

**NASA SPACE SCIENCE DATA COORDINATED ARCHIVE  
ARCHIVE PLAN FOR 2016 – 2017**

Ed Grayzeck  
National Space Science Data Center  
Greenbelt, Maryland 20771  
2016-03-29

## Abstract

This data management plan shows that the NASA Space Science Data Coordinated Archive (NSSDCA) expects to accept ~361 TB of data into the archive in 2016 and ~373 TB in 2017.

## 1. Introduction

NSSDCA provides a vital service as NASA's permanent multi-disciplinary Space Science archive. Its curation activities are essential to ensure that space science data will continue to be available and usable into the indefinite future. The need for long-term curation arises because in most cases the full value of any set of data cannot be known in advance. New science discoveries or changes in research and exploration priorities may make older data, seldom noticed before, suddenly highly relevant.

This data management plan summarizes the expected data inflow to NSSDCA (note the Acronym list at the end of this document) for the years 2016-2017. NSSDCA data management plans cover 2 years and are updated annually. Note these are estimates for planning purposes, not exact data projections.

### 1.1. Levels of Service

NSSDCA accepts and archives data under four levels of service, summarized in Table 1 below. The most familiar is the Permanent Archiving of data, but, as defined in MOUs with various data providers, it also provides Backup service, mostly for other Archives. The Analog Archive includes photos, maps, microfilm, microfiche, documents, etc, some analog copies of digital data and others supporting metadata; it is included in this list for completeness.

<b>Permanent Archive: AIPs</b>	Preservation of digital data in Archival Information Packages delivered by a data producer or created at NSSDCA. AIPs are re-written to new media within six years. Data is disseminated by NSSDCA if not available through an active archive or per MOU.
<b>Permanent Archive: non-AIP digital data</b>	Preservation of non-packaged data on various media types. Data will eventually be migrated from legacy media to AIPs, though no media refresh will be made in the meantime. Data is disseminated by NSSDCA if not available through an active archive or per MOU.
<b>Backup</b>	Storage of digital data at climate-controlled off-site facility to support another archive's contingency plan per MOU. Data will not be disseminated by NSSDCA.
<b>Analog Archive</b>	Storage of digital data at climate-controlled off-site facility to support another archive's contingency plan per MOU. Data will not be disseminated by NSSDCA.

### 1.2. Archive Information Packages (AIPs)

In Table 1 NSSDCA's permanent archive is digital data that is stored either as AIPs or not. The non-AIP digital data is stored on off-line media and tracked by the media on which it resides. The portion of the data stored near-line in LTO jukeboxes has been growing since 2000 and includes all new data inflows received via electronic transfer, plus some legacy data collections; it is notable not because of its media, but because those data are stored on LTOs in Archive Information Packages (AIPs).

An AIP is a single file container that holds one or many science data files, a number of attributes about each file that help NSSDCA manage its AIPs, and pointers to all of the

supporting documentation, including calibration information. Ideally this is enough information to allow a user to be able to utilize the data independently of the archive and the original producer of the data. No reformatting of the science data files is performed unless record boundaries need to be retained and are not already in the byte stream. Any files that are transformed may be returned to their original state using the NSSDCA defined attributes. Additionally, AIPs are media independent and platform independent, making AIPs the mechanism for long-term preservation. NSSDCA plans to convert most of the non-AIP data in the permanent archive to AIPs.

### **1.3. Active Archives**

NASA has established a set of Active Archives, which receive data from missions and provide electronic access to the missions' data, along with documentation and tools for accessing and using the data. NSSDCA's mission is to accept data from the Active Archives or sometimes directly from missions, then provide long-term curation of the data. This is a critical service, since the full value of any set of data cannot be known in advance. New science discoveries or changes in research and exploration priorities may make older data, seldom requested, suddenly highly relevant.

## **2. Data Management Plan**

The revised, detailed Data management Plan for NSSDCA for 2016-2017 is given below (next page) in Tables 2a and 2b.

