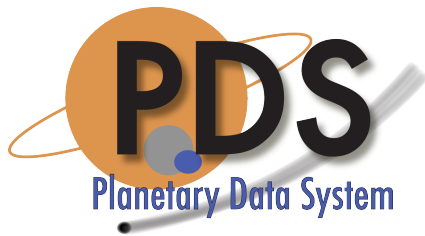




# Experiences with Data Providers from PDART/DAP Program

Nancy Chanover & Lynn Neakrase

PDS MC F2F Meeting  
Pasadena, CA  
19-20 April 2017



# Motivation



- This is the right time to start comparing notes about our interactions with providers of derived data, given that PDART14 selections are in their final year of funding
- We don't want to reinvent the wheel across all nodes – need to discuss what's working and what isn't
- Want to optimize efficiency given that the magnitude of the archiving requirement for derived data was not recognized at the time of the PDS CAN proposals
- Have already done lots of PDS4 “PR” – what information still seems to be lacking among data providers?



# Desired Outcomes



- Feedback to NASA HQ re: what information we need (and perhaps what else the proposers/awardees need)
- Discussion & sharing of strategies for dealing effectively with providers of derived data
- Understanding of most common issues/areas of concern
- Development of plan for future communications to potential data providers



# The Process

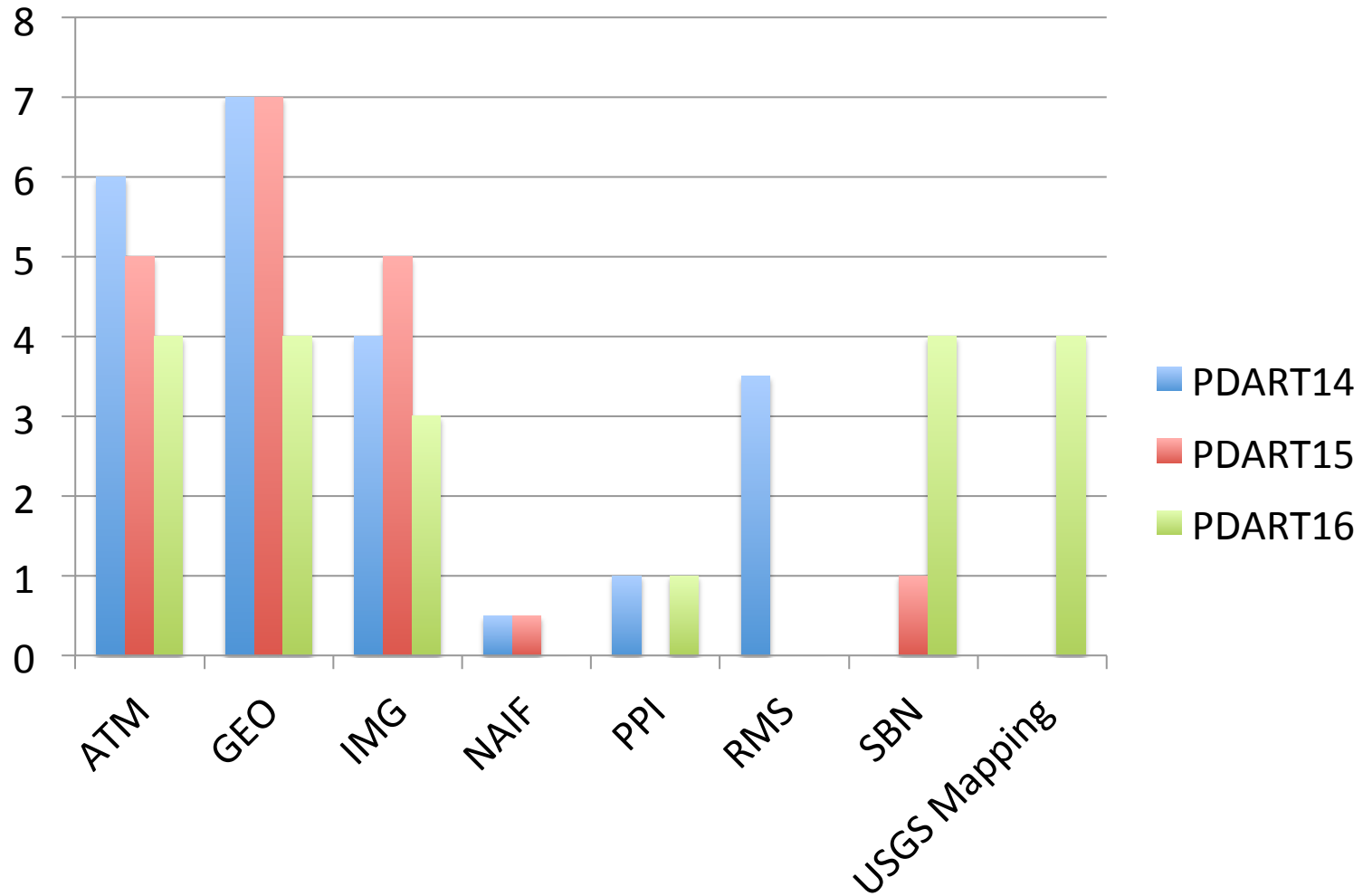
- I sent around some questions to collect some basic information about your interactions with providers of derived data (so far)... received responses from all nodes but one.
- I will present a synthesis of your responses, and we can either discuss them point by point or I can go through the complete set of questions and then we can discuss them as a group.
- First, some context...

