IPDA Recommendations for the Content of the Archiving Component of an MOU for International Missions

1. Introduction

The purpose of this document is to provide guidelines to be considered when developing an MOU for international planetary missions. It defines the aspects and requirements that are necessary to assure that all scientific data obtained from international planetary missions will be open to the science public and can be widely shared as human intellectual assets. The first step in assuring that this is the case is to set up well-defined requirements that are included in a Memorandum of Understanding (MOU), the agreement between two or multiple agencies that defines the terms of interaction and obligations associated with a mission. The MOU, or an official appendix, should include requirements defining how international access to the data will be achieved. This should address data creation, availability, access and preservation, specify intellectual property rights and define and limit the access to information concerning spacecraft and instrument development and operations.

The International Planetary Data Alliance (See https://planetarydata.org/) has established the following standards for data and archives containing data from planetary missions:

- Data shall adhere to standards formulated by the International Planetary Data Alliance (IPDA) for data formats and documentation.
- In adherence to these standards, each agency shall maintain its own archive or utilize another recognized archive (i.e., ESA/PSA).
- Archives shall be constructed under local guidance from individual agencies with compatible structure to allow international accessibility and optimization of interoperability.
- In cases of international missions where agencies work together in archiving, a lead authority shall be identified in the MOU.
- Well-documented ancillary data (i.e. pointing and geometry information) shall be included.
- Schedules for data delivery shall be defined and published releases of the data shall be made available at the individual agencies and via international announcements.

Members of the IPDA are interacting to develop an archiving structure based on the NASA PDS4 data model. By adhering to these standards, mission personnel and participating scientists can develop data archives that are universally accessible and adapted to the needs of the users of the mother agency.

2. Archiving aspects to be considered in an MOU

In order to develop the optimal archive, it is necessary to begin interaction among all parties who are involved in setting up the data center and in developing a data pipeline at an early phase in the mission. This recommends that the requirements of data development and archiving be established as early as possible. Because the goal of the mission is to produce data that will satisfy science goals and that data is the public reward for the costs and effort that has gone into the mission, it is logical that data development and archiving be defined at the time that the MOU is developed and signed. The following components of this paper further specify the task of creating optimal archives.

2.1. Responsibilities for implementing data creation

The MOU should define the process for data creation. A straightforward way to specify this aspect is a **4W2H** based consideration:

- **W**hat: Acknowledging the mission and instruments and suppliers that will be involved.
- **W**ho/where: Identifying the lead agency
- **W**hen: Specify the expected period of operation
- **W**hy: Define the intent of the archiving effort
- **H**ow: The MOU would indicate that a data management plan will be generated
- **H**ow much: who supports the budget for producing the various components of the mission archive.

2.2. Necessary components for assuring data access

The following are necessary components for acquiring and preserving usable data sets and managing access and preservation. It is recommended that the MOU either address these aspects or specify that they be done in a Data Management Plan (DMP) that is generated through collaboration between designated personnel from the mission and data centers.

2.2.1. Data Standards and Formats – The MOU should specify the data standards that will be used. International partners are moving to adopt the IPDA standards, which are derived from the NASA PDS4 data model that utilizes a comprehensive global information model to produce an XML-based service-oriented system. This approach is flexible enough to allow legacy data (PDS3 and other formats) to be registered for top-down searches while optimizing preservation and access to incoming data.

2.2.2. Data generation and responsibilities – The MOU should specify that a DMP will be developed. In order to maintain data standards it is necessary that personnel from mission teams and the appropriate archiving centers establish a detailed data management plan that describes a schedule of well defined mile-posts and assigns responsibilities that are required to achieve quality data and efficient data pipe lines.

2.2.3. Data availability – The availability of data is determined by the effectiveness of documentation and search capability of the system. The DMP should specify the requirements for achieving adequate documentation and clear descriptions of the structure. In a PDS4 system all data are tied to a common model to improve data validation and search. Each product (including user help pages) is assigned a URN (Equivalent to an ISBN for books or a DOI for journals) and when registered in the system registry is accessible on the web.

2.2.4. Technical Information – The MOU should recognize that the scope of publication of information concerning spacecraft and instrument development and operations must be addressed. This could be deferred for inclusion in the DMP.

2.2.5. Data validation processes – The MOU should specify that a single validation process be mutually agreed on. One of the initial actions of a working group that is involved in acquiring and validating the data should be to agree on a single set of validation tools that will be utilized by all participants. The IPDA should be consulted for best practices.

