

**NATIONAL SPACE SCIENCE DATA CENTER**

**ARCHIVE PLAN FOR 2003 - 2006**

Donald M. Sawyer  
National Space Science Data Center  
Greenbelt, Maryland 20771

2003-10-23

## Table of Contents

ABSTRACT.....	3
1. INTRODUCTION.....	3
2. ASTROHPYSICS.....	4
3. SPACE AND SOLAR PHYSICS.....	6
4. PLANETARY AND LUNAR SCIENCE.....	8
5. SUMMARY.....	9
Glossary.....	10

## **ABSTRACT**

This archive plan shows that NSSDC presently expects to bring 9-30 TB/year of space science data into its permanent archive over this and the next three years, and to make a significant fraction of space physics data network-accessible to the research community.

## **1. INTRODUCTION**

NSSDC serves as the permanent archive for most OSS mission data and it provides underlying support for the Sun-Earth Connection Active Archive (SECAA) within its Space Physics Data Facility partner at Goddard. This archive plan summarizes the expected data inflow to NSSDC (note the Glossary at the end of this Plan) for the years 2003-2006, and is the successor to several earlier plans covering 3-4 years each. The plan shows, by year and by mission, data volumes flowing to NSSDC as the permanent archive and the subset to be made network-accessible by NSSDC in support of the SECAA. Customers of NSSDC as the permanent archive are NASA's active archives. By agreement with the Planetary Data System, NSSDC makes PDS-provided data accessible to the general public via electronic and offline paths.

NSSDC's permanent archive environment and its externally accessible data dissemination environment are both in transition. The legacy permanent archive presently consists of: offline data volumes including 9-track tapes and 3480 tape cartridges holding older data; 4-mm and 8-mm tapes and CD-R disks holding more recently arrived data; CD-ROM disks wherein 2-3 copies are in the archive and 10's to 100's of copies are held for distribution; and non-digital media (film, etc.) almost always holding old data. NSSDC is gradually digitizing at least selected ones of its non-digital data sets. The new permanent archive, consisting of "nearline" DLT jukeboxes, has been growing since mid-2000. To date, most data moved to these jukeboxes are new data inflows or data previously user-accessible from NSSDC's 12"-WORM-optical-disk jukebox-based NDADS system.

NSSDC also provides electronic access to independently usable data from current and recent missions where the data are typically not electronically accessible from other sources. Two "levels" of electronic accessibility provided are (1) graphical-browse/subset/retrieve capabilities to CDF-formatted (Common Data Format) data via such systems as CDAWeb and OMNIWeb (ASCII output is a user option) and to data held as ASCII via "FTPBrowser" and ATMOWeb and (2) basic file-level ftp access to the CDF's underlying CDAWeb, etc., to the ASCII files also accessible through FTPBrowser and to various other mostly-ASCII data sets. A special client-server-based interface to IMAGE UDF files is also provided.

## 2. ASTROPHYSICS

NASA's astrophysics data environment is primarily characterized as a set of Science Archive Research Centers (SARC) which provide electronic access to NASA (and some non-NASA) mission data along with documentation about the data and tools for accessing and using the data. These include the High Energy Astrophysics SARC (HEASARC) at Goddard for gamma-ray and X-ray data and the Multi-mission Archive at Space Telescope Science Institute (MAST) for ultraviolet and optical data. Long wavelength data have their SARC functions done by Caltech/Infrared Processing and Analysis Center's (IPAC) Infrared Science Archive (IRSA), and the Legacy Archive for Microwave Background Data Analysis (LAMBDA) at Goddard. LAMBDA has a focus on MAP and COBE data and also provides access to IRAS and SWAS data. NSSDC operates solely as a permanent archive for astrophysics data. It also provides some derived products primarily of interest to the general public and education communities.

