EPOXI

Data Management and Archiving Plan

Document Number
JPL D-39796

Final Version: 6 September 2007
Approved: 26 September 2007

Revision 1: 19 September 2008
Approved: 30 October 2008

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109-8099

Approvals for Revision 1: 19 September 2008

EPOXI Project Approval Signatures:

Approved by: Michael A'Hearn via email Date: 19 Oct 2008
M. A'Hearn, UMD, Principal Investigator

Approved by: Drake Deming via email Date: 21 Oct 2008
D. Deming, GSFC, Deputy Principal Investigator

Approved by: Albert Nakata for T. Duxbury via email Date: 20 Sep 2008
Albert Nakata for T. Duxbury, JPL, Program Manager

Data Provider Approval Signatures:

Approved by: Joseph Veverka via email Date: 30 Oct 2008
J. Veverka, CU, Cornell Science Data Center
(applies to HRII, HRIV, MRI data)

Approved by: Charles Acton via email Date: 19 Sep 2008
C. Acton, JPL, NAIF
(applies to SPICE data)

Approved by: Donald Yeomans via email Date: 30 Oct 2008
D. Yeomans, JPL, SSD/Navigation
(applies to Radio Science data)

Data Recipient Approval Signatures:

Approved by: Edwin Grayzeck via email Date: 20 Oct 2008
E. Grayzeck, GSFC, PDS Program Manager
(applies to all EPOXI data)

Approved by: Ludmilla Kolokolova via email Date: 24 Sep 2008
L. Kolokolova, UMD, PDS Small Bodies Node
(applies to all EPOXI data)

Approved by: Richard White via email Date: 22 Sep 2008
R. White, STScI, Multimission Archive
(applies only to data from the EPOCh phase)

Approvals for Final Version, 6 September 2007

EPOXI Project Approval Signatures:

Approved by: Michael A'Hearn via email Date: Sept. 6, 2007
M. A'Hearn, UMD, Principal Investigator

Approved by: Drake Deming via email Date: Sept. 17, 2007
D. Deming, GSFC, Deputy Principal Investigator

Approved by: Thomas Duxbury via email Date: Sept. 25, 2007
T. Duxbury, JPL, Program Manager

Data Provider Approval Signatures:

Approved by: Joseph Veverka via email Date: Sept. 26, 2007
J. Veverka, CU, Cornell Science Data Center
(applies to HRII, HRIV, MRI data)

Approved by: Charles Acton via email Date: Sept. 6, 2007
C. Acton, JPL, NAIF
(applies to SPICE data)

Approved by: Donald Yeomans via email Date: Sept. 6, 2007
D. Yeomans, JPL, SSD/Navigation
(applies to Radio Science data)

Data Recipient Approval Signatures:

Approved by: Edwin Grayzeck via email Date: Sept. 14, 2007
E. Grayzeck, GSFC, PDS Program Manager
(applies to all EPOXI data)

Approved by: Ludmilla Kolokolova via email Date: Sept. 6, 2007
L. Kolokolova, UMD, PDS Small Bodies Node
(applies to all EPOXI data)

Approved by: Richard White via email Date: Sept. 6, 2007
R. White, STScI, Multimission Archive
(applies only to data from the EPOCh phase)

Revisions

Draft 1:	April 13, 2007	Prepared by S. McLaughlin, EPOXI/UMD
Initial Release:	May 2, 2007	Delivered to signatories for review and approval
Second Release:	May 10, 2007	Revised per comments received from signatories; Distributed for the JPL Site Review on May 17th
Third Release:	August 10, 2007	Distributed for final signatures; Changes since the second release for the site review: - Clarified the primary pipeline will reside at the Cornell Science Data Center with redundant systems at UMD and JPL (2.2, 5.1.2, Fig. 1) - Removed clause stating pipeline operations will take place at the MOS (2.2) - Changed EPOXI to EPOCh (3.3 e) paragraphs) - Changed Cruise 3 to Cruise 2 (Tab. 1) - Corrected the title for E. Grayzeck - Added SDC to the acronym list
Fourth Release:	August 22, 2007	Changes to Revision 0 (requestor in parentheses) - Reference Documents section: Removed (MA) - 2.2: Changed “final archival volumes” to “peer-review and final volumes” (MA) - 3.4: Stated best-available SPICE kernels will be used for reduced products (MA) - 3.4, 5.1.1.b, Tab. 3: Clarified RSS data will only by provided for Boethin encounter (DY) - 3.4, 5.1.1.b, Tab. 3,: Reclassified Boethin ephemeris file as a SPICE product (DY,CA) - 3.5: Stated packet and ancillary data will be assembled for safekeeping to match text in 5.1.1.a (CA) - 4.2, Append. A: Clarified liens are required actions that must be taken before a data set can be archived (SBN) - 5.1.2.a: Removed last sentence about pipeline software (MA) - Append. A: Modified data set definition (SBN)
Fifth Release: (Final)	Sept. 6, 2007	Changes to Revision 1 (requestor in parentheses) - Title Page: Inserted JPL document number - Table 1: Moved Earth obs. from Cruise 1 to EPOCh phase (DD, SM) - 3.4.b: Reworded so that I/F products are not required for EPOCh (BC) - App. B: Corrected typo in Level 1 entry (BC)

