

**NATIONAL SPACE SCIENCE DATA CENTER  
ARCHIVE PLAN FOR 2007 – 2010**

Ed Grayzeck  
National Space Science Data Center  
Greenbelt, Maryland 20771  
2006-09-01

**ABSTRACT**

This archive plan shows that NSSDC presently expects to accept 36-127 TB/year of space science data into the archive over the years 2007-2010.

**1. INTRODUCTION**

NSSDC provides a vital service as NASA's only 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 archive plan summarizes the expected data inflow to NSSDC (note the Acronym list at the end of this document), by year and by missions, for the years 2007-2010, and is the successor to several earlier plans covering 3-4 years each.

**1.1 Levels of Service**

NSSDC 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 Second Archive and Backup services, mostly for other Archives. The Analog Archive includes photos, maps, microfilm, microfiche, documents, etc, some analog copies of digital data and others supporting metadata.

<b>Table 1. NSSDC Archival Storage Services</b>	
<b>Permanent Archive: AIPs</b>	Preservation of digital data in Archival Information Packages delivered by a data producer or created at NSSDC. AIPs are re-written to new media within six years. Data is disseminated by NSSDC 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. Data is disseminated by NSSDC if not available through an active archive or per MOU.
<b>Second Archive</b>	Storage of digital data on distributable media that is also held by another archive. No media refreshment is performed. NSSDC may disseminate the data if authorized to do so by the primary archive as 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 NSSDC.
<b>Analog Archive</b>	Preservation of analog data on a variety of media with selected refreshment and selected digitization. Selected retention of original analog data after digitization. Data are copied and disseminated by NSSDC.

**1.2 Archive Information Packages (AIPs)**

In Table 1 NSSDC'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 DLT 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 DLTs as AIPs.

An Archive Information Package (AIP) is a single file container that holds one or many science data files, a number of attributes about each file that help NSSDC 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 NSSDC defined attributes. Additionally, AIPs are media independent and platform independent.

AIPs are the preferred delivery and storage means. To that end, NSSDC makes available NSSDC packaging software and is encouraging Data Providers to use it to create AIPs and submit data in that form to NSSDC. In the long-term most of the non-AIP data in the permanent archive is planned to be converted 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. NSSDC'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.

## 2.0 ARCHIVE PLAN

The revised, detailed Archive Plan for NSSDC for 2007-2010 is given below (next page) in Table 3. Table 3 lists the missions, their launch dates and the estimated data volume to be delivered each year. Also included are the level of service (Permanent Archive - with or without AIPs, Second Archive, Backup) defined by MOU for each data collection and the discipline (Astrophysics, Heliophysics, Planetary & Lunar) for each. For archives which require Backup service, the data volumes expected from individual missions are combined and listed in the table by the name of the archive, i.e. HEASARC, IRSA, MAST, and LAMBDA.

The totals in GB for each year show an exponential growth, i.e. 36, 39, 118, 127 TB/yr for 2007-2010, respectively. The greatest increases are for 2009 and 2010 and are due to the Mars (MRO) and Lunar (LRO) Reconnaissance Orbiters. The summary of Table 3 by level of service and by discipline is given here in Table 2. Clearly, planetary missions dominate, contributing 271 TB to the NSSDC permanent archive.

<b>TABLE 2</b>			
<b>Service Level</b>	<b>TB (2007-2010)</b>	<b>Discipline</b>	<b>TB (2007-2010)</b>
Permanent Archive	257	Astrophysics	31
Second Archive	3	Heliophysics	16
Backup	31	Planetary & Lunar	271
Undetermined	26		

**TABLE 3.** Summary of data expected at NSSDC, 2007-2010. The large increases starting in 2009 are due primarily to MRO and LRO. SDO and SOHO will archive through SDAC, their data won't reach NSSDC until after 2010.