NSSDC's holding of data whose user access is provided by the ensemble of SARCs is for permanent archiving purposes only. Only the SARCs are viewed as NSSDC's customers for these data, while the SARCs in turn are expected to provide all needed user access and support. NSSDC still holds old legacy data not supported by any SARC yet. Such legacy data are currently held offline, but are community-visible through the web-accessible NSSDC Master Catalog and may be ordered from NSSDC. Discussions on the disposition of such data are being pursued between NSSDC and the relevant SARCs. The only astrophysics data arriving at NSSDC as a permanent archive are from the SARCs, or with a SARC as an intermediary between a spaceflight project and NSSDC.

**HEASARC** is expected to provide the following data volumes (GB):

	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>
<b>ASCA</b>	-	520	-	-
<b>Astro-E II</b>	-	-	100	300
<b>BeppoSAX</b>	50	-	-	-
<b>GLAST</b>	-	-	-	300
<b>HETE 2</b>	25	25	25	-
<b>Integral</b>	300	1000	1000	1000
<b>RXTE</b>	170	170	170	170
<b>Swift</b>	-	180	360	360
<b>XMM-Newton</b>	330	330	330	330
	-----	-----	-----	-----
<b>Totals</b>	875	2225	1985	2460

The 520 GB ASCA flow in 2004 represents a final reprocessing of all ASCA data. The 50 GB BeppoSAX flow in 2003 represents Wide Field Camera data for whose archiving HEASARC has only recently reached agreement with the project.

MAST is expected to provide the following data volumes (GB) or, in the case of GALEX, to facilitate the flow from the project to MAST and to NSSDC in parallel:

	2003	2004	2005	2006
<b>FUSE</b>	230	230	90	90
<b>Orfeus</b>	-	5	-	-
<b>CHIPS</b>	-	1	-	-
<b>GALEX</b>	100	500	4500	-
<b>Kepler</b>	-	-	-	tbd
	-----	-----	-----	-----
<b>Totals</b>	330	736	4590	90+

FUSE is funded through FY 2006 and is expected to generate data as long as it is operational. The FUSE project has begun to reprocess their earlier data, so the FUSE totals assume that half the data will be processed in 2003, and half in 2004. GALEX will deliver Data Release 0 (DR0) in late 2003. They will deliver DR1 in 2004 and the remainder of the data will be delivered in 2005 as DR2.

IRSA is expected to start the flow of 2MASS data to NSSDC in 2003. The expectation is that this data will consist primarily of catalogs and ancillary tables, and will total 53GB. IRSA may also make available approximately 410 GB of 2MASS Atlas image data.

LAMBDA will provide to NSSDC the 6th SWAS data release (approximately 4 GB) in 2003. It is believed there will be at most 2 more SWAS data releases before the project ends. LAMBDA expects a final release to be ready for permanent archiving in 2004. The first-year WMAP data release is approximately 20 GB. The final 4-year data set will be approximately 160-320 GB.

Summarizing these astrophysics data inflows to NSSDC, we have (in GB):

	2003	2004	2005	2006
<b>HEASARC</b>	875	2225	1985	2460
<b>MAST</b>	330	736	4590	90+
<b>IRSA</b>	53	tbd	tbd	tbd
<b>LAMBDA</b>	24	tbd	tbd	tbd
	-----	-----	-----	-----
<b>Totals</b>	1282	2961+	6575+	2550+

### 3. SPACE AND SOLAR PHYSICS

Space Physics missions are virtually all Principal Investigator/Co-Investigator mode missions wherein the following "data products" become available for archiving: low processing level (LPL) data not yet irreversibly transformed; independently usable and analyzable digital data, typically but not necessarily irreversibly transformed; browse-only data (GIF's, film products, etc.)

The LPL data typically arrive at NSSDC on media and are inadequately documented to make their use outside the framework of the data-providing project a realistic option. Such data are not made electronically accessible to the external community and are designated NIU (not independently usable) in the tables below. The latter two product types are typically made electronically accessible, except for data sets for which there is sufficient electronic access provided elsewhere. Electronic access may be provided through CDAWeb or other active archive services (e.g., OMNIWeb, COHOWeb, FTPBrowser) managed by the SEC Active Archive (SECAA) in the Goddard's Space Physics Data Facility or by simple ftp access managed jointly with SECAA.