Release 6	July 31, 2008	Revised delivery dates & data volumes for EPOCh extension and new DIXI target.
Release 7 (Rev 1)	Aug 15, 2008	Revisions per or by MA: - Adjusted revision numbers; Inserted new signature page with Al Nakata to sign for Tom Duxbury. - Table 1 Mission Timeline: Changed "Yearly" to "Several" Earth Swingbys; Changed Cruise 1 & 2 to simply Cruise. - Section 2.2: Clarified 3rd ¶; Removed reference to temporal data sets in 4th ¶ because data will be grouped by type of target. - Section 3.2: In 1st ¶ moved reference Table 2 to previous sentence. - Table 2 Data Volume Components: Improved description of calibration directory to avoid confusion about its contents. - Table 3 Data Volumes: Changed "Cruise 1/2" to "Cruise"; Inserted "in-flight" before "calibrations"; Inserted "High-Level" before "Derived Products"; Changed "Reduced" data to "Calibrated" for consistency. - Section 3.4, Imager a) & Spectrometer a): Inserted "in-flight" before "calibrations" at end of ¶. - Section 3.4, Imager c) & Spectrometer c): Adjusted text in first ¶ to avoid confusion about calibrated data vs files used to calibrate. - Section 3.4, Imager c) & Spectrometer c): Adjusted text in first ¶ to avoid confusion about calibrated data vs files used to calibrate. - Section 3.4, High-Level Products: Inserted "Derived" into "High-Level Products" for consistency. - Section 3.5 Safed Data: Pointed to section 5.1.1.a for a brief discussion of possible products for safing. - Section 4.2: In 2nd ¶ corrected project end date & added end date for EPOCh. - Section 4.3: Inserted nominal delivery dates final lien-resolved data. - Table 4 Data Deliveries: Added entries for EPOCh & DIXI peer reviews; Removed temporal references to Cruise 1 & 2; Inserted "High-Level" in "Derived product".

Release 8 (Rev 1)	Sep 19, 2008	Minor changes per B.Carcich: - Section 5.1.2.b: Inserted "team" at end of second-to-last sentence. - Sections 4.1 and 5.1.2.c: Stated the PDS catalog files are prepared by the archive teams with input or text descriptions from the science teams. - Section 3.2 and Table 2: Inserted illustration of a single volume data set to reflect text in 3.2.

Acronyms

C&DH	Command and Data Handling
CCSDS	Consultative Committee on Space Data Systems
CD	Compact Disc
CFDP	CCSDS File Delivery Protocol
Co-I	Co-Investigator
CSDC	Cornell Science Data Center
CU	Cornell University
DAP	NASA Data Analysis Programs
DMAP	Data Management Archive Plan
DI	Deep Impact mission
DIXI	Deep Impact eXtended Investigation science investigations
DOM	Distributed Object Manager (database for files produced and used by DI)
DSMS	Deep Space Mission System, provided by IPN
DSN	Deep Space Network
DVD	Digital Versatile Disc
EDR	Experiment (Raw) Data Record
EPOCh	Extrasolar Planet Observations and Characterization science investigations
EPOXI	Mission for the combined EPOCh and DIXI science investigations
FITS	Flexible Image Transport System
GB	Gigabyte(s)
GDS	Ground Data System
GSE	Ground Support Equipment
GSFC	Goddard Space Flight Center
HRI_IR	High Resolution Spectrometer - Infrared
HRI_VIS	High Resolution Imager - Visual
IPN	Interplanetary Network Directorate (at JPL)
IRAS	Infrared Astronomical Survey spacecraft
JPL	Jet Propulsion Laboratory
MAST	Multimission Archive at Space Telescope Science Institute
MMO	Mission Management Office (part of IPN)
MOS	Deep Impact/EPOXI Mission Operations System
MRI	Medium Resolution Imager
NAIF	Navigation and Ancillary Information Facility (services provided under IPN)
NASCOM	NASA Communications
NAV	Navigation and related services provided by DSMS
NSSDC	National Space Science Data Center
PDS	Planetary Data System (of NASA)
PI	Principal Investigator
RDR	Reduced or Calibrated Data Record
RS	Radio Science

Acronyms (continued)

SBN	Small Bodies Node of the Planetary Data System
SDC	Science Data Center
SFDU	Standard Formatted Data Unit
SIS	Software Interface Specification
SPICE	Spacecraft, Planet, Instrument, C-matrix, Events
SSD	Solar System Dynamics Group (at JPL)
STScI	Space Telescope Science Institute
TC&DM	Telemetry, Command and Data Management (services provided under IPN)
TCM	Trajectory Correction Maneuver
TDS	Telemetry Data Server
TMOD	Tracking and Mission Operations Directorate (renamed to IPN)
UMD	University of Maryland
VRML	Virtual Reality Modeling Language

Table of Contents

1. Introduction

This document defines the plans for archiving data from the EPOXI mission with the Planetary Data System (PDS), including the design, generation, validation, and transfer of the data archive to the PDS and the Multimission Archive at Space Telescope Science Institute (MAST). The archive will include raw and calibrated instrument data, calibration and navigation data necessary for the interpretation of the instrument data, and documentation. Data management for EPOXI will be nearly identical to that for Deep Impact (DI).

Section 2 gives an overview of the EPOXI mission including the Ground Data System by means of which the data stream will be converted into science data products. Section 3 provides an overview and design of the EPOXI data archive. Section 4 describes the steps of the archive generation process. Section 5 specifies the roles of each of the participants in the archiving process and assigns responsibility for each of the archiving functions. Section 6 provides the schedule for data archiving. Section 7 specifies the data release policy for the EPOXI mission.

2. Overview of the EPOXI Mission

2.1 Mission Overview

The EPOXI mission is the combination of two independently proposed scientific investigations using the flyby spacecraft of Deep Impact in an extended mission. EPOCh (Extrasolar Planet Observation and Characterization) will utilize the HRI CCD and IR spectrometer on the DI flyby spacecraft to a) observe transits by known extrasolar planets (hot Jupiters) to look for perturbations due to terrestrial planets in low-order resonances, b) search those same systems for secondary occultations of the hot Jupiter by the star, and c) characterize the Earth as a remotely sensed planet. DIXI (Deep Impact eXtended Investigation) consists of a flyby of 103P/Hartley 2 in order to study a second short-period comet with the same set of instrumentation as was used at 9P/Tempel 1. The original target was 85P/Boethin. When the comet could not definitively be recovered after extensive observations during 2007, the mission selected the backup, 103P/Hartley 2, as the new target for DIXI.