2.2.6. Peer-review – The process should be defined in the DMP. The review consists of 2 parts: assuring that the data structure is consistent with defined standards and ascertaining that the data are of acceptable quality and sufficiently documented to allow "non-team" members to utilize the data.

2.2.7. Data accessibility – The MOU is an agreement between two or multiple agencies; therefore, the accessibility could be limited in these agencies or preferably be internationally accessible. The MOU should ensure direct access to the mission data by specifying the archive location and the means by which the scope of data use and accessibility will be determined through the generation of a joint DMP. The level of access (web, ftp, script, PDAP, etc.) in large part depends on the nature and quantity of the data.

2.2.8. Data rights and Intellectual Property Rights – In order that planetary scientific data can be used freely, the rules of intellectual property rights, including the individual agencies policies, should be clearly defined in the MOU.

2.2.9. Data delivery schedule – The manner, location/locations and delivery schedule should be specified in the MOU.

2.2.10. Data release policy – The MOU should specify any constraints on usage of the data. Preferably it would be universally accessible with proper citation.

2.2.11. Long-term preservation – An MOU is often prepared in each mission, and most of the provisions are available only when the mission is active. For data, provisions should be written considering long-term preservation after the mission ends. It is important that the MOU require that all instrument data acquired in accordance with this MOU be archived in digital form and properly maintained for the duration of this MOU.

2.3. Definitions of terms related to data archiving

The following are a list of proposed definitions to be used as a standardized vocabulary in negotiations of Memoranda of Understanding between space agencies.

Ancillary Data	This term is not recommended. Although it is used throughout the space science endeavor, it does not have a clear meaning. "Observation Geometry Data", "Engineering Data," and "Housekeeping Data" are preferred.
Archival	Of, relating to, contained in, suitable for, or constituting archives. [3]
Archive	(As a verb) To collect and store materials (such as recordings, documents, or computer files) so that they can be found and used when they are needed. [3] (As a noun) A collection of documents and materials or a place where they are held.
Calibrated Data	Raw data corrected for instrument properties and converted to physical units, which makes values independent of the instrument.
Data Management Plan	A plan for generation, validation, and transfer to the archives from the mission. The plan defines accepted data standards and specifies required data levels and documentation and how the communities will access the data.
Derived Data	Results that have been distilled from one or more calibrated data products (for example, maps, gravity or magnetic fields, or ring particle size distributions).
Engineering Data	Data about the operations, health, and safety state of a spacecraft, science instruments and/or supporting ground facilities and their subsystems (*e.g.,* temperatures, activities, system flags).
Housekeeping Data	An informal way to refer to Engineering Data.
Navigation Data	A set of measurements, properties, and supplementary information used to find the past, present and imminent future position, orbit and orientation of a spacecraft [1].
Observation Geometry Data	A specific type of supplementary information necessary to calibrate and interpret science measurements including: • Spacecraft location, • Spacecraft and instrument orientation (pointing), • Location, size, shape, and orientation of the target being observed, and • What events were occurring on the spacecraft or ground that might affect interpretation of science observations. [2]

Peer Review	A process by which scholarly work (such as a paper, documentation or data) is checked by a group of experts in the same field to make sure it meets the necessary standards before it is published or accepted. [3]
Radio Science Data	Science data collected using spacecraft and/or ground facility radio instrumentation for the purpose of precision navigation, inference of gravitational forces and relativistic effects, and/or study of propagation phenomena. Radio science also requires ancillary, engineering, and geometry data for proper computation and interpretation.
Raw Data	Original data from an instrument. If compression, reformatting, packetization, or other translation has been applied to facilitate data transmission or storage, those processes are reversed so that the archived data are in a PDS4-approved archive format.
Science Data	Any data from (or related to) science instruments, both received and processed.
Supplementary Information	Any data or documentation needed to take properties or measurements and incorporate them correctly into numerical computations and interpretations. [1] Sometimes referred to as "ancillary information."

1. CCSDS 500.0-G-3, *Navigation Data—Definitions and Conventions, Informational Report*, Issue 3 (May 2010)
2. The Navigation and Ancillary Information Facility SPICE Tutorials, http://naif.jpl.nasa.gov/naif/tutorials. Accessed on 14 May 2014.
3. Merriam-Webster Dictionary

2.4. Structure of Archives of Involved Space Agencies

The management structures of the data centers of IPDA members are quite diverse and frequently undergo revisions. However, this is an important piece of information, and in an effort to assure that updated information is available, a link on the IPDA website has been created (http://atmos.nmsu.edu/IPDA/ipda.html) where the current organization of PDS, PSA, JAXA, ISRO and other partners are described and regularly updated.