# Deep Archive (NSSDC)

## PDS Management Council Meeting

## Washington, D.C.

April 2008

http://pds.nasa.gov
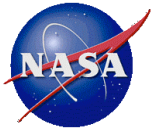
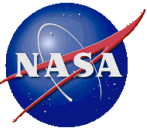# Topics
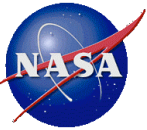
- Requirement

- Background

- Delivery Inventory Status

- Electronic Delivery Status

- Large-scale Data Delivery Status

- Next Steps

- PDS Level 3 requirement 4.1.5:

  "PDS will meet U.S. federal regulations for the preservation and management of the data through its Memorandum of Understanding (MOU) with the National Space Science Data Center (NSSDC)."

# Background

- MOU established with NSSDC for long-term preservation

- Data set volumes were mailed to NSSDC in the past using hard media (including tapes, CDs, DVDs), old NSSDC IDs were assigned to them

- New NSSDC tape archive system, was implemented, which enables electronic delivery

- Software was provided by NSSDC to facilitate electronic delivery

- With the new NSSDC system, new NSSDC IDs are created

- Effort initiated in gathering delivery inventory per MC direction

- NSSDC Delivery Working Group (NDWG) was formed to address large scale data delivery

- NSSDC Interface project page can be found at: http://pds-engineering.jpl.nasa.gov/index.cfm?pid=100&cid=105
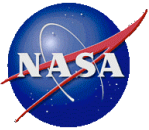
# Delivery Inventory Status

- NSSDC provided a list of PDS data sets in their holding with corresponding new and old NSSDC IDs

- EN generated a list of PDS data sets currently in PDS Catalog with corresponding new and old NSSDC IDs as well as Curating Node

- Two lists were merged and the results were sent to the Nodes for review

- In progress of gathering comments and performing analysis

- Disparities found

  - Missing data sets in NSSDC
    - Data sets that have yet to be delivered to NSSDC
    - Data sets that have been delivered to NSSDC but were not located at NSSDC, require research if they need to be re-delivered
  - Incorrect data set names and IDs in NSSDC
  - Missing data sets in PDS Catalog
    - Data sets found in NSSDC only, require research if they should be ingested into PDS Catalog
  - Incorrect or missing NSSDC IDs in PDS Catalog
    - New NSSDC IDs were not ingested into the PDS Catalog
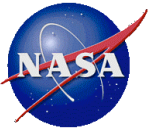
# Electronic Delivery Status

- Linux version of the MPGA (Multi-file Package Generator/Analyzer) software was provided to PDS Nodes to package volumes into AIPs (Archive Info Package) for delivery via FTP

- Upon approval by the MC at the December 07 F2F meeting, the software was made available to ATMOS, IMG, NAIF, PPI, SBN and RINGS

- AIPs are restricted to 300 GBs in size due to NSSDC archival media constraints

- ATMOS and SBN have been making deliveries of data sets < 300 GBs electronically using this software
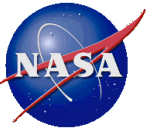
# Large-scale Data Delivery Status

- NSSDC Delivery Working Group was formed to address large data sets (> 300 GB) delivery to the NSSDC

- Working Group members include D.Crichton, E.Grayzeck, S.Hughes, C.Isbell, E.Law, P.McCaslin, and T.Stein

- Produced a white paper which describes proposed processes to submit, retrieve and replace archive volumes
  - Data will be sent via "data brick" to the NSSDC
  - NSSDC will be responsible for transferring the data from the data bricks to their archive system
  - Minimal impact to the nodes

- NSSDC provided software to generate submission manifests that describe what's included in the data brick

- In progress of alpha testing the proposed processes using data sets and volumes from Geosciences and Imaging Nodes
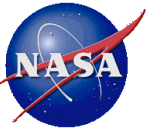
# Next Steps

- Work with Nodes and NSSDC to complete delivery inventory (May 2008)

- Complete alpha testing (April 2008)

- Fine tune large data set delivery processes based on alpha test results and finalize the white paper (April 2008)

- Perform beta tests by Nodes (May 2008)

- Present final processes and beta test results to the MC and get approval (June 2008)

- Work with the MC to develop delivery and tracking policy for both small and large data sets delivery (June 2008)

- Work with the Nodes to develop delivery and tracking plans (including PDS Catalog update) (July 2008)