Table 1 presents a high-level timeline for the mission. It includes revisions for the change of target for DIXI and the extension of EPOCh observations to recover from the suspension of imaging for most of March 2008 due to a problem with the downlink signal.

Table 1 **EPOXI Mission Timeline**		
Phases and Events	**Start Date**	**Stop Date**
Cruise	**2007-09-01**	**2008-01-21**
Earth Encounter/Flyby	2007-12-31	2007-12-31
In-flight Calibrations	2007-09-01	2008-01-21
EPOCh	**2008-01-22**	**2008-08-31**
Extrasolar Planet and Earth Observations; In-flight Calibrations	2008-01-22	2008-08-31
Cruise	**2008-09-01**	**2010-09-03**
In-flight Calibrations; Several Earth Flybys	2008-09-01	2010-09-03
DIXI	**2010-09-04**	**2010-12-16**
Hartley 2 Approach	2010-09-04	2010-11-03
Hartley 2 Encounter/Flyby	2010-11-04	2010-11-04
Lookback Imaging	2010-11-04	2010-11-25
In-flight Calibrations	2010-09-04	2010-12-16

2.2 Ground Data System Overview

EPOXI will use the existing DI Ground Data System (GDS) and Mission Operation System (MOS). The GDS is the mechanism by which the raw spacecraft data stream will be converted to science data products, including those to be archived with the Planetary Data System under project direction. The MOS, implemented for DI at the Jet Propulsion Laboratory in Pasadena, will be responsible for monitoring the status of the spacecraft and payloads, commanding the spacecraft and payloads, coordinating real-time mission planning, operating the telemetry data services and producing navigation and ancillary data in the form of SPICE kernels. The MOS will receive packets from the DSN and place these on the Telemetry Data Server (TDS). The MOS File Delivery Protocol (CFDP) process will produce reconstructed science images and spectra, navigation images, and instrument temperatures from the telemetry data and place these on the Distributed Object Manager (DOM). The DOM will push data to the primary science data pipeline running at the Cornell Science Data Center (CSDC) and to redundant pipelines located at the University of Maryland (UMD) and the Jet Propulsion Laboratory (JPL). Patterned after the DI data flow, the CSDC will provide the processing, analysis, and archive preparation of the science and navigation data, and will allow members of the EPOCh and DIXI science teams to access the raw and calibrated data located at Cornell, UMD, or JPL.

The CSDC will accept and process images and spectra into Experiment (Raw) Data Records (EDR) and Reduced Data Records (RDR) as FITS files using the data processing and calibration pipeline developed and implemented for DI. The CSDC will also collect and format instrument thermal telemetry data from DOM. For DI, a subcontractor

externally produced the PDS labels for the EDR and RDR products. For EPOXI, the CSDC will enhance the existing DI data pipeline to generate PDS labels for all EDR and RDR products. The CSDC will also receive and organize calibration files, and will coordinate with the JPL navigation team to develop the comet kinematics model and shape model.

Once science EDR and RDR products and PDS labels have been produced by the CSDC they will be transferred to the PI teams at the UMD and Goddard Space Flight Center (GSFC) where the content and format will be validated and the volumes for the peer review, and later the lien-resolved volumes, will be prepared for delivery to PDS Small Bodies Node (SBN).

In addition to delivering all data to PDS-SBN, the data relevant to EPOCh will be delivered as FITS files to the Multimission Archive at Space Telescope Science Institute (MAST). Since the FITS files and labels are internal to the PDS archive, this is merely a second delivery path. The organization of the archive will be patterned after DI.

A comprehensive description of the GDS is provided in the DI Mission Operations Plan. A data flow diagram for the GDS is shown in Figure 1 inserted at the end of this document.

3. The EPOXI Archive

3.1 Overview

The archive for EPOXI will be organized similarly to that for Deep Impact. Raw and calibrated data of the EPOCh transit targets, the EPOCh Earth target, and the DIXI target will be grouped into separate data sets by target, instrument, and processing level. The EPOCh transit targets will be combined in one data set. All raw calibration-related data from all phases will be gathered into one data set per instrument. EPOXI will use the same formats as DI for data files and PDS labels.

Like the DI archive, the EPOXI archive will contain science data products from each of the instruments, instrument calibration data, navigation and ancillary data in the form of SPICE kernel files, and sufficient documentation of the data and the mission to enable scientists to understand and use the archive well into the future. To produce this archive a number of steps need to be carried out, including design of the archive structure and contents, generation of the archive components, peer review with the PDS, and final packaging and delivery. The science data products form the core of the archive; a list of the expected data products from each of the instruments is given in Table 3. The entire EPOXI archive is expected to be 300-320 gigabytes (GB) in size. The archive will be online at the PDS Small Bodies Node, consistent with current PDS practice. MAST will also make the EPOCh data available online. PDS Small Bodies Node is responsible for transferring EPOXI data to the National Space Science Data Center (NSSDC) for the deep archive and will negotiate the media at the time of the delivery.

3.2 Archive Structure

The EPOXI archive will be broken down into logical data sets by target type (Epoch star, Epoch Earth, Hartley 2, or calibration), instrument, and reduction level. Logical data sets larger than 4.7 GB will be broken into physical volumes by specified time intervals. Each volume will contain the top-level directories required by PDS, described in Table 2. The requirement may be relaxed as SBN has started to allow a large data set (> 4.5 GB) to be delivered as one logical and physical volume. Figure 2 illustrates a logical data set organized as one physical volume or separated into multiple physical volumes. Data types and data volume estimates for each archive component and for the total archive are shown in Table 3.

