A SYSTEMS ENGINEERING APPROACH TO PLANETARY DATA ARCHIVE DEVELOPMENT.

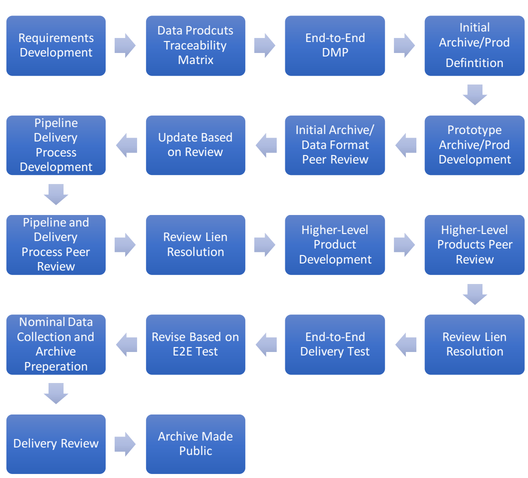
M. K. Crombie[1], [1]Indigo Information Services, LLC, Tucson, AZ (crombie@indiginfos.com).

Introduction: Since 1989, the Planetary Data System (PDS) has served as the archive for data products produced by NASA planetary missions [1]. These archives, apart from sample return missions, are the mission's final lasting scientific legacy. NASA requires missions to archive raw and calibrated scientific instrument data and encourages archiving of derived data products through language included in their mission Announcements of Opportunity (AO). NASA's policy is that all scientific mission data are in the public-domain and are to be released to the public through a NASA approved-data archive within a specified timeframe, usually not more than six months after data collection [2,3].

AO specifications give proposers links to PDS guidelines and standards for archive and data product preparation. Through these links the PDS provides a comprehensive set of standards that define the format, documentation and peer review necessary to deposit data products in the archive [4]. The mission-specific Data Management Plan (DMP) is the method by which proposers address AO specifications for description of the end-to-end data management of the mission including data policy, access and archiving. The DMP sets the stage for the project archive development lifecycle.

This abstract outlines a systems engineering approach to archive-lifecycle development from proposal though archive delivery. The systems engineering approach looks at the archive as an integral part of the complete mission system. Using this approach and the process shown in Figure 1, the archive is developed in coordination with mission requirements, mission operational objectives, and the ground data-processing system. Archive development that is integrated into the mission development process rather than as an afterthought, yields a more complete and comprehensive archive. By integrating the archive into the mission development process, investigation team members can provide insight and subject matter expertise during data product definition and design to ensure that products contain the correct information, are scientifically useful, and can be correlated or combined with other mission or outside data products.

Figure 1. Archive System Planning and Development Process



Archive System Planning: The first step in the archive system planning is to identify all data management requirements within the AO. Each AO will have specific requirements listed, but generally these will include items such as the selection of the NASA-approved archive, inclusion of raw, calibrated and derived data products, data formats, types and volumes, and delivery latency. As the mission concept is formulated, Level-1 and Level-2 science requirements (Figure 1, box 1) are developed to address the scientific objectives of the mission. Mission requirements are then evaluated against AO specifications to ensure all specifications are met. During this evaluation, the NASA-approved archive should be selected, and if PDS, a discipline node contacted.

Once Level-2 science requirements are developed, the data products necessary to meet these requirements can be formulated. The data products that satisfy Level-2 science requirements are often derived products that require input or analysis of lower-level products. A traceability matrix (Figure 1, box 2), a common systems engineering tool, that shows the relationship of lower-level products to the Level-2 science requirement data products can be used to develop the data product content of the mission archive from raw to calibrated to higher-level products. Once data product content is known, ancillary information necessary to document or interpret those products can be outlined. The data product traceability matrix including ancillary information is then used to develop the DMP.

Data Management Plan. The development of a project-level DMP (Figure 1, box 3) reduces mission risk by organizing each phase of data archive development from archive structure identification, data product development and documentation through peer review and delivery. Risk reduction is accomplished by the early planning and integration of the end-to-end data processing and data production procedures necessary to produce the data products that meet mission Level-2 science requirements. DMP concepts can also be used as input for ground data system design and development.