The following data inflows (in GB) are expected from active missions:

MISSION	2003		2004		2005		2006		Notes
	NIU	IU	NIU	IU	NIU	IU	NIU	IU	
<b>ACE</b>	12	2	12	2	12	2	12	2	
<b>Cluster</b>	350	10	350	10	350	10	350	10	
<b>FAST</b>	660	6	660	6	660	6	660	6	
<b>GENESIS</b>	-	1	-	1	-	-	-	-	
<b>Geotail</b>	200	4	200	4	200	4	200	4	
<b>IMAGE</b>	120	200	120	200	120	200	120	200	(1)
<b>IMP 8</b>	-	76	-	76	-	2	-	2	(2)
<b>Polar</b>	250	40	250	40	170	30	1	1	(3)
<b>RHESSI</b>	1100	-	730	-	730	-	730	-	
<b>SAMPEX</b>	18	6	9	3	-	-	-	-	(4)
<b>SNOE</b>	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
<b>STEREO</b>	-	-	-	-	18	1	220	10	
<b>TIMED</b>	-	-	-	20	-	30	-	2000	
<b>Ulysses</b>	8	1	8	1	8	1	8	1	
<b>Voyager</b>	-	.02	-	.02	-	.02	-	.02	
<b>Wind</b>	40	8	40	8	40	8	40	8	
<b>TOTAL</b>	2758	354	2379	371	2308	294	2341	2244	

Notes:

1. IMAGE IU (independently usable) data are in 3 data sets: 100 GB of CDF'd data to CDAWeb, 50 GB of tarred and zipped UDF files to DIONAS/nssdcftp, and 50 GB extracted from the nssdcftp UDF's to "IMAGE Archive" on lewes.
2. Large 2003 and 2004 submissions are for a 0.32-s magnetic field data set.
3. Polar will cease operations in 2005 but some residual data expected in 2006

4. SAMPEX will cease operations in 2004.

In addition to inflows from these active missions, NSSDC expects to continue receiving 60 GB/year of data from the ISIS ionogram digitization project at Goddard through 2003. The only other likely voluminous inflow is from the Atlas 1 Imaging Spectrometric Observatory (ISO), at 20 GB for 2004. Folding these in, a net summary for space physics becomes:

	2003		2004		2005		2006	
	NIU	IU	NIU	IU	NIU	IU	NIU	IU
<b>Active Missions</b>	2758	354	2379	371	2308	294	2341	2244
<b>ISIS</b>	-	60	-	-	-	-	-	-
<b>ISO</b>	-	-	-	20	-	-	-	-
<b>Totals</b>	-----	----	-----	----	-----	----	-----	-----
	2758	414	2379	391	2308	294	2341	2244

The Solar Data Analysis Center (SDAC) at Goddard functions much like an astrophysics SARC for the solar remote sensing data from current and recent solar missions (except that the voluminous helioseismology data are managed at, and made electronically accessible from, Stanford U.). Key current missions supported at SDAC are SOHO and TRACE. Data from these are managed by SDAC in both an online-accessible mode and an offsite media-storage mode. Copies will be provided to NSSDC after conclusion of those missions.

## 4. PLANETARY AND LUNAR SCIENCE

NSSDC's main source of these data is the multi-nodal Planetary Data System (PDS). PDS "publishes" data sets in that they accumulate data and needed supporting material onto media that are peer reviewed for adherence to PDS standards and for their correct and independent usability. These data arrive at NSSDC on CD-ROM or on DVD and are not made electronically accessible by NSSDC because PDS provides such access. (NSSDC provides electronic access to selected popular images as part of its support of the general public; this will not be further described herein.)