Table 2
Top-Level Components of a Data Volume

Directories	Contents
CALIB	PDS-labeled files, such as darks and flat fields, and any other parameters needed to convert raw data to calibrated data or other higher-level products.
CATALOG	The catalog files (text files) required by PDS to document the mission, spacecraft, instruments, and data set.
DATA	Data products such as raw HRIV CCD images or calibrated HRII spectral images and PDS labels, depending on the data set or volume. Within this directory, data will be subdivided by reduction level (raw, reduced, reduced and cleaned) then by year and day of year.
DOCUMENT	Text files serving as documentation for the archive; Documentation may be delivered as a separate PDS volume as was done for DI.
INDEX	Index files to enable the user to find the data of interest.

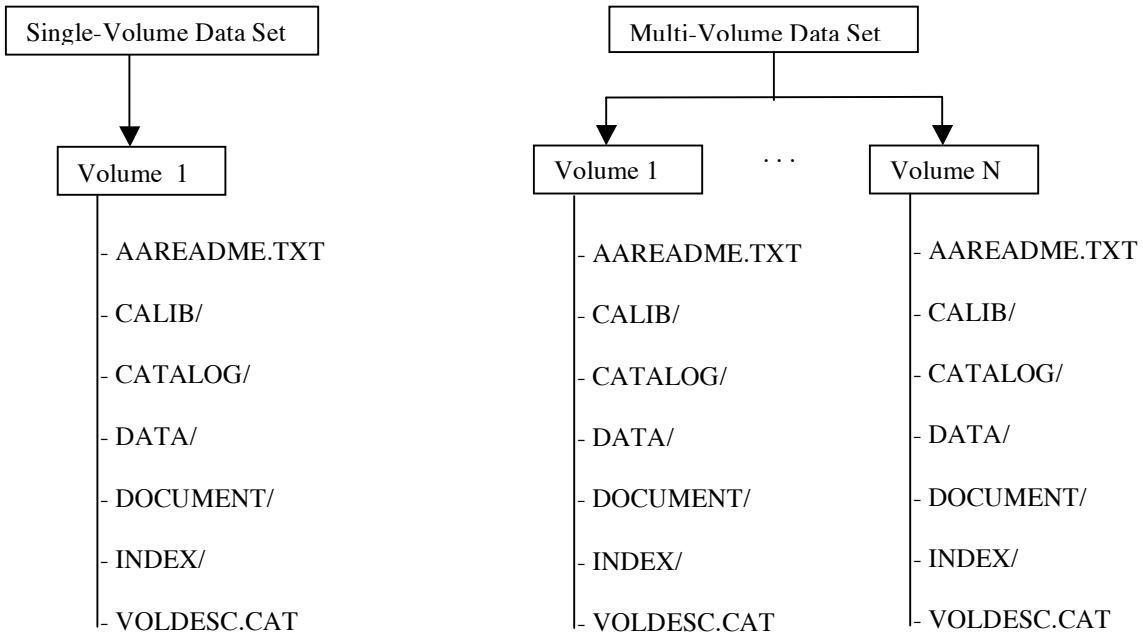

Figure 2: Left - One logical data set organized as one physical volume; Right - One logical data set organized into multiple physical volumes

Table 3
Data Volumes for Archive Components

Archive Component	Target or Phase, Data Products	# Data Sets	Est. Vol. (GB)
Imagers (HRIV, MRI)	All (Cruise, EPOCh, DIXI)		
	- Raw in-flight calibration images	2	15
	EPOCh		
	- Raw science images	2	70+
	- Calibrated science images, calibration files	2	120+
	DIXI		
	- Raw science images	2	5
	- Raw Nav images	2	5
	- Calibrated science images, calibration files	2	10
	- Calibrated Nav images, calibration files	2	10
Spectrometer (HRII)	All (Cruise, EPOCh, DIXI)		
	- Raw in-flight calibration spectra	1	20
	EPOCh		
	- Raw science spectra	1	10
	- Calibrated science spectra, calibration files	1	25
	DIXI		
	- Raw science and calibration spectra	1	5
	- Calibrated science spectra, calibration files	1	15
Instrument Thermal Data	Tables of averaged instrument temperature measurements from telemetry for all phases	1	0.2
Radio Science	Trajectory estimates and supporting products used to determine Radio Science results for the Hartley 2 encounter	1	0.2
SPICE	Navigation and ancillary kernels for all phases; comet Hartley 2 ephemeris kernel	1	5
High-Level, Derived Products	- Shape models	1	0.2
	- Surface temperature maps	1	0.2
	- Photometry (EPOCh, DIXI)	2	0.2
Documentation	- Observing logs - HRIV/HRII slit context maps - Calibration documentation - Data pipeline software (IDL source code as documentation only) - Publications - DMAP, Software Interface Specifications (SIS)	n/a	1.0

3.3 Volume Documentation Files

PDS requires a number of documentation files for each volume. One is an aareadme.txt, a text file describing the contents of the volume. Also required is voldesc.cat, a catalog of all the files residing on the volume. Each of the sub-directories under the top-level directory, except for the data sub-directory, also requires one file to document the contents of that directory. The details of these files are specified in the PDS Standards Reference .

3.4 Data Products

A summary of the expected data products, to be provided by the CSDC and science team members, is provided next. The EPOXI data products to be delivered to the PDS and MAST will have the same format as those generated by Deep Impact for the PDS archive.