A complete DMP contains a top-level description of the mission data processing elements, their roles and responsibilities, and relationship to one another. There are several ways to set up mission data processing, including centralized or distributed processing systems. Any method can work, however there must be strong agreements within the investigation team on how, when, by whom, and to what standard products are produced. These agreements are documented in the DMP.

The DMP also outlines the relationship between the mission and the PDS. It is essential for mission management to support the integrated archiving process and for the PDS to have clear insight into the mission and archive development including product development, archive peer review and data delivery schedules. Finally, as the mission and archive development process proceeds, the DMP is updated and expanded to document changes and evolution of the data processing and archive systems.

Archive System Development: Once a mission is selected the archive system can be fully developed per the DMP, continuing the process shown in Figure 1. It is essential that a Mission Archive Scientist be identified and incorporated into the science team. Early development should re-visit the first three steps of the process to ensure that the archive and data products are compliant with mission requirements. The next step is to develop an initial outline of the archive and the products and documentation to be contained within. The initial outline of the archive includes how data are to be bundled, the collections of information within the bundles and the products within the collections. Once the initial archive outline is complete, the format of products to be archived can be defined using information from the data products traceability matrix.

Prototype archive and data product format development (Figure 1, box 5) is an essential part of the process. Prototypes are delivered to the PDS for an initial archive and data format peer review (Figure 1, box 6). This review happens early in the mission lifecycle. This step is necessary to ensure that the proposed archive is complete, and designed to PDS4 specifications and best practices. It also ensures that products either are or can be made compliant with the archive standards before in-depth development has started. The prototype archive and data products are updated based on review results. If necessary, a second review can be performed after changes are made.

Data products that are produced by automated processing procedures are called "pipeline" products. Generally, these are raw, reduced and calibrated instrument data products that are produced by the same algorithms throughout the mission lifecycle. The pipeline products, including collection and bundle information are developed (Figure 1, box 8) and peer reviewed to meet mission requirements and PDS4 standards (Figure 1, box 9). Liens from these reviews are addressed (Figure 1, box 10) before final acceptance. The lien resolution process can be iterative if necessary. High-level products (Figure 1, boxes 11 and 12) follow a similar development path to pipeline products.

Once the final peer review liens have been resolved, the pipeline and high-level data product production are put under configuration control. Changes can be made to the pipeline or high-level processes, but a change-control process is used to document and test the revisions. The scope of the change determines if additional peer reviewers are required. The change control process should be documented in a mission-to-PDS configuration control plan that sets the agreement on the processes and procedures used for configuration changes.

The configuration-controlled data processing pipeline is tested with real data, possibly collected during flight calibration activities, to perform an end-to-end delivery test (Figure 1, box 14). This test mimics how data will be delivered during nominal mission deliveries. Any liens generated by this test are addressed (Figure 1, box 15) through the configuration control process, and implemented before nominal deliveries commence.

Finally, data are collected as the mission plan specifies and the archive is generated based on the delivery schedule outlined in the DMP (Figure 1, box 16). The archive is delivered to the PDS, where a final delivery review (Figure 1, box 17) is conducted. Since the archive is configuration controlled, the delivery review is a rapid validation rather than a full peer review. The rapid delivery review allows the archive to be made public (Figure 1, box 18) shortly after the archive is delivered to the PDS.

Conclusions: The systems engineering approach to the archive planning and development process, including reviews, is general enough to be used as a guideline for many missions. It can be used by both archive scientists and mission systems engineers to develop requirements, plan data products, write a DMP, influence project ground system design, and design and develop configuration-controlled data processing pipelines resulting in a complete mission science archive that preserves data for future generations.

References:

[1] https://pds.nasa.gov/about/about.shtml [2] New Frontiers 4 AO (2016) NNH16ZDA008J. [3] Discovery 2014 AO (2014) NNH14ZDA014O. [4] https://pds.nasa.gov/pds4/about/