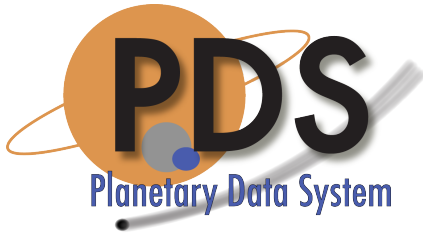


# Relevant Timeline

| When     | What  |
|----------|---|
| Feb 2014 | Release of ROSES-14, containing first PDART solicitation                                |
| Dec 2014 | Release of <i>NASA Plan for Increasing Access to the Results of Scientific Research</i> |
| Feb 2015 | PDART-14 selections announced; 25 proposals selected for funding                        |
| Feb 2015 | Release of ROSES-15, containing first requirement for Data Management Plan              |
| Dec 2015 | PDART-15 selections announced; 24 proposals selected for funding                        |
| Nov 2016 | PDART-16 selections announced; 20 proposals selected for funding                        |

# PDART Selections





# Additional Programs



Providers of derived data are receiving funding from other programs aside from PDART:

|              |
|--------------|
| CDAP         |
| PMDAP → DDAP |
| MDAP         |
| LDAP         |
| SSO          |
| SSW          |
| Others?      |

**Plus unsolicited  
and/or  
“encouraged”  
submissions of  
derived data**



## *Q1a: Describe your level of contact with PDART14 data providers*



- ATM: extensive, monthly-bimonthly emails, met F2F w/3 at LPSC and other conferences in the last year
- NAIF: none
- IMG: initially once/yr, then twice/yr ,then monthly
- GEO: initial contact requesting timeframe for completion of archiving tasks, then ongoing
- RMS: ongoing contact, substantial involvement in 2/3
- SBN: no PDART14 providers, but in contact with 4 others





*Q1b: How far along in the archiving process are your PDART14 data providers?*



- ATM: 4 likely to submit by end of summer, 1 working with RMS, 1 may request NCE
- IMG: 2 on schedule to deliver this year, 3 in NCE w/archiving still pending
- GEO: 1 completed peer review of first of two deliveries, 2 in label design process, 3 not ready to start archiving until FY18
- RMS: 2 making satisfactory progress, 1 hasn't sent detailed delivery plans or schedules
- SBN: one peer reviewed and in lien resolution, 1 regularly archiving already, others in development



## Q1 Take-Away Points



- The data providers don't always know which node they should be talking to!
- Should we consider standardizing our initial contact letter once they have been notified of selection?
- Frequent contact/reminders tend to work well for keeping the archiving task on their radar.



*Q2: Have you provided label templates and/or PDS4 training for Your data providers?*



- ATM: label templates, partially filled in references, 3 actively producing labels on schedule
- NAIF: No, plan to work one-on-one w/provider (@ JPL)
- IMG: label templates for the 2 on schedule
- GEO: label templates (asked for samples of data and designed labels for them), also offer in-house label generation software
- RMS: No, but effectively yes for 2 (as members of proposal team, not as PDS)
- SBN: have training and sample data for OLAF so don't need special PDS4 training



## Q2 Take-Away Points



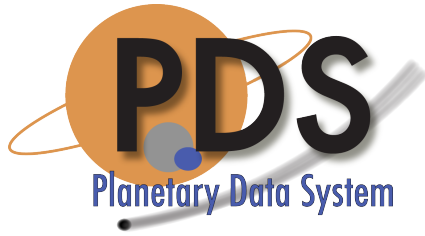
- Hands-on PDS4 training to support PDART/DAP archiving efforts is needed! This will be offered at the upcoming Planetary Data Workshop. **Should the PDS consider offering these trainings every ~ 6 months?**
- The PDS4 documentation is not mature enough yet for data providers to be expected to do this on their own. [Susie has been revising the DPH to help address this!]



