

**NATIONAL SPACE SCIENCE DATA CENTER
ARCHIVE PLAN FOR 2010 – 2013**

Ed Grayzeck
National Space Science Data Center
Greenbelt, Maryland 20771
2009-09-01

ABSTRACT

This archive plan shows that NSSDC presently expects to accept ~100+ TB/year of space science data into the archive over the years 2010-2013.

1. INTRODUCTION

NSSDC 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 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 2010-2013, and is the successor to several earlier plans covering 3-4 years each. Note these are estimates for planning purposes, not exact data projections.

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; it is included in this list for completeness.

Permanent Archive: AIPs	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.
Permanent Archive: non-AIP digital data	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.
Second Archive	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.
Backup	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.
Analog Archive	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, making AIPs the preferred delivery and storage means. 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 2010-2013 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, PDS, and SPDF.

The totals in TB for each year show roughly similar numbers from year to year, i.e. 136, 188, 161, 61 TB/yr for 2010-2013, respectively. The largest data sets expected are those for the Lunar (LRO) and Mars (MRO) Reconnaissance Orbiters. The summary of Table 3 by level of service and by discipline is given here in Table 2. Clearly, planetary missions dominate, their coming contribution to the NSSDC permanent archive is estimated to be 476 TB.

TABLE 2			
Service Level	TB (2010-2013)	Discipline	TB (2010-2013)
Permanent Archive	14	Astrophysics	34
Second Archive	476	Heliophysics	36
Backup	56	Planetary & Lunar	476
Undetermined	TBD		

TABLE 3. Summary of data expected at NSSDC, 2010-2013.

Project	Service Level* & Discipline+		Launch Date	Data Volume (GB)				Totals (GB)
				2010	2011	2012	2013	
ACE	A	H	1997 August	30	30	30	30	120
AIM	A	H	2006 September	6,000				6,000
Cassini	S	P	1997 October	1,000	1,000	1,000	1,000	4,000
Chandrayaan	S	P	2008 October	10,000	40,000	10,000		60,000
EPOXI	S	P	2005 January	110	90			200
FAST	A	H	1996 August	400	400	400		1,200
Fermi	A	A	2008 June	2,000	2,000	2,000		6,000
GALEX	B	A	2003 April	1,500	1,500	1,500		4,500
GPB	A	A	2004 April	10				10
GRAIL	S	P	2011 September				100	100
Hinode	U	H	2006 September	TBD				TBD
IBEX	A	H	2008 June	15	15	15	15	60
ISIS-Alouette 2	A	H	1965 November	50	50			100
Kaguya	S	P	2007 September	12,000				12,000
LADEE	S	P	2012 May			2		2
LCROSS	S	P	2009 June	4				4
LRO	S	P	2009 June	50,000	100,000	100,000	50,000	300,000
Mars Odyssey	S	P	2001 April	4,000				4,000
Mars Science Lab	S	P	2011 Fall			10	10	20
MER	S	P	2004 January	2,000				2,000
MESSENGER	S	P	2004 August		2,400	2,400		4,800
MRO	S	P	2005 August	30,000	25,000	34,000		89,000
NExT	S	P	1999 February			8		8
RHESSI	A	H	2002 February	600				600
SOHO	U	H	1995 December		TBD			TBD
THEMIS	A	H	2006 October	100				100
TWINS	A	H	2008 March	5	5	5	5	20
Ulysses	A	H	1990 October	1	1	1	1	4
Voyager	S	H	1977 August, Sept.	<1 GB				<1 GB
HEASARC	B	A		5,000	5,000	5,000	5,000	20,000
IRSA (WISE)	B	A		2,000				2,000
MAST (FUSE,others)	B	A		800	800			1,600
PDS (Other)	S	P		40	40	40	40	160
SPDF (CDAWeb)	B	H		8,000	10,000	5,000	5,000	28,000
Totals (GB)				135,665	188,331	161,411	61,201	546,608

*Service Levels: A = Permanent Archive (AIP or non-AIP); B = Backup; S = Second Site; U = Undetermined.

+Discipline: A = Astrophysics; H = Heliophysics; P = Planetary & Lunar.

Glossary

ACE	Advanced Composition Explorer
aka	also known as
AIM	Aeronomy of Ice in the Mesosphere
AIP	Archive Information Package
CDAWeb	Coordinated Data Analysis Web
EPOXI	Extrasolar Planet Observation and Characterization (EPOCh) plus Deep Impact Extended Investigation (DIXI)
FAST	Fast Auroral Snapshot Explorer
FUSE	Far Ultraviolet Spectroscopic Explorer
GALEX	Galaxy Evolution Explorer
GB	Gigabyte
GPB	Gravity Probe-B
GRAIL	Gravity Recovery and Interior Laboratory
HEASARC	High Energy Astrophysics Science Archive Research Center
Hinode	Sunrise, aka Solar-B
IBEX	Interstellar Boundary Explorer
IRSA	InfraRed Science Archive
ISIS	International Satellites for Ionospheric Studies
Kaguya	aka SELenological and ENgineering Explorer (SELENE)
LADEE	Lunar Atmosphere and Dust Environment Explorer
LCROSS	Lunar CRater Observation and Sensing Satellite
LRO	Lunar Reconnaissance Orbiter
MAST	Multi-mission Archive at Space Telescope Science Institute
MER	Mars Exploration Rover
MESSENGER	Mercury Surface, Space Environment, Geochemistry and Ranging
MRO	Mars Reconnaissance Orbiter
NExT	New Exploration of Tempel-1, follow-on to STARDUST
NSSDC	National Space Science Data Center
PDS	Planetary Data System
RHESSI	Reuven Ramaty High Energy Solar Spectroscopic Imager
SARC	Science Archive Research Center
SDAC	Solar Data Analysis Center
SOHO	Solar and Heliospheric Observatory
SPDF	Space Physics Data Facility
TB	Terabyte
THEMIS	Time History of Events & Macroscale Interactions during Substorms
TWINS	Two Wide-angle Imaging Neutral-atom Spectrometers
WISE	Wide-field Infrared Survey Explorer