Project	Service Level* & Discipline+		Launch Date	Data Volume (GB)							Totals (GB)		TOTAL Planned 2007-10	
				2005		2006		2007	2008	2009	2010	2005-06		2005-06*
				Planned	Actual	Planned	Actual*	Planned	Planned	Planned	Planned	Planned		Actual*
ACE	A	H	1997 Aug	14	20	20	6	20	20	20		34	26	60
AIM	A	H	2006 Sep					2,000	2,000	2,000		0	0	6,000
AMPTE	A	H	1984 Aug									0	0	0
Cassini	A	P	1997 Oct	5,029		2,000		2,000	4,000	11,000		7029	0	17,000
CDAWeb	A	H			627		441					0	1068	0
Clementine	S	P	1994 Jan	56								56	0	0
Cluster	U	H	2000 Jul/Aug	360			28					360	28	0
CNOFS	A	H	2006 Mar					1,000	1,000	1,000		0	0	3,000
DE	A	H	1981 Aug		0.3							0	0	0
FAST	A	H	1996 Aug	666		2,000	109	2,000				2666	109	2,000
GALEX	A	A	2003 Apr				43					0	43	0
Galileo	S	P	1989 Oct	93								93	0	0
Genesis	A	P	2001 Aug		8							0	8	0
Geotail	A	H	1992 Jul	204	9	1	62	1	1	1		205	71	3
GP-B	A	A	2004 Apr					3,000				0	0	3,000
HEASARC	B	A		1,985	4,100	4,100	12,800	3,500	3,400	3,900	3,500	6085	16900	14,300
Hayabusa	S	P	2003 May									0	0	0
IBEX	A	H	2008 Jun					3	8	3		0	0	14
IMAGE	A	H	2000 Mar	320	44	150		150	150	150		470	44	450
IMP8	A	H	1973 Oct	2	0.6	76		2				78	1	2
INJUN 5	A	H	1968 Aug									0	0	0
IRSA	B	A						5,000	2,000	2,000	2,000	0	0	11,000
ISEE	A	H	1977 Oct		3		1					0	4	0
ISIS-Alouette	A	H	1965 Nov		200		22					0	222	0
ISO	B	A	1995 Nov									0	0	0
LAMBDA	B	A				351		512			810	351	0	1,322
LRO	A	P	2009 Fall							50,000	100,000	0	0	150,000
Lunar-A	U	P	2008?									0	0	0
Mariner 9	S	P	1971 May									0	0	0
Mariner 10	S	P	1973 Nov	8								8	0	0
MAST	B	A		4,590		700	217	4,000				5290	217	4,000
Mars Express	A	P	2001 Apr	4,400		4,000		6,000				8400	0	6,000
Mars Odyssey	S	P	2004 Jan	628			54	3,000				628	54	3,000
MER	A	P	2003 Jun/Jul	100				2,000				100	0	2,000
MESSENGER	S	P	2004 Aug	8								8	0	0
MGS	S	P	1996 Nov	1,222	44	500	236					1722	280	0
MRO	A	P	2005 Aug						25,000	34,000	8,000	0	0	67,000
MSC	U	A										0	0	0
NEAR	S	P	1996 Feb	8								8	0	0
New Horizons	S	P	2006 Jan									0	0	0
Nozomi	U	P	1998 Jul	1,200								1200	0	0
PDS-Other**	S	P			92		15					0	107	0
Phoenix	S	P										0	0	0
Polar	A	H	1996 Feb	200			193					200	193	0
RHESSI	A	H	2002 Feb	730	796	1,000	382	1,000	1,000	1,000		1730	1178	3,000
Rosetta	S	P	2004 Mar									0	0	0
S <sup>3</sup> -A	A	H	1971 Nov									0	0	0
SAMPEX	A	H	1992 Jul									0	0	0
SDO	U	H	2008 Apr									0	0	0
SELENE	U	P	2007 Jul							12,780	12,780	0	0	25,560
SMART 1	U	P	2003 Sep									0	0	0
SNOE	A	H	1998 Feb	0.20		0.10		0.00				0	0	0
SOHO	U	H	1995 Dec									0	0	0
Sp Phys-Other	A	H			92.00		0.04					0	92	0
STEREO	U	H	2006 May	19				21	210	210		19	0	441
THEMIS	A	H	2006 Oct					100	100	100	100	0	0	400
TIMED	A	H	2001 Dec	30		350		350				380	0	350
TWINS	A	H	2006									0	0	0
Ulysses	A	H	1990 Oct	9	1		3					9	4	0
Venus Express	U	P	2005 Nov									0	0	0
Viking	S	P	1975 Aug/Sep									0	0	0
Voyager	A	H	1977 Aug/Sep	226	0.04	0.04		0.04	0.04	0.04	0.04	226	0	0
Wind	A	H	1994 Nov	48	0.3	2.70	0.15	2.50				51	0	3
Yohkoh	A	A	1991 Aug									0	0	0
<b>Totals (GB)</b>				<b>22,154</b>	<b>6,037</b>	<b>15,251</b>	<b>14,612</b>	<b>35,662</b>	<b>38,889</b>	<b>118,164</b>	<b>127,190</b>	<b>37,405</b>	<b>20,649</b>	<b>319,905</b>

