Catalog Cleanup

PDS Management Council Meeting

Mountain View, California

March 2006

http://pds.nasa.gov
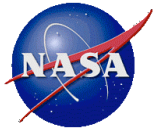
Background

- Major PDS catalog and DN data holding synchronization effort performed in March – May 2004 for D02
  - DN and CN synchronized PDS catalog resource links to Nodes holding

- Noticeable discrepancies
  - Set #1: PDS Catalog values missing or mismatch
    - Data set name
    - Data set terse description
    - Citation and Abstract
    - Label revision note
    - Archive status
    - Resource links
    - SBN Target Locator
  - Set #2: Missing data sets in Catalog
    - Data sets in EN queue, yet to be ingested
    - Data sets did not reach EN
    - Archive status in lien, therefore not delivered to EN

- Investigation needed for other discrepancies not apparent
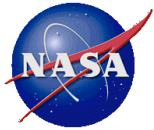
- Catalog/DN data holding reconciliation report
  - Wrote program to extract data holding info from GEO, PPI and SBN, compared with Catalog, generated and reported discrepancies to GEO, PPI and SBN (Jan-Feb 2006)
  - Generate reports for other Nodes (Mar 2006)

- Reconciliation
  - Work with Nodes to investigate unobvious discrepancies (Mar-Apr 2006)
  - Work with Nodes to identify reconciliation plan (Mar-Apr 2006)
  - Present to MC for approval to proceed (May 2006)
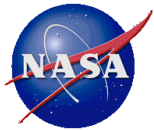
- **Primary / Backup online repository**
  - Corrected Catalog resource links: (Jan 2006)
    - Replaced CN Online Archive by EN Backup Volumes Online
    - Replaced CN Volume Offline by EN Backup Volumes Office
    - Removed resource links pointing to CN Online Archive for PPI data sets as requested
  - Wrote program to query Node product servers, compared with Catalog, generated and reported "PDS Primary and Backup Repository Report" to ATMOS, GEO and IMG (Feb 2006)
  - Generate reports for other Nodes (Mar 2006)
  - Work with Nodes to implement PDS secondary online repository requirement (TBD by MC)
    - Only ATMOS identified EN as its backup online repository
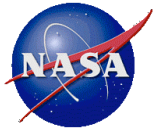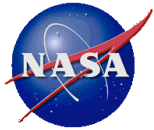
Tools

- Tools that can help resolve discrepancies

- Citation update tool
  - Created offline tool to update citations (Jan 2006)
  - Currently working with SBN

- Resource Manager
  - Released online tool to allow Nodes to update data set resources (Jan 2006)
 - http://pds-engineering.jpl.nasa.gov/index.cfm?pid=48

- Ingest and release process
  - Review and document current process (Feb-Mar 2006)
  - Identify problems and difficulties (Mar-Apr 2006)
  - Present proposal to MC (May 2006)

- Node/EN delivery tracking and notification tool
  - Define requirements (Feb-Mar 2006)
  - Evaluate if feasible to be included in CATS (Apr 2006)
  - Present proposal to MC (May 2006)
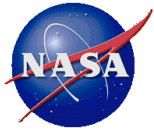
- NSSDC Identifiers
  - Currently, PDS Catalog contains NSSDC IDs
  - Question: Should they be maintained in the Catalog
  - If yes, need to identify the process to update and maintain them