Planetary Data Ecosystem

NASA defined the PDE as the ad hoc connected framework of activities and products that are built upon and support the data collected by planetary space missions and research programs, which primarily are NASA funded.

Types of information in the PDE (not exhaustive):
- Data returned from space missions and ground-based facilities, generated by research and analysis projects, and generated by citizen scientists
  - In addition to observational and research data, this also includes laboratory results, physical samples, Earth analog site field data, and contributions to collaborative citizen-science services
- Standards for planetary science data and metadata
- Software and Tools
  - Data processing pipelines, mission support, analysis tools, search and browse tools, display tools, and simulation tools
- Publications
  - Articles, books, conference abstracts, reports, posters, and presentations
- Education and communication products

Planetary Data Ecosystem – Org Chart

PSD Director – Lori Glaze
PSD Deputy Director – Eric Ianson

Planetary Data Officer – Robin Fergason

PDS Deputy Program Scientist - K.C. Hansen

PDE Program Scientist – Becky McCauley Rench

PDE Chief Scientist – Moses Milazzo

PDE Program Executive – Meagan Thompson

PDS Program Executive - Bo Trieu

We are all available to answer questions - hq-pde@mail.nasa.gov

Acknowledgements

- Planetary Data Ecosystem Team
  - Megan Ansdell
  - Robin Fergason
  - K.C. Hansen
  - Becky McCauley Rench
  - Moses Milazzo
  - Meagan Thompson

- Planetary Data System and Astromat

- SMD Office of the Chief Science Data Officer (OCSDO)
  - Kevin Murphy, Chief Science Data Officer
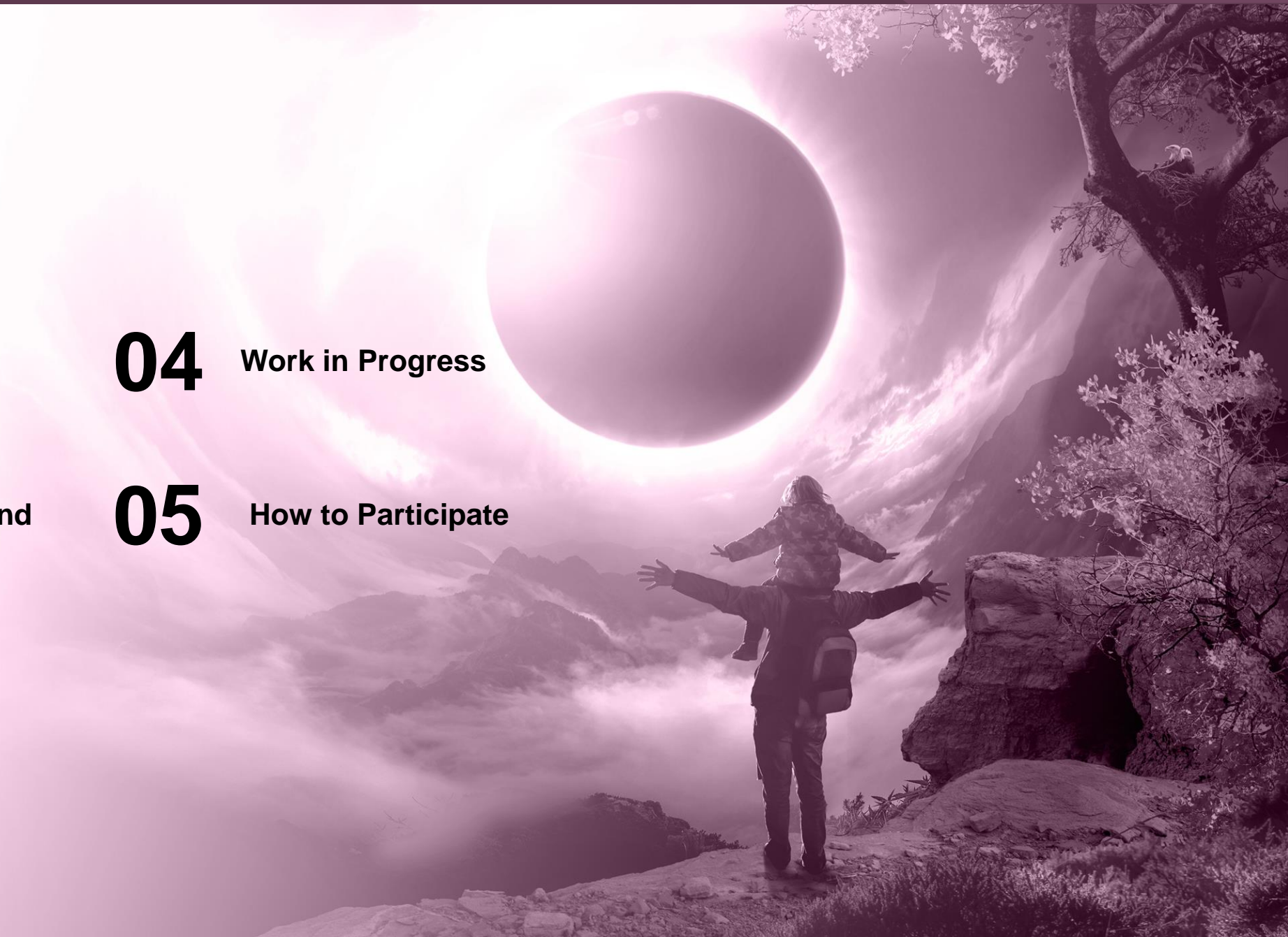 - Steve Crawford
  - Andy Mitchell