Imagers (HRIV and MRI)

a) Raw Products

Each validated raw data product (EDR) will be stored as a 2-dimensional FITS image with an image extension for a pixel quality map. A detached PDS label will be generated for each EDR FITS file. EDRs will be generated for mission targets and calibration sources. EDRs will be grouped into data sets by science or navigation mode, target type (Epoch Star, Epoch Earth, Hartley 2, or in-flight calibration), and instrument.

b) Reduced Products

Raw images of the mission targets will be calibrated and included in the archive. Calibrated data products (RDR), suitable for each mission phase as determined by the science teams, will be provided such as radiance (reversible, i.e., not cleaned), cleaned radiance, and cleaned irradance-over-flux (I-over-F). Each RDR will be stored as a 2-dimensional FITS image with image extensions for a pixel quality map and a signal-to-noise ratio map. Pointing and observational parameters will be computed using the best available SPICE data and will be included in the FITS header of each RDR. A detached PDS label will be generated for each RDR FITS file. RDRs will be grouped into data sets by science or navigation mode, target type (Epoch Star, Epoch Earth or Hartley 2; “reduced” calibration spectra will not be delivered), and instrument.

c) Calibration Files

Calibration files that are used to produce calibrated and higher level products from raw products, such as dark frames and filter transmission data, will be provided as

FITS images or ASCII tables with detached PDS labels. Calibration files will be included in data sets containing RDRs.

d) Documentation Files

Observing logs, a calibration report, and calibration software will be provided as documentation. Following Deep Impact as an example, EPOXI will provide documentation volumes for the EPOCh and DIXI portions of the mission.

Spectrometer (HRII)

a) Raw Products

Each validated raw data product (EDR) will be stored as a 2-dimensional FITS image (along-slit distance versus wavelength) with an image extension for a pixel quality map. A detached PDS label will be generated for each EDR FITS file. EDRs will be generated for mission targets and calibration sources. EDRs will be grouped into data sets by target type (Epoch Earth, Hartley 2, or calibration) and instrument.

b) Reduced Products

Raw spectra of the mission targets will be calibrated and included in the archive. Two different types of calibrated data products (RDR) will be provided: radiance (reversible, i.e., not cleaned) and cleaned radiance. Each RDR will be stored as a 2-dimensional FITS image with image extensions for a pixel quality map, a pixel wavelength map, a pixel resolution map, and a signal-to-noise ratio map. Pointing and observational parameters will be computed using the best-available SPICE data and will be included in the FITS header of each RDR. A detached PDS label will be generated for each RDR FITS file. RDRs will be grouped into data sets by (Epoch Earth or Hartley 2; "reduced" calibration spectra will not be delivered) phase and instrument.

c) Calibration files

Calibration files that are used to produce calibrated and higher level products from raw products, such as bad pixel maps and dark frames, will be provided as FITS images or ASCII tables with detached PDS labels. Calibration files will be included in data sets containing RDRs.

e) Documentation Files

HRIV/HRII slit context maps, observing logs, a calibration report, and calibration software will be provided as documentation. Following Deep Impact as an example, EPOXI will provide documentation volumes for the EPOCh and DIXI portions of the mission.

Based on experience from Deep Impact, spectra acquired during the DIXI encounter phase may not be well calibrated before mission funds expire. For this scenario, the best possible set of calibrated data will be delivered. Then traditional DAP funding will be used to improve the calibration and deliver revised, calibrated data sets.

Instrument Thermal Data

Temperature measurements from sensors located in or near the HRII, HRIV, and MRI instruments as well as the HRI and MRI telescopes will be extracted from raw thermal data and provided as ASCII tables with detached PDS labels. Data will nominally be divided by EPOCh and DIXI and delivered as two data sets with appropriate documentation.

Radio Science

Raw tracking data files containing Doppler and sequential range measurements will be produced for the encounter with comet Hartley 2. Additional products will include calibration files for the ionosphere and troposphere and a list of spacecraft maneuvers during the encounter. All radio science data products will be provided as ASCII tables with detached PDS labels. Data will delivered as one data set with appropriate documentation.

SPICE

SPICE data will include files (kernels) of spacecraft, planet and comet Hartley 2 ephemerides, instrument mounting alignments, spacecraft orientation, and data needed for relevant time conversions. The SPICE data will be provided using the standard SPICE formats for ASCII or binary kernels with detached PDS labels. Data will nominally be divided by EPOCh and DIXI and delivered as two data sets with appropriate documentation.

High-Level Derived Products

Although high-level derived products will be started for DIXI, these products may not be delivered before mission funds expire. For this scenario, traditional funding from the NASA Data Analysis Programs (DAP) will be used to complete and deliver these products.

High-level products such as shape models in ASCII table format and surface temperature maps in FITS will be produced by the science team and delivered as separate data sets with appropriate documentation.

3.5 Safed Data

Some packet data and ancillary files, as described in section 5.1.1.a, will be assembled and placed on CD-R or DVD media for long term safekeeping in the event problems are

discovered with the formally archived higher-level data products. Copies of these data and allied documentation will be provided to at least the PIs, the CSDC at Cornell University, and Navigation and Ancillary Information Facility (NAIF). These data will not be delivered to PDS-SBN and will not undergo a PDS peer review.

These same data will be kept by JPL/IPN for a minimum of three years after end of mission.

4. The Archive Generation Process

The major steps to be followed in generation of the EPOXI archive are described in this section. Responsibilities for generating archive components are specified in Section 5.

4.1 Archive Preparation

As described in the previous section, data products will be generated as FITS files or ASCII tables with formats compatible with both the PDS and MAST. EPOXI data files will be formatted using conventions already approved by PDS for Deep Impact. Each data file will be accompanied by a PDS label (i.e., a separate ASCII text file that describes the content and structure of the accompanying data file). For HRII spectra and HRIV and MRI images, ancillary information describing the observing conditions and spacecraft state will be extracted from FITS headers and placed in the PDS labels.

