It seems to me that we're preparing a 'pirates code' for a mission lead node – not really rules, more like guidelines. The idea is that once a node is identified as a 'lead node', that node identifies their primary mission rep who then uses the guidelines to help ensure that critical items happen in a timely fashion. All of us work to update the guidelines and their associated appendixes. I've inserted placeholders for some of the other entries in the guidelines. – Mitch

New Mission Lead Node Guidelines

I. Establish communications channels

1. Immediately after a node is identified as a lead node, they decide on which communication options they want to use and then set them up. Options include sending email, establishing email distribution lists, dedicated web pages – either as places for the lead node to post information and schedules, or as a wiki for multiple member posting. The lead node may want to consider setting up separate channels with some restricted to PDS only (e.g., one email distribution list for which all of the members are PDS nodes involved with the mission, and another for all participants of the AWG). In addition to other options, a mission specific web page should be established on the MIWG wiki.

II. Data Dictionary Additions (DDA)

1. Early, the lead node should produce a set of mission specific keyword values; see the DDA Appendix for a list of potential keywords.

2. Email the list of proposed keyword values to all PDS nodes for review, post on the mission page on the MIWG web site, and make discussion of the mission keywords an agenda item for the next MIWG telecon.

3. The MIWG telecon should establish consensus on the values. Once this is reached, supply the list of values to the mission – direct email to all AWG members, mission and instrument team leads, update the MIWG wiki entry, and post on lead node's mission web site if you made one.

4. Go over the keyword values and use requirements during a subsequent AWG and solicit input from the AWG members (especially mission and instrument reps) to insure that the list is complete.

5. Use these initially in a local data dictionary.

6. Submit to EN for inclusion in the PDS data dictionary once the mission SIS or equivalent document has been approved (*not sure this is the correct milestone for this step – MIWG discussion point*).

III. Catalog Files

Last updated 7-29-08 by Betty Sword

1. The lead node should provide mission representatives with templates and examples for MISSION and INSTHOST files and provide instrument representatives with templates and examples for DATASET, INSTRUMENT, PERSON, VOLDESC and REF files. If the lead node is maintaining a web site for the mission archive task, the examples should be posted there.

2. Catalog file due dates should be included in the archive schedule. Separate dates may be included for skeleton and/or draft and final versions. Skeleton or draft versions are adequate for End-to-End Tests.

3. Catalog due dates of the Engineering Node (EN) depend upon the circumstances.

Catalog Files	Version	Need Date
Required files		
MISSION.CAT INSTHOST.CAT INST.CAT - per instrument DATASET.CAT – per data set REF.CAT VOLDESC.CAT PERSON.CAT	Skeleton or Draft	In case of E2E testing, End of E2E test less 30 days or per the test schedule
(Same as above)	Final	First release less 8 weeks
Optional files		
SOFTWARE.CAT		Not ingested
DS_MAP_PROJECTION		Not ingested
TARGET.CAT		Only for new targets
*RELEASE.CAT		Two weeks before release
*HOUSEKEEPING.CAT		Two weeks before release
**Revised files		
Any of above that need to be ingested	Revised final	Release date minus 2-3 weeks

*Engineering Node would prefer that the information in the RELEASE.CAT and HOUSEKEEPING.CAT files be provided via GUI interfaces on the EN Catalog Tools site. The information is required; use of RELEASE and HOUSEKEEPING catalog files is optional. In cases such as MER, where there are over 100 release objects for each release, the EN DE can create a master file containing all release objects in the form of SQL statements.

**If previously ingested files are being revised, the provider should use the most recently ingested file as the starting point, since EN data engineers frequently have done some reformatting and/or corrections prior to ingestion. Ask your EN DE to provide the most recent version, or extract it from the database by using the Catalog Template Search on the EN web page. Allow several weeks lead time for updated files to be ingested.

4. Catalog files should be validated with PDS validation software prior to delivery to assure they are correct; they should also be correctly formatted.

IV. Peer Reviews

Last updated 7/24/08 by Susie Slavney

The approach to conducting a peer review varies somewhat from node to node. The main difference in approach is determined by whether the data set accumulates over a long period of time, as for a planetary mission, or is completed in a single delivery, as for an observatory or laboratory data set. With a single delivery, the products that are reviewed are the products that will be archived. In the case of a mission that will deliver over several years, it is not reasonable to conduct a review for every delivery, nor to wait until the end of the mission to review the complete archive. The review must take place in time for any recommended changes to be made to the processing pipeline software, well before data acquisition begins. So the review covers a representative sample of pipeline data products and a Data Product Software Interface Specification (SIS) document. When actual data are delivered, they are validated by the receiving node for compliance with PDS standards and for compliance with the peer-reviewed SIS.