Agenda

01 PDE History and Background

02 PDE IRB Background

03 Completed Efforts

04 Work in Progress

05 How to Participate

PDE History and Background

Open Science – Brief History

- Open Data

  - 2013: Executive Order (M-13-13) – Open Data Policy
    - Federal Agencies are required to make information resources accessible, discoverable, and usable by the public.
    - Additional Executive Orders and Memorandums followed, providing more specific requirements for implementation.
  - 2018: Congress passed the OPEN Government Data Act
  - Commercial and academic communities were adopting open principles, and particularly surrounding software development
    - GitHub became a popular repository for developing and sharing software.

- Government, industry, journals and professional organizations, and data communities have been moving in the direction of providing greater access to science results, more open access to data, and open software development.

- ~2019: Formal identification in PSD of a critical and broader Planetary Data Ecosystem that enabled data analysis, interpretation, and new science results

  - Included data, software, tools, infrastructure, policies, developers and users
  - Recognized the need to intentionally develop and manage this Ecosystem to support innovative science and effectively manage resources.

2013   2019

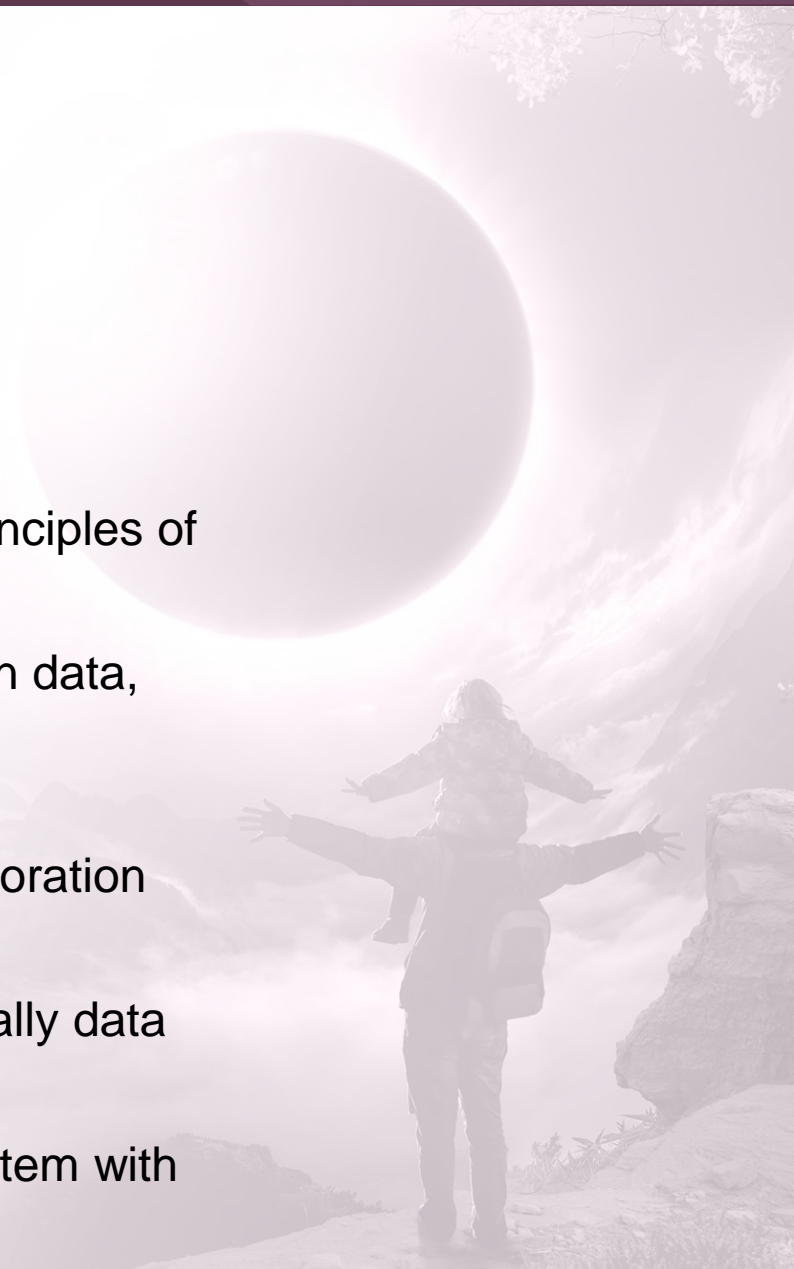# Planetary Data Ecosystem Independent Review Board (PDE IRB)