*Q3: Have your interactions w/derived data providers necessitated an adjustment in FTEs dedicated to this task?*



- ATM: Yes, 2 students for ELSA development
- NAIF: N/A
- IMG: Yes, 1 new FTE @ USGS, 3 new hires at JPL
- GEO: Yes, hired a student
- RMS: Yes, estimate 0.2 FTE/DAP recipient
- SBN: dedicating 2.5 FTE over next 2 yrs for OLAF overhaul to support R&A archiving



## Q3 Take-Away Points



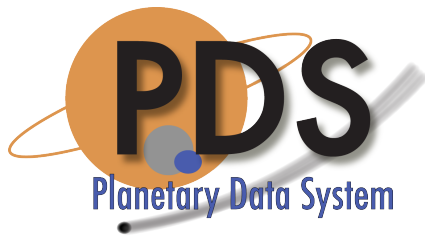
- Mission requirements have not gone away so the work associated with supporting PDART/other DAP archiving is in addition to (not replacing) other node tasks.
- Providers of derived data generally do not have the instrument team infrastructure that mission data providers would have; they are frequently working alone or with one other person (potentially a student) on the archiving tasks.
- There is a learning curve associated with PDS4 so the nodes naturally need to provide a substantial amount of support at this stage.



*Q4: How helpful was it to receive the selection list from NASA HQ, and did it result in earlier communication with the data providers?*



- ATM: Very helpful; node initiated contact in all cases
- NAIF: Useful
- IMG: Very helpful
- GEO: Very helpful; node initiated contact in all cases
- RMS: Very helpful; node initiated contact in all cases
- SBN: list not communicated to PSI



## Q4 Take-Away Points



- Everyone agrees that it is incredibly useful for the nodes to receive the selection lists ASAP.
- The nodes would like to receive these lists for other R&A programs besides just PDART!
- Who decides what node is assigned as the contact node? Presumably this is done at the pre-proposal stage when proposers request letters of support, but there have been a few mix-ups. We should all agree on the final list of assignments shortly after the selections are made!
- Can we add a question in the Program Specific Information section of the cover sheets for non-PDART programs that would flag a proposer's plans to submit derived data to the PDS? (this would then require program officers to communicate this to the PDS or individual nodes)





## Q5: *Describe interactions with other (non-PDART14) providers of derived data*



- ATM: several CDAPs (15,16) and a few other submissions with emails concentrated at beginning and end of work; provided most of the PDS4 expertise for generating templates with providers for 2 completed last year; several other small non-funded/non-R&A projects waiting to be completed
- NAIF: N/A
- IMG: contacted all of them; responses vary
- GEO: of 7 PDART15, 4 PDART16, 9 misc.: 2 completed archives, others range from starting to lien resolution, a few others turned up after receiving funding but without having made initial contact
- RMS: several CDAPS, contacts range from introductory to extensive
- SBN: no contact from PDART15 or PDART16 awardees



## Q5 Take-Away Points



- There is a strong desire for the nodes to be made aware of all selections from all R&A programs that involve PDS archiving shortly after selections are made.



## *Q6: Other comments or lessons learned from working with providers of derived data?*



- ATM: working top-down for bundle development provides most of the needed references early on for filling out templates; our interactions have been showing small providers understand the bulk PDS4 structure quicker
- NAIF: N/A
- IMG: nodes need to coordinate training tool development better and need to schedule regular training opportunities for user community
- GEO: we need better documentation, examples and tools (esp. to generate labels from a template and validate a complete archive bundle), designed for non-programmers
- RMS: need more integrated approach for PDS4 training for small data providers
- SBN: need more PDS4 training for providers



## Q6 Take-Away Points



- The need for more/better tools and training seems to be a common theme that emerged from the comments received so far.
- Frequent communication (in both directions) is the most effective way to ensure that the archiving gets done correctly. This is hard because EVERYONE IS BUSY.



# Where do we go from here?



- Training:
  - Planetary Data Workshop
  - Regular offerings at DPS, LPSC, AGU?
  - Longer term: online training pages like NAIF has for SPICE?
- Tools:
  - PLAID, OLAF, ELSA, ... ??
- Requests for more info from NASA HQ program managers
  - Who (among selected proposals) is promising to archive in PDS?