities	Archives	Online Repositories, Registries, and Portals	Public Communications
• NASA Center Archives • Spectroscopy Labs • Astromaterials Acquisition and Curation	• PDS • CNEOS • HITRAN • IRSA • MAST • NSSDCA • STI • SPDF • ESA's PSA	• AstroMat • ADS • Institutional Repositories • AHED • NASA Exoplanet Archive • NASA Open Data and Software Portal • Scientific Journals	• Active Mission Websites and Communications • Raw Images from Active NASA Missions • Instrument-specific Websites • NASA Photo Galleries • Science Nuggets • Treks • News Media Information
Sources of New Data	**Data Standards**	**Software**	**K-99+ Education Portals**
• Missions • DAPs • PDAR(T) • R&A • Scientists • NASA's • Domestic and International Partners	• PDS • FGDC • IAU • ISO • FAIR (access) • Open (access) • …	• AMMOS • Autopilot • ISIS3 + 4 • JMARS • SBMT • ASP • …	• Challenger Centers • NASA Kids' Club • NASA Science Space Place • NASA STEM Engagement • Universities • Junior and Comm Colleges • Citizen Science

Core Values

The IRB developed a set of core values

- **First, do no harm:** Avoid the law of unintended consequences.
- **FAIR:** Facilitate participation in the PDE by adhering to FAIR data principles of Findability, Accessibility, Interoperability, and Reusability.
- **Open:** Advocate open science practices, including open access, open data, open code, open software/tools, and others.
- **Collaborative:** Encourage international collaboration. Welcome new participants from both inside and outside the professional space exploration community.
- **Effective:** Provide timely, useful support to user communities, especially data producers.
- **Practical:** Pursuit of ideal solutions may sometimes leave the Ecosystem with no solution at all rather than a solution that is sufficient.

Planetary Data Ecosystem Chief Scientist Report

Presented by Moses Milazzo
PDE Chief Scientist

The PDE IRB report is available at:
https://science.nasa.gov/researchers/science-data
(scroll to the bottom on the page)

67 findings and 65 recommendations were organized into 5 themes:

- The Planetary Data Ecosystem Concept
- Planetary Data Stewardship
- Systemic Barriers to Data Preservation
- Barriers to Access and Usability
- Barriers to Development

+ a concluding “Pathway Toward an Ideal State” section

PLANETARY DATA ECOSYSTEM INDEPENDENT REVIEW

The highest priority recommendations fell into three groups:

- Develop the Ecosystem
- Address Data Preservation Needs
- Address Barriers to Use and Development

Develop the Ecosystem

Develop the Ecosystem

Plain language findings and recommendations summary

- PDE is a good idea and should be formalized
- PDE ≠ PDS; it is much broader
- Lack of communication among PDE elements causes inefficiencies and data losses
- NASA is an important leader, but it could lead more effectively by participating more in established communities
- NASA needs to learn from non-planetary communities to increase accessibility and use of planetary data
- NASA should work with international partners to expand access to the PDE and to learn from their successes

Develop the Ecosystem High Priority Recommendations

Group 1: Develop the Ecosystem

- ✓ Establish a sustained, community-led coordinating organization for the PDE that mirrors the other Planetary Assessment or Analysis Groups, reports to the Planetary Science Advisory Committee, and meets regularly.
  NASA Chose to establish the position of the Chief Scientist of the PDE.
- ➢ Refine the full scope of the Planetary Data Ecosystem and build community consensus around it. The responsibilities, accountabilities, governance, and service levels for elements of the Ecosystem that are funded by NASA Planetary Science Division should be clearly defined.
  This is ongoing.
- ➢ The prioritized goals and scope of PDS need to be carefully and explicitly defined and clearly articulated to the community. The differing responsibilities and expectations of the data preservation mission versus distribution of usable data need to be clarified. PDS should not be given unfunded mandates.
  This is ongoing.

Data Preservation Needs High Priority Recommendations

Group 2: Address Data Preservation Needs

- Establish an archive for planetary radar data either within the PDS Small Bodies Node or separately. Time is of the essence to prevent irretrievable data loss.
  This is ongoing.
- Establish a requirement for the preservation of mission-supported laboratory analyses of returned sample material. Require data preservation with appropriate metadata in an approved archive or repository for data produced by laboratory analysis of returned samples supported by ROSES Data Analysis Programs.
  This is ongoing.

- Establish a carefully crafted strategy to identify and prioritize the data preservation needs of the planetary science community that are not currently being addressed.
- Consider ways of archiving outside of the PDS that are amenable to creating FAIR and standards-based archives of these growing data sets.

Barriers to Use & Development High Priority Recommendations

Group 3: Address Barriers to Use and Development - 1

- Include early funding for mission data acquisition, processing, and archiving of data and foundational data products so that they are planned well in advance of data acquisition.
- Training and outreach
  - ➢ Develop outreach to user communities within the Planetary Data Ecosystem, assess user needs, and develop focused educational and documentation materials that meet highest-priority needs.
  - ➢ Provide regular, accessible, and effective training programs for researchers, data producers, mission specialists, and others who need to archive with the PDS. Address data preparation from the perspective of reusability and interoperability, such as the Earth Science Data Systems Working Group (ESDSWG) Data Product Development Guide (DPDG) for Data Producers.
  - ➢ Expand opportunities for intermediate to advanced technical training in topics related to accessing, using, and processing planetary data.

Barriers to Use & Development High Priority Recommendations

Group 3: Address Barriers to Use and Development - 2

- Support the delivery of higher-level and analysis-ready data products in well-documented and broadly used protocols and formats.
- Broaden support across the Ecosystem for a wider variety of data and information formats, such as engineering data; model input and output data; sound data; and physical collections.
- Expand intra- and inter-agency and international efforts to ensure that best practices, lessons learned, and appropriate technologies are shared and implemented across Planetary Data Ecosystem elements.
- The Planetary Data Ecosystem should regularly assess the Findability, Accessibility, Interoperability, and Reusability (FAIR) of data across each PDE element for machine-actionable access to data.

Barriers to Use & Development High Priority Recommendations

Summary:

- The concept of the PDE is meant to acknowledge that we have a wide diversity of people, data, data types, ways to store and preserve data, ways to access data, etc.
- Along with the richness of the wide diversity in the PDE comes a challenge of making sure everyone has access to these data… in perpetuity.
- I am working to lay the foundation for future Chief Scientists of the PDE to help NASA address this fundamental challenge.
- All of this needs your help.