- PDE IRB Charter
  - Commissioned to conduct a holistic review of the Ecosystem.
  - Goals:
    - Define the full environment.
    - Identify missing or overly redundant elements.
    - Provide findings and prioritized, actionable recommendations from which an optimized strategy can be developed for the PDE.

- Study convened from November 2020 to March 2021

- Outcome
  - Final report of prioritized findings and recommendations
  - ***The guiding document for the development of the Planetary Data Ecosystem.***

2013   2019   2020   2021

PDE IRB Core Values

- The PDE IRB team developed a set of core values that encapsulate their overarching vision:

    - 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.

PDE IRB Final Report

Planetary Data Ecosystem Independent Review Board (PDE IRB)

PDE IRB final report included 67 Findings and 65 Recommendations:
- Continued Strategic Development of the Overall Ecosystem
  - Sections 2 and 3
- Barriers to Data Preservation
  - Section 4
- Barriers to Access, Usability, and Development
  - Sections 5 and 6

PSD status update is presented to the Planetary Advisory Council (PAC) annually for feedback and accountability.
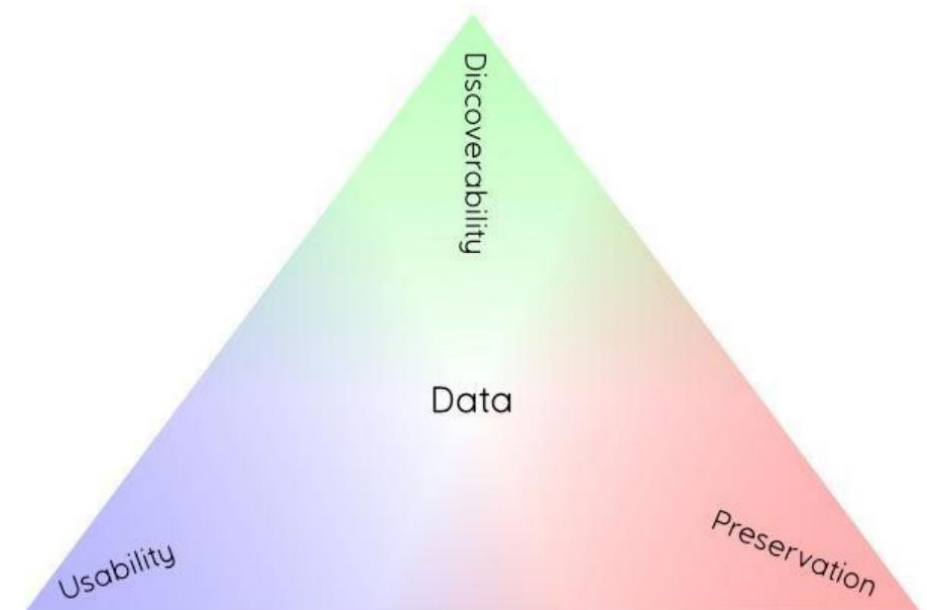
Planetary Science Division PDE IRB Response