Parties generating the data products will prepare files documenting the archive components. In general, all information necessary to interpret and use the data is to be included in the archive.

PDS standards call for the documentation of the mission, spacecraft, instruments, and data products with special ASCII text files called catalog files that must be filled out with prescribed information for each data set. These catalog files will contain much of the information necessary to document the archive and should make it possible for scientists to make correct use of the data in the future when the mission personnel are not available to support them. EPOXI will use catalog files from the Deep Impact data sets as templates. Science team members will provide text descriptions of the mission, spacecraft, instruments, and data sets for the catalog files that will be prepared by the EPOCh and DIXI archive teams.

4.2 Archive Validation

Data validation falls into two types, validation of the data themselves and validation of the compliance of the archive with PDS archiving requirements. The science team will perform the first validation, and the second will be overseen by the PDS-SBN in coordination with the science team. The separate delivery of data by mission phases will facilitate validation by ensuring that problems in the early deliveries are resolved by the time of the later deliveries.

The formal validation of data content and visibility, adequacy of documentation, and adherence to PDS archiving standards is finalized through external peer reviews coordinated by the PDS-SBN. Liens (action items) placed by the reviewers or PDS personnel must be corrected before the affected data set(s) can be archived. The data providers identified in section 5 must resolve liens if received in time to address them (i.e., at least 45 days before the end of EPOCh on September 30, 2009 and for DIXI at least 45 days before the end of the project on September 30, 2011). Any other changes (not liens) recommended during the review will be addressed as time, resources, and feasibility allow. PDS will spot-check lien resolution of the final products.

4.3 Archive Packaging and Delivery

The delivery of data to PDS for peer reviews will take place in stages, as specified in Table 4. After the PDS peer reviews, the final lien-resolved data will be delivered to PDS-SBN and the MAST, nominally by October 2009 for EPOCh and October 2011 for DIXI. Each delivery will be made by electronic transfer or on a hard medium such as CD or DVD as negotiated between EPOXI and PDS-SBN and MAST at the time of delivery.

Table 4
Data Delivery Schedule to PDS for Peer Review

Product	Delivery Date	Est. Vol. (GB)
Raw Calibration data (Mission start through EPOCh)	2009-02-28	35
Raw EPOCh data	2009-02-28	80+
Calibrated EPOCh data with calibration files	2009-02-28	145+
Instrument Thermal data (Mission start through EPOCh)	2009-02-28	0.1
SPICE data (Mission start through EPOCh)	2009-02-28	3
EPOCh Peer Review	2009-04 to 06	n/a
Raw Calibration data (End of EPOCh through DIXI)	2011-06-30	25
Raw DIXI data	2011-06-30	15
Calibrated DIXI data with calibration files	2011-06-30	35
Final SPICE data (for the entire mission)	2011-06-30	5
Radio Science data for DIXI	2011-06-30	0.2
Instrument Thermal data (Cruise 2 through DIXI)	2011-06-30	0.1
High-level Derived products	2011-06-30	0.5
DIXI Peer Review	2011-07 to 08	n/a

5. Roles and Responsibilities

This section describes the roles and responsibilities of the personnel and organizations involved in generating, validating, transferring, archiving and distributing the EPOXI archive.

5.1 Responsibilities of the EPOXI Project

The EPOXI archive consists of products generated by the project Mission Operations System (MOS) and by the Project Science Team, comprised of the PIs, Co-Is, the EPOXI archive team, and the Cornell Science Data Center.

5.1.1 Responsibilities of the Project Mission Operations System

Essentially all elements of the EPOXI operations team at JPL contribute directly or indirectly to the EPOXI archive. As such all MOS elements are responsible for ensuring the data provided by them are as complete and correct as possible under the terms of each element's contract with the EPOXI project or NASA. Major MOS archiving responsibilities are described below.

a) The Deep Space Mission Systems (DSMS) is responsible for making packet and CCSDS file data available to the Project Science Team and other project teams via the TDS and DOM servers, respectively. This service includes ensuring that 95+% of all data received at DSN stations are made available to the EPOXI teams. The DSMS also provides images reconstructed from transfer frames and raw instrumental thermal data; these are placed in the DOM database. EPOXI project elements needing these data, such as the CSDC, are responsible for querying for the data from the TDS and/or DOM as needed. Query tools and/or instructions are provided by DSMS. The DSMS services are multi-mission services provided by JPL's Interplanetary Network Directorate (IPN).

   DSMS is also responsible for producing a "safed" set comprised of all packet data and a selection of important file products produced by DSMS, such as the SCLKvSCET file and the Predicted Events File (PEF). This "safed" set will include any Software Interface Specifications (SIS) for the file products and appropriate documentation for all the contents. Its construction will generally follow PDS standards, but the product will not be peer reviewed or formally catalogued. Copies will be provided to the PIs, the CSDC, and the NAIF Node of the PDS.

b) The NAIF is responsible for producing a full suite of SPICE kernels (including an ephemeris file for comet Hartley 2), preparing PDS-formatted data sets of SPICE files, and delivering the data sets to PDS-SBN. As the navigation node of the PDS, NAIF will maintain copies of these SPICE files and make them available to the science community, along with the associated SPICE Toolkit: but this is accomplished as a responsibility of the PDS. As for DI, during mission operations NAIF will place the EPOXI SPICE kernels on the "naïf" server at JPL from which

PDS-SBN and other science team members, if interested, can download them. JPL's IPN Directorate provides NAIF multi-mission services.

c) The Solar Systems Dynamics group (SSD) at JPL is responsible for producing Radio Science data files for the encounter with comet Hartley 2 during the DIXI phase of the mission. SSD will provide a PDS-formatted data set to the EPOXI archive team for delivery to PDS-SBN.

d) The EPOXI Project Office will make available all non-proprietary and non-ITAR-restricted documents for possible inclusion on the EPOXI data volumes.