1. What should be included in a peer review?

Data set delivered over time:	**Single-delivery data set:**
Required	
Sample data products with PDS labels	Actual data products with PDS labels
Data Product SIS, including: - examples of labels - table of keyword definitions	Complete archive volume, including: - catalog files - index table - documentation (not necessarily a SIS)
Archive Volume SIS, including: - columns in index table	
Desirable	
Catalog files	
Complete archive volume (that is, all directories populated as for a delivery, but with sample data in the data directory)	
Local or "working" data dictionary to use with validation tools	

The exception to the above is SPICE data. Sample SPICE data is not sufficient for a review.

2. When should the review take place?

Data set delivered over time:	**Single-delivery data set:**
Before final data processing pipeline is frozen, but not too far ahead; requires PDS to work closely with Project on scheduling	When archive is completely assembled
Before any end-to-end delivery tests	

3. How should the review take place?

Most reviews take place via email and teleconference over a period of several weeks. Only the Small Bodies Node conducts some reviews on site on a regular basis. The following guidelines apply to reviews conducted electronically.

The PDS review coordinator should maintain a password-protected web site accessible to reviewers, instrument team members, and PDS, on which the schedule, review materials, comments and responses are posted.

In general, the procedure is as follows.

a. The review coordinator assembles a list of potential reviewers and invites them by email.
b. The review coordinator announces the opening of the review and provides access to the web site.
c. Reviewers have about two weeks to examine the material and email their comments to the review coordinator.
d. The review coordinator posts comments (either separately, or collated in one list, or both) on the web site.
e. The data providers have about two weeks to email their response to the comments, which the review coordinator posts on the web site.
f. The review coordinator decides whether there is reason to hold a review teleconference, based on the data providers' response. If so, a teleconference is arranged at a time convenient for reviewers, data providers, and PDS reps.
g. At the teleconference the unresolved issues are discussed one by one. The result should be a list of actions, or liens, that the data providers must perform to enable the data set to pass the review. The data providers may decline to follow a reviewer's recommendation if they can justify doing so (for example, because of insufficient resources), as long as the resulting data set would still be PDS-compliant. A record of the resolution of each issue is posted on the review web site.
h. The data provider resolves the liens within a given period of time, say from two weeks to a month. (The schedule should take into consideration other mission activities going on; in particular the liens should be resolved before delivery testing begins.)
i. The data provider delivers revised material to the review coordinator, who posts it on the review web site.
j. The review coordinator invites the reviewers to examine the revised material to be sure their comments have been addressed. If the reviewers are satisfied, the review is complete.

4. Who should be on the review panel?

- At least three scientists who have experience with the type of data
- At least one scientist not from the instrument team; more are better
- Scientists who are on the team but not involved in the data set design; since they have a real stake in the outcome, they may be the most critical reviewers

- PDS Data Engineer for the mission
- PDS representative from a node involved in another instrument team on the mission
- PDS representative from a node not involved in the mission

V. End-to-End Testing
Last updated 8/26/09 by Susie Slavney

Purpose. End-to-end testing allows for a rehearsal of a PDS delivery. Data is created just as it will be in a real delivery to PDS, and all of the interfaces between instrument teams, mission data handlers, PDS Nodes, Data Nodes, and the Engineering Node are exercised. These tests help to bring to the surface any problems before the actual first delivery. To date end-to-end tests have been conducted for the MRO, Phoenix, LRO, and LCROSS missions.

Role and responsibilities. The Lead Node works with the mission via the mission archive working group to (1) determine whether or not to conduct end-to-end tests, (2) to establish responsibilities and schedule, and (3) to conduct the tests. At least the following positions will be assigned: Test Manager (most likely a mission person), PDS Test Lead (from the Lead Node), Test Leads from each instrument team and participating nodes.

Documentation. The following documents aid in guiding and recording the end-to-end activities:

- A *Test Plan* defining tests, responsibilities, schedules. Authored by mission with the aid of the Lead Node.

- *Test Procedures* – one set per test for each instrument team / PDS node, includes very detailed steps.

- *Test Reports* – summary of results per test, compiled by mission lead with input from all.

Typical Tests. Generally, end-to-end tests are broken up in such a way as to build upon each other and culminate in a "final dress rehearsal". An example 3-test scenario might be:
Test 1: Delivery Path / Handshaking Test – tests paths between instrument teams and PDS by passing at least one file that does not need to be an actual data product file. Validation is by checksum.
Test 2: Archive Volume Skeleton Test – data flow includes a properly formed archive volume for each instrument with at least one valid data product, though the test may be limited to EDRs. The volumes should include all documentation, catalog files, index tables, and other files that would be part of a normal delivery. Some of these files may be dummy files. Optimally, final Data Product SISs and Volume SISs would be available. Manifests and checksums are included.