Final Report
of the
Planetary Data Ecosystem Independent Review Board
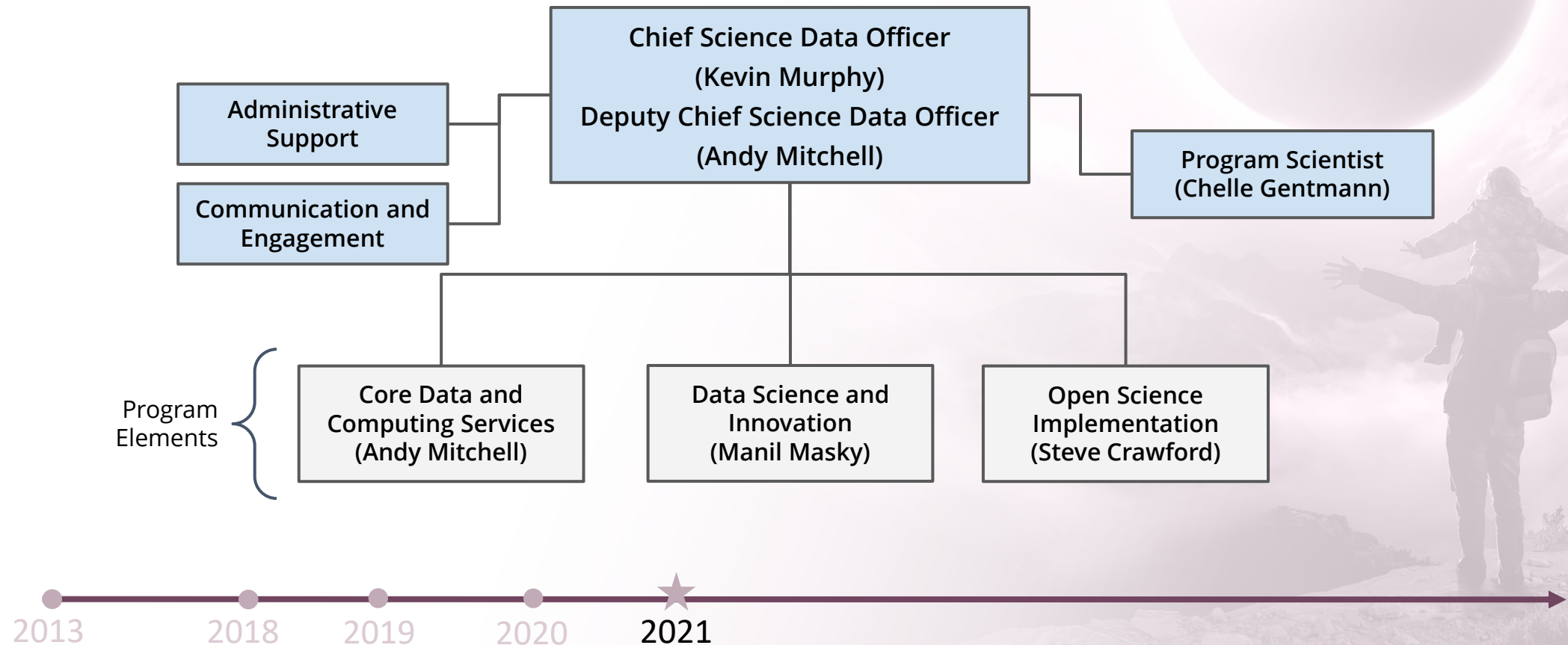**April 2021**

Image Credit: PDE IRB Final Report, 2021

SMD Office of the Chief Science Data Officer (OCSDO)

2021: SMD hired a Chief Science Data Officer and began the development of an SMD data office

Chief Science Data Officer
(Kevin Murphy)
Deputy Chief Science Data Officer
(Andy Mitchell)

Administrative Support

Communication and Engagement

Program Scientist
(Chelle Gentmann)

Program Elements

Core Data and Computing Services
(Andy Mitchell)

Data Science and Innovation
(Manil Masky)

Open Science Implementation
(Steve Crawford)

2013   2018   2019   2020   2021

Initial Planetary Science Division Actions

- Feb 2022: PSD contracted the PDE Chief Scientist – Dr. Moses Milazzo
  - Contractor and external to NASA, providing independence.
  - Gather information, concerns, and feedback from the community and communicate those needs to NASA and the PDS.

- May 2023: PSD hired a Planetary Data Officer – Dr. Robin Fergason
  - NASA Civil Servant, operations and implementation
  - Develop, maintain, and evolve the PDE, using the PDE IRB final report as a guiding document
  - Develop programs and identify infrastructural needs
  - Support Open Science policies and compliance with SPD-41a
  - Work closely with the Office of the Chief Science Data Officer (OCSDO)

2013   2018   2019   2020   2021   2022   2023

Key Takeaway Points

- The PDE IRB final report is taken very seriously.
    - The time put into this effort is greatly appreciated.
    - This final report is the guiding document for PDE development and prioritization.