We expect to receive the following numbers of physical data volumes from PDS in the coming years:

MISSION	2003		2004		2005		2006	
	CD	DVD	CD	DVD	CD	DVD	CD	DVD
Mars Global Surveyor	1540	47	784	184	784	184	784	84
Mars Odyssey	99	523	96	543	96	142	96	142
Mars Recon. Orbiter	-	-	-	-	-	-	60	2702
Mars Express	-	-	-	1400	-	1100	-	750
Mars Explorer Rover	-	10	-	27	-	25	-	-
Nozomi	-	-	-	-	-	300	-	300
Viking Lander	2	5	-	-	-	-	-	-
Galileo	10	-	300	-	150	-	-	-
Cassini	-	15	112	451	111	1240	111	1540
MESSENGER	10	-	10	2	-	2	-	2
Clementine	-	15	-	15	-	14	-	14
Mariner 10	-	-	-	3	-	2	-	2
Voyager	832	-	545	-	364	-	60	-
Dawn	-	-	-	-	-	-	-	2
Deep Impact	-	15	-	11	-	12	-	11
DS1	-	1	-	3	-	2	-	-
Muses-C	-	-	-	-	-	1	-	-
NEAR	-	-	-	2	-	2	-	2
Rosetta	-	-	-	2	-	2	-	3
Stardust	-	2	-	4	-	4	-	1
Huygens	10	-	-	-	-	-	-	-
Mariner 9	-	2	-	-	-	-	-	-
Other	521	70	71	361	71	356	74	506
<b>TOTAL</b>	<b>3024</b>	<b>705</b>	<b>1918</b>	<b>3008</b>	<b>1576</b>	<b>3386</b>	<b>1185</b>	<b>5993</b>

If we assume that each CD has 620 MB and each DVD has 4 GB, then in the years 2003-2006, we expect inflows as follows:

	2003		2004		2005		2006	
	CD	DVD	CD	DVD	CD	DVD	CD	DVD
<b>Total (TB)</b>	1.87	2.82	1.19	12.0	0.98	13.5	0.73	24

## 5. SUMMARY

The following shows the total data volumes (GB) expected to flow to NSSDC in the coming years and the subsets of these volumes that NSSDC/SECAA expects to make network-accessible to the research community. Note that NSSDC's magnetic disk needs are a combination of these new data plus disk space to make still-important but still-offline data network-accessible.

DISCIPLINE	2003		2004		2005		2006	
	TOTAL	DISK	TOTAL	DISK	TOTAL	DISK	TOTAL	DISK
<b>Astrophysics</b>	1282	0	2961+	0	6575+	0	2550+	0
<b>Space Physics</b>	3172	414	2770	391	2602	294	4585	2244
<b>Planetary</b>	4690	0	13190	0	14480	0	24730	0
	-----	-----	-----	-----	-----	-----	-----	-----
<b>Total (GB)</b>	9,144	414	18,921+	391	23,654	294	31,866	2244
<b>Total (TB)</b>	9.1	0.41	18.9+	0.39	23.7+	0.30	31.9+	2.2

## 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
MAP	Microwave Anisotropy Probe
MAST	Multi-mission Archive at Space Telescope Science Institute
MESSENGER	Mercury Surface, Space Environment, Geochemistry and Ranging
NDADS	NSSDC Data Archive and Dissemination System
NEAR	Near Earth Asteroid Rendezvous
NIU	Not Independently Usable
NSSDC	National Space Science Data Center
OMNI	(not an acronym)
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

SEC	Sun Earth Connection
SECAA	Sun-Earth Connection Active Archive
SOHO	Solar and Heliospheric Observatory
SNOE	Student Nitric Oxide Explorer
STEREO	Solar Terrestrial Relations Observatory
SWAS	Submillimeter Wave Astronomy Satellite
TB	Terabyte
TIMED	Thermosphere Ionosphere Mesosphere Energetics and Dynamics
TRACE	Transition Region and Coronal Explorer
2MASS	2 Micron All Sky Survey
UDF	Universal Data Format
WORM	Write Once Read Many
XMM	X-ray Mapping Mission