Validation by both the mission teams and PDS nodes includes validation for PDS standards and verification that data products follow the design in the Data Product SIS.
Test 3: Full Archive Volume Test - Full archive volumes for all anticipated data sets are created, catalog files are ingested, PDS release notice generated. Though some files may still need to be "dummied up," this test should resemble a real delivery as nearly as possible. This test uses the PDS test bed at Engineering Node to exercise EN functions.

VI. Release Announcements
Last updated 7/29/09 by Susie Slavney

The PDS Engineering Node maintains a list of users who subscribe to receive email announcements of selected data releases, documentation, and software. Any user may use the link on the PDS home page to be added to the list, removed from the list, or to change the selection criteria for announcements he or she wishes to receive.

Announcements are sent to subscribers for the following events:

1. A scheduled release of data from a mission.
2. A release of data from an individual data provider.
3. A substantial revision, correction, or addition to a data set, including new or revised calibration data. The data provider and the PDS representative determine whether the event is substantial enough for an announcement.
4. A release of PDS software.
5. A release of PDS documentation.
6. Other events of importance to PDS users.

This policy addresses announcements for data releases, not for PDS documentation or software.

All data set releases (not just those from missions) should be announced.

The announcement for a mission's first release is sent to all subscribers on the list, whether or not they have signed up to receive announcements for that mission. The announcement includes instructions for signing up to receive future emails about that mission, and a warning that users who do not sign up will not receive further announcements. This policy may be extended to subsequent releases if PDS reps think it necessary (for example, if there are only a few subscribers after the first announcement). A pre-release announcement may be sent a few weeks before the first release to allow subscribers to sign up early.

All announcements are sent to the editor of the Planetary Exploration Newsletter (PEN), pen_editor@psi.edu. If the announcement is too long or otherwise does not meet PEN guidelines, an edited version should be sent.

Announcements of a scheduled data release are sent on the day of the release, if possible, but no more than two working days later.

The contents of a data release announcement follow a standard format agreed upon by the data provider(s) and the PDS rep(s). Any amendments to the standard announcement must be

submitted by the PDS rep to the release announcement coordinator at EN at least 1 week before the announcement is to be sent.

The standard data release announcement includes, at minimum, the following information:

- Mission and instrument names, if this is a mission-related release
- Data set names and/or descriptions in general terms where practical, understandable by users not already familiar with the data
- URL for access to data set(s)
- Instructions for subscribing and unsubscribing to the email list.

Release announcements should also include information about reprocessing or calibration updates to a data set, if there are any. The lead node should poll the other nodes for this information for each release, and coordinate the text for the release announcement.

Recommendations for release announcements from Mark Sykes are given in Appendix 2.

VII. Versioning of data sets and products

Last updated 2/5/10 by Susie Slavney

When a data set is first delivered to PDS the Data Set ID should have the version number 1.0.

The Data Set ID version is incremented rarely, and only when the entire data set has been revised. Individual products may be revised without requiring a change of the Data Set ID version. It is to be expected that during the course of an ongoing mission some data products may be revised one or more times as knowledge of them improves. Product revisions should be tracked using the label keywords PRODUCT_CREATION_TIME, PRODUCT_VERSION_ID, and other versioning keywords if present such as SOFTWARE_VERSION_ID. Product revisions should also be recorded in the ERRATA.TXT file for the archive.

When a Data Set ID version is incremented, a new data set comes into existence. The previous version still exists as a PDS data set, although its archive status is set to SUPERSEDED. As of this writing, discussion is ongoing within PDS about the proper disposition of SUPERSEDED data sets. When a conclusion is reached this text will be updated.

When incrementing a Data Set ID version number, the instructions in section 6.3.2, item 8, of the PDS Standards Reference should be followed.

VIII. Data Transfer Protocol Recommendations

Last updated 7/28/10 by Emily Law

Recommendations for data transfer protocol to use between data providers and the nodes can be found on the MIWG wiki home page http://oodt.jpl.nasa.gov/wiki/display/MIWG/Home under Documents section.

Appendix 1 (DDA) Mission Specific Keywords

Suggested keywords. Lead node decides which are relevant for the mission and proposes mission specific values.

- DATA_SET_NAME
- DATA_SET_ID
- INSTRUMENT_HOST_NAME
- INSTRUMENT_HOST_ ID
- INSTRUMENT_NAME
- INSTRUMENT_ ID
- MISSION_NAME
- MISSION_ALIAS_NAME
- VOLUME_SERIES_NAME
- VOLUME_SET_NAME
- VOLUME_SET_ID

Appendix 2. Recommendations for Release Announcements

[Copied in its entirety from *Guidelines for PDS Data Release Announcements* by Mark Sykes, http://www.psi.edu/~sykes/PDS/datarelease/.]