- Completed Efforts
    - Astromaterial sample data is archived in Astromat
    - Arecibo and radar data can be archived in the PDS
    - Developing policies and infrastructural services (via the OCSDO)
        - Open science requirements defined (SPD-41a)
        - Methods for repositories to assign NASA DOIs is established

- Work in Progress
    - Program Office development
    - Community building
    - Website development - communication and streamlined data access
    - Developing and expanding available data and software repositories
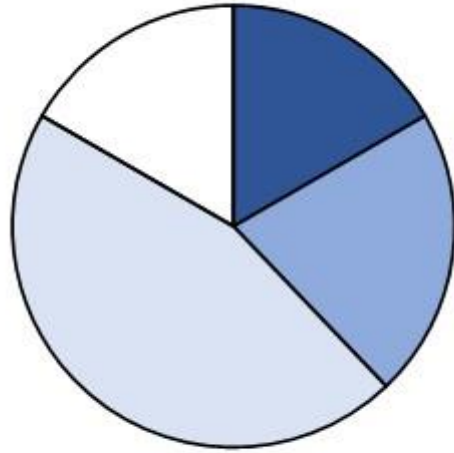
- How to Participate

Status Update:
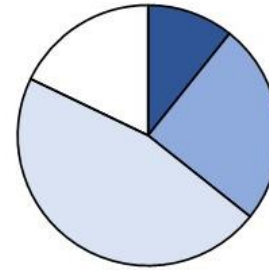Completed Efforts

PDE IRB Response Status - September 2023

Completed (11; 17%) Maintenance (14; 21%)
In Progress (30; 45%) Not Yet Started (11;17%)

Planetary Science Division PDE IRB Response

Continued Strategic Development - September 2023

Completed (3; 11%) Maintenance (7; 25%)
In Progress (13; 46%) Not Yet Started (5; 18%)

Barriers to Data Preservation - September 2023

Completed (3; 17%) Maintenance (2; 11%)
In Progress (8; 44%) Not Yet Started (5; 28%)

Barriers to Access, Usability, and Development - September 2023

Completed (5; 25%) Maintenance (5; 25%)
In Progress (9; 45%) Not Yet Started (1; 5%)

FY23 status addressing recommendations (as of September 2023):
- 11 (17%): Completed
- 14 (21%): Maintenance; completed, but require ongoing efforts to maintain
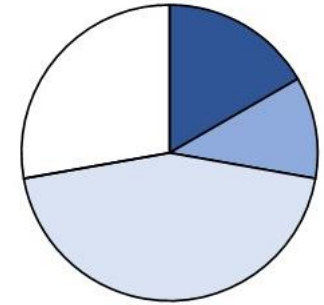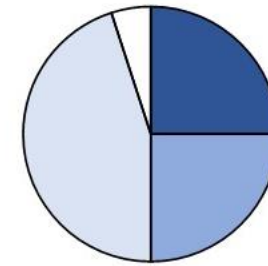- 30 (45%): In progress
- 11 (17%): Not yet started

Astromaterial Sample Data

https://www.astromat.org/

- Astromat was selected in 2023 as the official NASA archive for astromaterial sample data

  - Accepts data collected both from:
    - Mission-supported efforts (e.g., OREx)
    - Research and Analysis (R&A) funded efforts (including laboratory analyses of returned samples and meteorites)
  - Astromat has been selected for funding through 2028. Team is led by Dr. Kerstin Lehnert at Columbia University.

- Addresses R33 and R34

  - R33: NASA should establish a requirement for the preservation of mission-supported laboratory analyses of returned sample material that makes the information accessible to the planetary science community.

  - R34: NASA should 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 (DAP).