\* Service Level: A = Permanent Archive (with or without AIPs); B = Backup; S = Second Site; U = Undetermined  
 + Discipline: A = Astrophysics; H = Heliophysics; P = Planetary & Lunar

## Glossary

ACE Advanced Composition Explorer  
ADC Astronomical Data Center  
ASCA Advanced Satellite for Cosmology and Astrophysics  
ASCII American Standard Code for Information Interchange  
CD Compact Disk  
CD-R CD-Recordable  
CD-ROM CD-Read Only Memory  
CDAW Coordinated Data Analysis Workshop  
CDF Common Data Format  
CHIPS Cosmic Hot Interstellar Plasma Spectrometer  
COBE Cosmic Background Explorer  
COHO Coordinated Heliospheric Observations  
DLT Digital Linear Tape  
DR0 Data Release Zero  
DR1 Data Release One  
DR2 Data Release Two  
DS1 Deep Space 1  
DVD Digital Versatile Disk  
EDR Experiment Data Record  
FAST Fast Auroral Snapshot Explorer  
FTP File Transfer Protocol  
FUSE Far Ultraviolet Spectroscopic Explorer  
GALEX Galaxy Evolution Explorer  
GB Gigabyte  
GLAST Gamma-Ray Large Area Space Telescope  
HEASARC High Energy Astrophysics Science Archive Research Center  
HETE High Energy Transient Explorer  
IMAGE Imager for Magnetopause-to-Aurora Global Exploration  
IMP Interplanetary Monitoring Platform  
IPAC Infrared Processing and Analysis Center  
IRAS Infrared Astronomy Satellite  
IRSA Infrared Science Archive  
ISIS International Satellites for Ionospheric Studies  
ISO Imaging Spectrometric Observatory  
IU Independently Usable  
LAMBDA Legacy Archive for Microwave Background Data Analysis  
LPL Low Processing Level  
LRO Lunar Reconnaissance Orbiter  
MAP Microwave Anisotropy Probe  
MAST Multi-mission Archive at Space Telescope Science Institute  
MESSENGER Mercury Surface, Space Environment, Geochemistry and Ranging  
MRO Mars Reconnaissance Orbiter  
NDADS NSSDC Data Archive and Dissemination System  
NEAR Near Earth Asteroid Rendezvous  
NIU Not Independently Usable  
NSSDC National Space Science Data Center  
OSS Office of Space Science  
PDS Planetary Data Center  
RHESSI Reuven Ramaty High Energy Solar Spectroscopic Imager  
RXTE Roentgen X-ray Timing Explorer  
SAMPEX Solar Anomalous and Magnetospheric Particle Explorer  
SARC Science Archive Research Center  
SDAC Solar Data Analysis Center  
SDO Solar Dynamics Observatory  
SOHO Solar and Heliospheric Observatory