GUIDELINES FOR PDS DATA RELEASE ANNOUNCEMENTS

General Principles:

- Releases should always be by the "PDS" and not the discipline nodes. There is only one PDS. Links at the DN dataset locations should guide people to other resources provided by the DNs for the community they serve.
- An announcement should describe what has been released and provide a link to a page through which a user can access that data. Data Set IDs do not convey useful information. Data Set Names do, as does Data Set Terse Description (or it should...).
- The URL for the dataset should be to a page where one can directly access that particular dataset and the information necessary for its use. The number of clicks from announcement to data download should be minimum.
- Announcements should be pithy.

Sample Template 1

```
PDS RELEASES [DATA SET IDENTIFICATIONS]

The NASA Planetary Data System announces the availability of
the following data set[s]: [or other optional introductory message]

[Data Set Name 1]
[URL 1]

[Data Set Terse Description 1 - optional]

[Data Set Name 2]
[URL 2]

[Data Set Terse Description 2 - optional]

For access to all data archived in the PDS, go to:

http://pds.nasa.gov
```

Sample Template 2

```
PDS RELEASES [DATA SET IDENTIFICATIONS]

The NASA Planetary Data System is pleased to announce the availability
of data from the [MISSION] for the following instruments [or other
optional introductory message]:

[Instrument 1]
[Instrument 2]
```

[Instrument 3]

These data, as well as mission and instrument information, may be accessed at:

[URL]

For access to all data archived in the PDS, go to:

http://pds.nasa.gov

EXAMPLES (MODIFIED)

February 3, 2008

[NASA] PDS RELEASES NEW HORIZONS DATA

The NASA Planetary Data System is pleased to announce the first delivery of the data from the NEW HORIZONS mission. The delivery includes POST-LAUNCH CHECKOUT and JUPITER FLYBY row and calibrated data for the following New Horizons instruments:

MVIC (Multispectral Visible Imaging Camera)
LEISA (Linear Etalon Image Spectral Array)
Alice (UV imaging spectrometer)
SWAP (Solar Wind Around Pluto)
PEPSSI (Pluto Energetic Particle Spectrometer Sciences Investigation)
LORRI (Long Range Reconnaissance Imager)
SDC (Student Dust Counter)
SPICE data are also available.

To see and download the data as well as mission and instrument information, go to

http://pdssbn.astro.umd.edu/missions/newhorizons/index.html

For access to all data archived in the PDS, go to:

http://pds.nasa.gov

January 13, 2008

[NASA] PDS ANNOUNCES FIRST MESSENGER DATA RELEASE

The Planetary Data System announces the first release of data from the MErcury Surface, Space ENvironment, GEochemistry, and Ranging (MESSENGER) mission. The release includes EDR level data acquired during cruise, the Earth flyby (Aug 2, 2005), and the 1st (Oct 24, 2006) and 2nd (June 5, 2007) Venus flybys. Not all instruments acquired data during each of these mission phases.

Data sets from the following experiments are now available:

EPPS (Energetic Particle and Plasma Spectrometer)
GRNS (Gamma Ray and Neutron Spectrometer)

MAG (Magnetometer)
MASCS (Mercury Atmosphere and Surface Composition Spectrometer)
MDIS(Mercury Dual Imaging System)
MLA (Mercury Laser Altimeter)
RS (Radio Science)
XRS (X-Ray Spectrometer)
SPICE data for the mission

The data can be accessed from:

http://pds.nasa.gov/messenger_fake_url/

To access all data archived in the PDS, go to:

http://pds.nasa.gov.

January 6, 2008

[NASA] PDS ANNOUNCES NEW DELIVERY OF MARS ODYSSEY RADIO SCIENCE DATA

The Planetary Data System (PDS) is pleased to announce a new delivery of Odyssey data for the RSS instrument.

To access the RSS data, please visit the following link:
http://pds.nasa.gov/subscription_service/SS-20080103.html

To access all data archived in the PDS, go to:
http://pds.nasa.gov.

November 25, 2007

[NASA] PDS RELEASES NEW MARS SOIL ANALOG LABORATORY DATA SET

The Planetary Data System announces the release of a new laboratory data set, the Mars Analog Soil Observations from the Bloomsburg University Goniometer (BUG) Laboratory. These are bidirectional reflectance distribution function measurements of five Mars soil analogs, acquired using the MER Pancam flight spare filters. The data set was provided by Michael Shepard, Bloomsburg University, Bloomsburg, Pennsylvania.

The data is available at:
http://pds-geosciences.wustl.edu/missions/labdata/marsbug.htm

For access to all data archived in the PDS, go to:
http://pds.nasa.gov