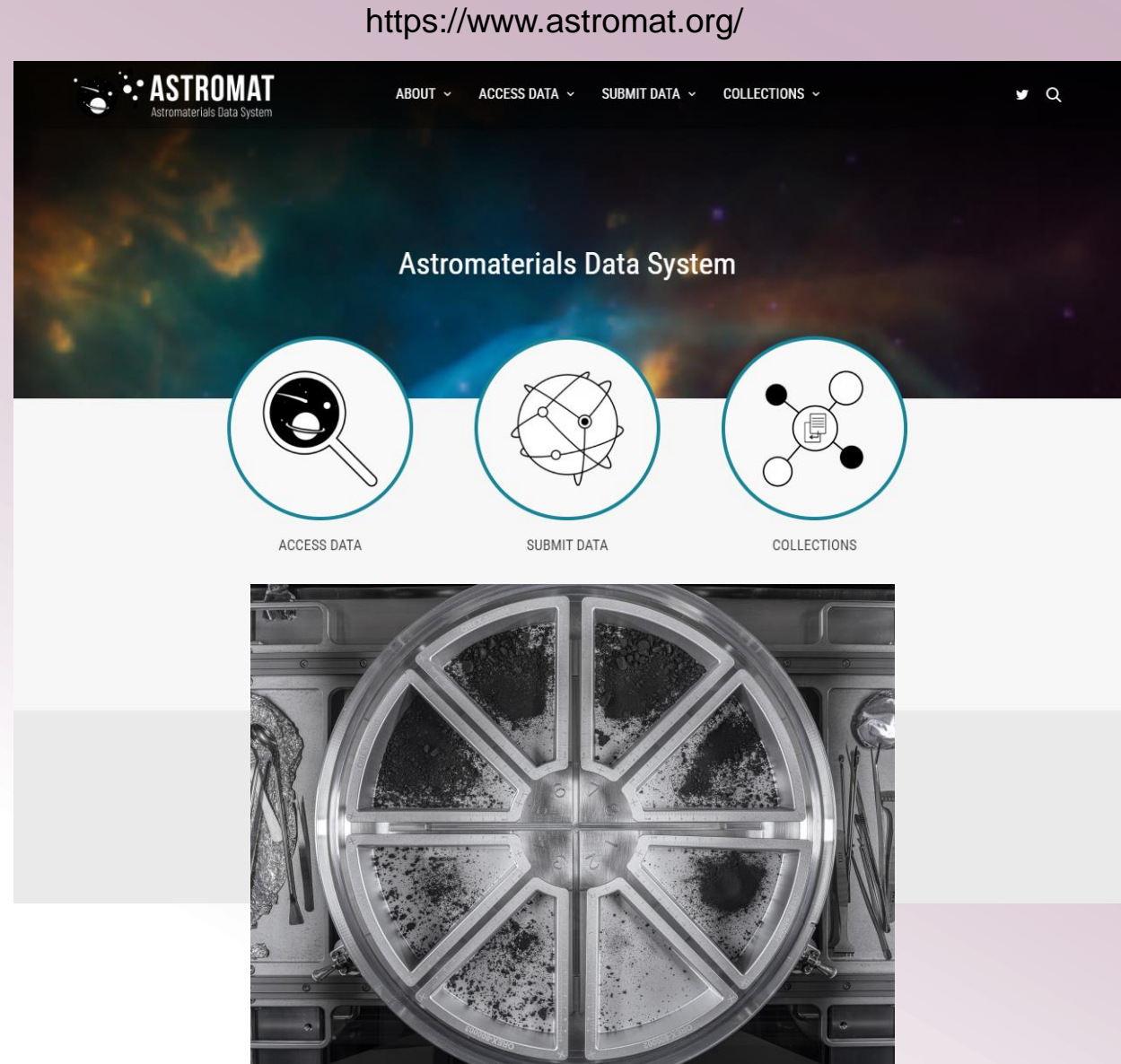 - Supports addressing R10, R16, R35, R47, R53

Image credit and press release: https://blogs.nasa.gov/osiris-rex/2024/02/15/nasa-announces-osiris-rex-bulk-sample-mass/

Archiving Arecibo and Radar Data

Image credit and press release: https://www.nasa.gov/solar-system/nasa-statement-on-nsfs-planned-controlled-decommissioning-of-arecibo-radio-telescope/

- The PDS Small Bodies Node (SBN) accepts planetary radar data
  - Arecibo data is being prepared for archive in the SBD and is currently undergoing review.
  - Asteroid radar data is archived at the SBN
- Radio Science Sub-Node was also established in collaboration with the Planetary Radar and Radio Sciences Group (PRRSG) at JPL
  - Provides a Planetary Radar Advisory Role to the PDS
  - Small Bodies Node Radio Science
    - Deep Impact, Giotto, New Horizons, Rosetta, DART, LICIACube
  - Geosciences Node Radio Science
    - Mercury, Venus, Moon, Mars
- Addresses R31
  - R31: NASA should establish an archive for planetary radar data either within the PDS Small Bodies Node or separately. This archive should facilitate preservation and usability of data at all processing levels by preservation of data processing procedures (or software).