Table 2a lists data providers that will submit data to NSSDCA where the data will be archived under NSSDCA's Permanent Archive level of service, as defined by MOU for each data collection. For each provider the estimated data volume to be delivered and the provider's discipline (Planetary and Lunar, Astrophysics, or Heliophysics) is identified as well. Currently, only Planetary and Lunar data received from the PDS data nodes receive Permanent Archive level of service.

Table 2b conveys the same information as above for data providers that require NSSDCA's Backup Archive level of service. Both the Astrophysics and the Heliophysics disciplines are represented in this group of providers.

Table 2a identifies the data expected in 2016-2017 that will receive Permanent archival service. The volume of these data, all from PDS data nodes, is estimated at 259 TB. Table 2b identifies data, expected in 2016-2017, that will be stored according to NSSDCA's Backup service level. The volume of these data, from Astrophysics and Heliophysics providers, are estimated at 472 TB bringing the total volume of expected

The totals in Tables 2a and 2b show that NSSDCA is planning for ~361 TB of data arriving at the archive in 2016 and ~373 TB in 2017. 259 TB of those data will be archived under NSSDCA's Permanent Archive level of service and 474 TB will be stored under NSSDCA's Backup level of service.

**TABLE 2a**  
**Summary of Permanent Archive service level data expected at NSSDCA, 2016-2017.**

Permanent Archive Service Level			Expected Data Volume (TB)		Totals (TB)
			Discipline <sup>†</sup>		
			2016	2017	
<i>PDS Nodes</i>	ATMOS	P	6	7	13
	GEO	P	75	6	81
	IMAGING	P	65	65	130
	NAIF	P	0.55	0.75	1.3
	PPI	P	10	10	20
	RINGS	P	0.38	0.04	0.42
	SBN-PSI <sup>‡</sup>	P	4	5	9
	SBN-UMD	P	2	2	4
<b>Totals (TB)</b>			<b>162.93</b>	<b>95.79</b>	<b>258.72</b>

**TABLE 2b**  
**Summary of Backup service level data expected at NSSDCA, 2016-2017**

Backup Service Level			Expected Data Volume (TB)		Totals (TB)
			Discipline <sup>†</sup>		
			2016	2017	
<i>Missions</i>	FERMI	A	7	7	14
	RHESSI	H	.6	.6	1.2
	WISE	A	0	0	0
<i>Active Archives</i>	HEASARC	A	90	90	180
	IRSA	A	0	0	0
	MAST	A	0	0	0
	SPDF	H	100	180	280
<b>Totals (TB)</b>			<b>197.6</b>	<b>277.6</b>	<b>475.2</b>

<sup>†</sup>Discipline: A = Astrophysics; H = Heliophysics; P = Planetary & Lunar.

<sup>‡</sup>Subnode of SBN-UMD (PDS Small Bodies Node at the University of Maryland)

Table 3 summarizes the estimated data volumes by discipline. The heliophysics missions will make the largest contributions, closely followed by the PDS data nodes.

**TABLE 3**  
**Data expected at NSSDCA, 2016-2017, by discipline**

<b>Discipline</b>	<b>TB (2016-2017)</b>
Astrophysics	194
Heliophysics	281
Planetary & Lunar	258

## Glossary

AIP	Archive Information Package
GB	Gigabyte
HEASARC	High Energy Astrophysics Science Archive Research Center
MAST	Multi-mission Archive at Space Telescope Science Institute
NSSDC	National Space Science Data Center (now NSSDCA)
NSSDCA	NASA Space Science Data Coordinated Archive
PDS	Planetary Data System
PDS_ATM	PDS Atmospheres Node
PDS_GEO	PDS Geosciences Node
PDS_IMG	PDS Imaging Node
PDS_NAI	PDS Navigation and Ancillary Information Node
PDS_PPI	PDS Planetary Plasma Interactions Node
PDS_PSI	PDS Planetary Science Institute
PDS_RING	PDS Planetary Rings Node
PDS_SBN	PDS Small Bodies Node
RHESSI	Reuven Ramaty High Energy Solar Spectroscopic Imager
SPDF	Space Physics Data Facility
TB	Terabyte
WISE	Wide-field Infrared Survey Explorer