5.1.2 Responsibilities of the Project Science Teams

The EPOCh and DIXI science teams have the overall responsibility of setting observing priorities for the spacecraft instruments in order to meet the fundamental science goals of the EPOXI project. The ensemble data from these observations constitute the raw science data - the fundamental element of the archive.

Each of the entities comprising the Science Teams has major responsibilities in producing and validating the peer-review volumes and resolving any liens if provided in a timely manner (see 4.2):

a) Cornell Science Data Center

The CSDC is responsible for generating and validating the raw and calibrated science data files and PDS labels, the instrument temperature tables, the comet shape and kinematical models, and the source code and associated documentation of software used to generate these data products. The CSDC will provide the source code strictly as documentation and be responsible for archiving it; however, the CSDC will not support the code after the end of the mission.

The CSDC will operate and maintain the primary data pipeline at CU and the redundant pipelines at UMD and JPL.

The CSDC will participate in the PDS-sponsored peer review of data products, and will take actions needed to clear any liens against the data it has provided.

b) PIs and Co-Is

The EPOCh and DIXI PIs have the ultimate responsibility for delivering to the project a complete and validated archive of all files necessary for final data calibration.

The PI and Co-Is who generate high-level or derived digital data products delineated in Table 3 are responsible for producing archival quality data sets, including calibration data, relevant software, and documentation consistent with PDS standards.

The PIs are responsible for clearing all liens that result from the PDS peer review process on data products generated by the science team. The Co-Is will participate in this process if their data products are involved.

c) Archive Teams

The EPOCh and DIXI archive teams are responsible for producing all PDS-compliant data volumes: a preliminary set for PDS peer reviews and a final set for archiving in the PDS and MAST. The PIs and the archive teams at GSFC and UMD will coordinate this work. Approved copies of the final volumes will be distributed to the science team members (if requested), the PDS, and the NSSDC. The final raw and reduced volumes for the EPOCh phase of the mission will also be distributed to the MAST.

The archive teams will prepare the Software Interface Specification (SIS) documents in consultation with the EPOXI project. These documents define the structure and content of the data products and data sets and are needed both for internal use and for future users of the archive. The archive teams are also responsible for preparing the catalog files required by PDS (with input from the science teams) as well as organizing any documentation.

5.2 Responsibilities of the PDS

The SBN is the lead PDS node for interfacing with the DI mission and will be supported by the NAIF and the Engineering nodes. Specific functions of PDS are listed below.

a) Consult on Archive Generation: Support the generation of the archive by advising the project/science teams on PDS archive standards, requirements and documentation needs. The SBN will support the data validation activity to ensure that the formal peer review process, a requirement for data ingestion into PDS, proceeds with a minimum of problems.

b) Conduct formal peer reviews of the data sets. This is a PDS mandated step before any data can be ingested into PDS.

c) Offer support to the PIs, the EPOCh archive team, and CSDC in the resolution of liens that arise in the course of the peer review.

d) Maintain the EPOCh archive collection on-line for access by the planetary science community. PDS is also to offer expert advice to customers requesting help with the use of these products.

e) Replicate the archival volumes for distribution to the NASA supported science community whenever physical media are judged to be appropriate

f) Provide a copy of the archival volume set to the NSSDC for deep archive purposes, and for replication to serve any requests coming from the general public.

5.3 Responsibilities of the MAST

The Multimission Archive at Space Telescope Science Institute will archive raw and reduced data from the EPOCh phase of the mission. Data submitted to MAST must be the final version delivered to the PDS (i.e., peer-reviewed and lien-resolved). Specific functions of the MAST are listed below.

a) Consult on archive generation: Support the generation of the archive by advising the project/science teams on the MAST archive standards, requirements and documentation needs.

b) Maintain the EPOCh archive collection online for access by the "Origins" and astronomical communities.

c) Provide assistance for preparing indices required for loading raw and reduced EPOCh data into the MAST.

6. Schedule for Archive Generation, Validation, and Delivery

The principal archive elements, namely the science data products defined in Section 3.4, will be generated during the course of the mission, as will many ancillary products such as SPICE files. The general guideline for Discovery missions is that they deliver archive quality volumes to PDS at intervals not exceeding six months after receipt of the data used to make the products contained on the volumes. The planned timeline for archive delivery to PDS for peer review was previously shown in Table 4.

Following the delivery of data volumes to PDS is a several-month-long period in which the data will be peer reviewed by PDS. The Project and the appropriate science team members must resolve all liens identified during the peer review process before operations cease. Final acceptance of the data by PDS will occur only after all liens have been cleared.

7. Data Release Policy

There are no proprietary data rights for the EPOXI Mission. Science team members do have a limited amount of exclusive time–not to exceed six months–for validation of data prior to delivery to the PDS. MAST will publicly release the EPOCh data on a date negotiated between MAST and the PI of EPOCh.

Fully reduced, calibrated and corrected data products will be produced by the science team for delivery to PDS per the schedule given in Table 4 and Section 6. The PIs and science teams are responsible for coordinating all scientific investigations involving the use of calibrated data from their respective instruments and ensuring that all science data products are delivered in a timely fashion.

Appendix A. Glossary of PDS Terms

Catalog file	A catalog file is a PDS-required text file formally documenting the details of a mission, spacecraft, instrument, or data set.
Data product	A data product is a single instance of the output of an instrument, such as an image or spectrum. For instruments having a continuous output (such as a particle counter) there need be an artificial binning applied to create discrete data products.
Data set	A data set is a collection of data products, such as images or spectra.
Label	A label is an attached or detached text header that formally describes the structure and content of a data product.
Lien	A lien is placed by reviewers or PDS personnel and is an action that is required to be taken before a data set can be archived.
Volume	A volume is a single CD, DVD, or other type of a storage medium. One volume may contain one full data set or multiple data sets. Large data sets may span several physical volumes.