Developing Policies and Infrastructural Services - OCSDO

- Developed policies to define initial requirements to support data and software preservation.

  - [SPD-41a](#) establishes SMD requirements to meet Federal requirements (R03)
    - Supports the realization of Findability, Accessibility, Interoperability, and Reusability (FAIR) data principles (R29, R41)
    - Defines a software preservation and archiving strategy that ensures discoverable, accessible, and usable software tools (R42, R62)
    - Defines desirable criteria for data and software repositories (R40)

- Expanded and enabled the use of persistent identifiers, such as DOIs (R47, R53)

  - DOIs provide a data linkage mechanism and enhance findability and discoverability of science products.

- Developing Infrastructural Services (*Ongoing Effort*)

  - Supports High End Computing and Cloud Computing Services (R27)
    - NASA has an Authority to Operate (ATO) with all major cloud-providers (i.e., Amazon Web Services (AWS), Azure, and Google Cloud).
  - OCSDO recently completed a "*NASA SMD Data and Computing Architecture Study*" and findings from that study will be made publicly available in the coming months.
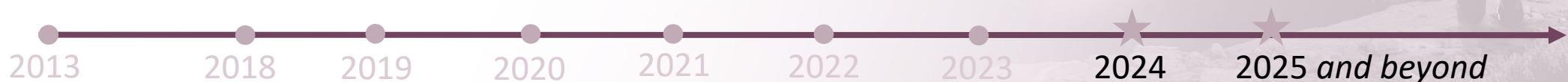
Work In Progress

All future planning - and the execution of such plans - are dependent on available future funding.

Program Office Development

- Developing a (small) Program Office at NASA HQ to manage, develop, evolve the PDE.
  - Currently have a fantastic team!! – Dedicated team is needed to effectively manage the PDE and address the recommendations in the PDE IRB.
  - Defining duties and responsibilities to ensure that we have adequate staffing.
  - Growth of this Program Office would occur in parallel to the development of the overall PDE.
    - The team doesn't need to grow faster (or slower) than the needs of the PDE.

- Enables us to Better Address (*not comprehensive*):
  - R01: NASA should proceed with developing the concept of the Planetary Data Ecosystem.
  - R02: NASA should lead work to refine the full scope of the Planetary Data Ecosystem and build community consensus around the Ecosystem.
  - R03 – NASA should ensure that responsibilities, accountabilities, governance, and service levels for those elements in the Ecosystem by the PSD are clearly defined.
  - E.g. R14, R28, R40, R59

2013   2018   2019   2020   2021   2022   2023   2024   2025 *and beyond*

All future planning - and the execution of such plans - are dependent on available future funding.

Community Building

- With the planetary community (R04; R05; R08; R48)

  - Attend and present at AG meetings
  - Attend and present at community conferences and workshops
  - Organizing future workshops and engagement opportunities

- Within SMD Office of the Chief Science Data Officer (OCSDO) and the Planetary Science Division (PSD) (R26; R48)

- Across Divisions (within SMD), NASA Directorates, National, and International Agencies (R05; R07; R08)

- Facilitating training in areas related to the PDE (R23; R64)

  - Funding training-focused workshops and curriculum development aimed at addressing PDE-related training needs through F2. TWSC and F.14 TOPST.
  - SMD OCSDO developed TOPS Open Science-101 curriculum

2013   2018   2019   2020   2021   2022   2023   2024   2025 *and beyond*
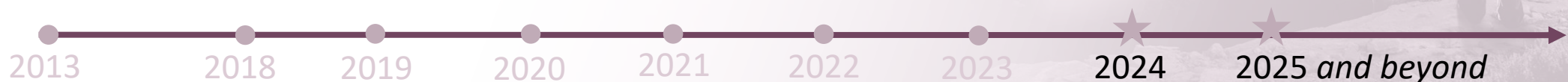
Website Development

- *F08:There is no mechanism for transparent, widespread communication of plans, timelines, or developments of new or changed capabilities, either from NASA to other Planetary Data Ecosystem communities and elements, or among the elements of the Ecosystem.*

- Two website modernization efforts are underway:
  - Planetary Data System Engineering Node is developing a new PDS main page and search (*and it's awesome!*)
    - Significant feedback and user testing; UI/UX expertise (R45; R46)
  - PDE Team is modernizing the Planetary Data Ecosystem website, as part of larger NASA and SMD efforts (F08; R45; R46).
    - Planetary Data website provides a mechanism to communicate plans, timelines, or developments of new or changed capabilities widely.

https://science.nasa.gov/planetary-science/data/

Overview ⌄   Focus Areas ⌄   Data ⌄   Programs ⌄   Resources ⌄

Planetary Data Ecosystem

NASA and CHORUS have signed a participation agreement to expand public access to the results of NASA-funded research. Publications available in CHORUS comply with NASA policy with no further action required by researchers.

Find links to PDE websites ⬤

What is the Planetary Data Ecosystem?

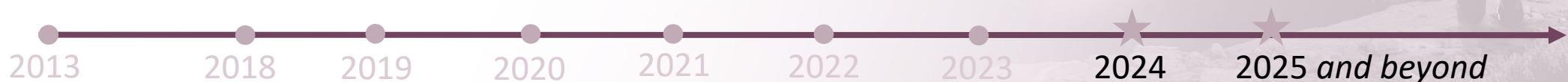Planetary Science Division Information and Data Policy (PDF)

NASA defined the Planetary Data Ecosystem (PDE) as the ad hoc connected framework of activities and products that are built upon and support the data collected by planetary space missions and research programs, which primarily are NASA funded.

2013   2018   2019   2020   2021   2022   2023   2024

All future planning - and the execution of such plans - are dependent on available future funding.

Data/Software Repository Architecture

- Addressing barriers to data/software preservation and access is necessary for a healthy Ecosystem.

  - Data/software preservation and access is foundational to all PSD-funded research and scientific discovery.
  - Data/software use and reuse are built on the prior tasks of archiving and accessibility.
  - Users' needs and expectations have recently evolved to include expectations of higher-level data products and improved access to software and tools.
  - Expanded repository infrastructure is necessary for the planetary community to meet the data and software preservation requirements of SPD-41a.

- Addresses:

  - Metadata standards and reuse: R10; R35; R53
  - FAIR data principles: R11; R29
  - Responsibilities and governance: R03; R14; R18; R48
  - Preservation options for all data/software types: R20; R28; R32; R37; R38; R39; R40; R41; R42; R43; R44; R52; R59; R63

2013   2018   2019   2020   2021   2022   2023   2024   2025 *and beyond*

Activities Over the Next Two Years

- Summer 2024: Develop initial Objectives for a PSD data and software preservation architecture.

  - Apply a NASA Systems Engineering-based approach

  - Build from existing community documents (e.g., 2017 PDS Roadmap; PDE IRB Final Report)

  - Identify objectives, needs, and characteristics for the data and software preservation architecture

- Summer/Fall 2024: Seek community feedback to inform revisions to Objectives.

  - Be on the lookout for a Request for Information!

- Continue to engage the community through RFIs, Workshops, and Special Action Teams to develop Level 1 system requirements and an initial architecture design.

How to Participate

- Status update on PSD actions in response to the PDE IRB findings and recommendations is available.
  - The report documents the results of an internal assessment by the PSD towards developing a more connected and effective PDE.
  - PDF: Planetary Science Division PDE IRB Response
    - Email comments or questions to: hq-pde@mail.nasa.gov
  - GitHub: https://github.com/PDE-CS/Planetary-Data-Ecosystem
    - PDE Chief Scientist (Moses Milazzo) GitHub repo
    - Enables conversations as a community around these recommendations.
    - Post an Issue on GitHub or email hq-pde@mail.nasa.gov

- Email any comments, questions, or ideas to: hq-pde@mail.nasa.gov

Image Credit: https://depositphotos.com/photos/feedback.html

Key Takeaway Points

- ## The PDE IRB final report is taken very seriously.

  - The time put into this effort is greatly appreciated.
  - This final report is the guiding document for PDE development and prioritization.

- ## Completed Efforts

  - Astromaterial sample data is archived in Astromat
  - Arecibo and radar data can be archived in the PDS
  - Developing policies and infrastructural services (via the OCSDO)
    - Open science requirements defined (SPD-41a)
    - Methods for repositories to assign NASA DOIs is established

- ## Work in Progress

  - Program Office development
  - Community building
  - Website development - Communication and streamlined data access
  - Developing and expanding available data and software repositories

- ## How to Participate

THANK YOU!
hq-pde@mail.nasa.gov