Appendix B. Data Processing Level Definitions

The following data processing level definitions can also be found in the PDS Standards Reference document

CODMAC Level	Proc. Type	Data Processing Level Description
1	Raw Data	Telemetry with data embedded.
2	Edited Data	Corrected for telemetry errors and split or decommutated into a data set for a given instrument. Sometimes called Experimental Data Record (EDR). Data are also tagged with time and location of acquisition. Corresponds to NASA Level 0 data.
3	Calibrated Data	Edited data that are in units produced by instrument, but have been corrected so that values are expressed in or are proportional to some physical unit such as radiance. No resampling, so edited data can be reconstructed. Corresponds to NASA Level 1A.
4	Resampled data	Data that have been resampled in the time or space domains in such a way that the original edited data cannot be reconstructed. Could be calibrated in addition to being resampled. Corresponds to NASA Level 1B.
5	Derived Data	Derived results, as maps, reports, graphics, etc. Corresponds to NASA Levels 2 through 5
6	Ancillary Data	Non-Science data needed to generate calibrated or resampled data sets. Consists of instrument gains, offsets; pointing information for scan platforms, etc.
7	Correlative Data	Other science data needed to interpret space-borne data sets. May include ground based data observations such as soil type or ocean buoy measurements of wind drift.
8	User Description	Description of why the data were required, any peculiarities associated with the data sets, and enough documentation to allow secondary user to extract information from the data.
N	N	Not Applicable

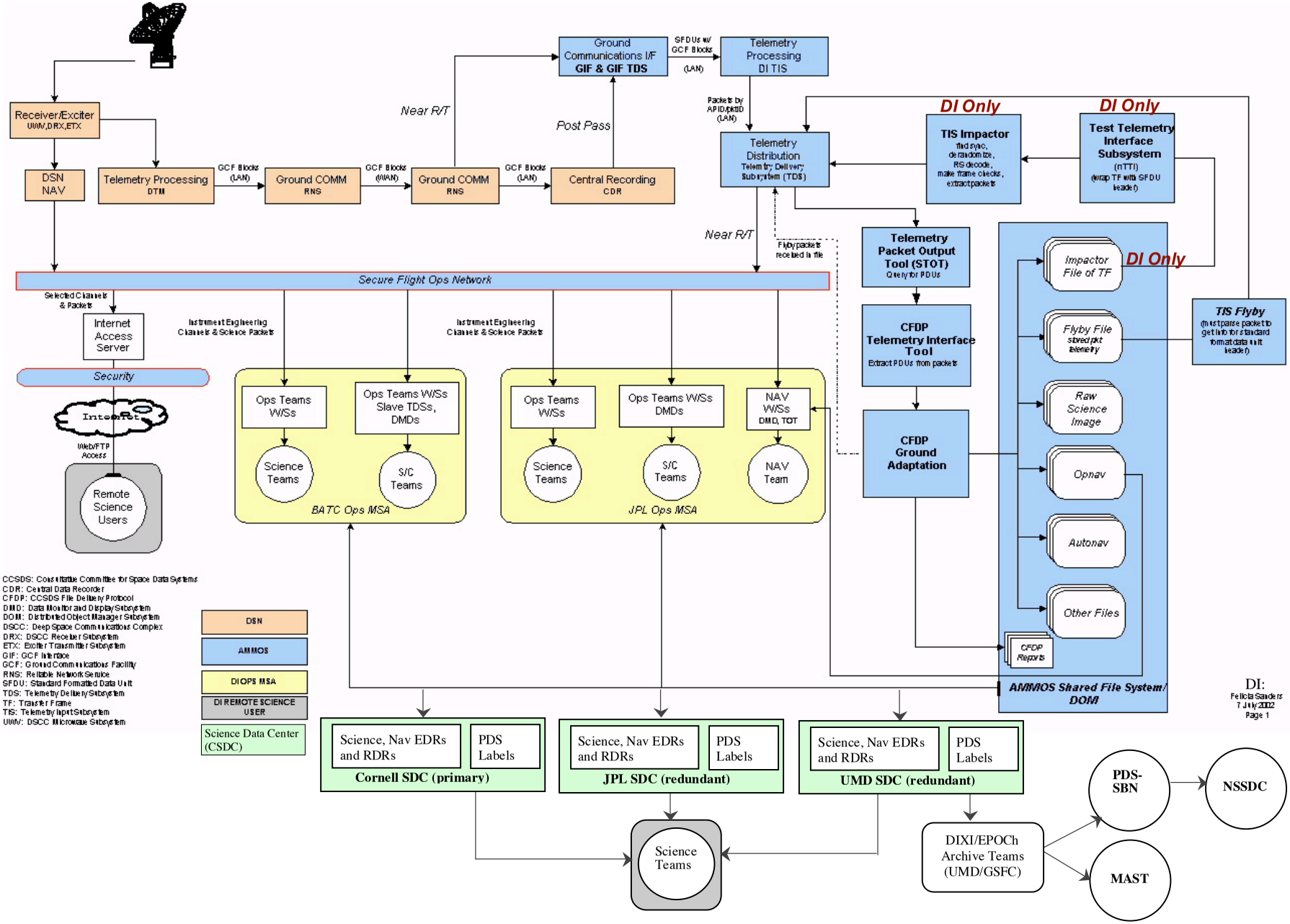

Figure 1: EPOXI/DI Ground Data System including downlink and SDC data flows