

2011 ANNUAL STATISTICS AND HIGHLIGHTS
for
THE NATIONAL SPACE SCIENCE DATA CENTER

Edwin J. Grayzeck

National Space Science Data Center

Greenbelt, Maryland 20771

2012-Mar-01

Table of Contents

PREFACE

1. INTRODUCTION

2. HIGHLIGHTS

3. DATA MANAGED AT NSSDC

3.1. Data Inflow for 2011

3.2. Data Outflow for 2011

4. ADDITIONAL NSSDC SERVICES

4.1. NSSDC Information Management System (NIMS)

4.2. SPASE and Virtual Observatories (VOs)

4.3. Consultative Committee for Space Science Data Systems (CCSDS)

4.4. Sun-Earth Day (SED)

Glossary

PREFACE

The National Space Science Data Center (NSSDC) serves as the permanent archive for NASA's Space Science community. A major component of its mission is to ensure future data accessibility and usability. NSSDC also provides current data access, complementary to the efforts of other NASA "active archives," in support of the NASA and international research enterprises. Finally, NSSDC is a conduit for the general public and education community to acquire NASA space science data that may interest them.

Herein we report on the activities of the NSSDC for the calendar year 2011. As much as possible, we report the same statistics as in previous years to enable interested parties who wish to compare accomplishments year-to-year. Nevertheless, as NSSDC evolves, some statistical tables have been updated to better reflect current operations. Note specifically that many statistics in this report are only comparable to those from the 2005 and later, since our NIMS central database was revamped in 2005.

NSSDC is pleased to issue this 2011 Annual Report describing the 2011 growth and evolution of NSSDC's data archives, access pathways, and other tools and services, as well as the access to those data and services by NSSDC's customer communities. This report has been made WWW-accessible in the hope that readers will avail themselves of the opportunity to link to the services reported herein.

I welcome suggestions from users for improvements to this Annual Report and to NSSDC services.

Edwin J. Grayzeck

Head, National Space Science Data Center

1. INTRODUCTION

This report characterizes NSSDC's data holdings, metadata holdings, access pathways, and value-added data products, tools, and services at the end of 2011, with a focus on the 2011 activities leading to that end-of-year state. In addition this report characterizes the nature and access to NSSDC's data and services by its many users from various communities. It is assumed the reader will have a general familiarity with NSSDC and its mission. For more information see the top level NSSDC web page at <http://nssdc.gsfc.nasa.gov/>.

In 2005 we assembled an external user group, the NSSDC User Group (NUG), which meets periodically to provide findings regarding NSSDC goals. The group's reports, most recently from March 2009, are on-line and can be linked to from the NSSDC homepage. The next NUG meeting will be planned for 2012.

2. HIGHLIGHTS

The center of this report is the 13 Tables which summarize NSSDC activities in 2011. In most cases these numbers speak for themselves, though it is irresistible to address a few highlights.

The most important result of NSSDC's 2011 continuing activities is the preservation of growing space science data volumes, ensuring their continuing and future accessibility to the space science, education and general public communities. NSSDC's archive has now grown to 130 TB of space science data and an additional 1.6 TB of Earth science data. During 2011, 23 TB of data were added to the NSSDC archive; the largest components of that were in our holdings for Fermi, HEASARC and PDS.

In 2001 NSSDC adopted the Archival Information Package (AIP) as its preferred mechanism for the preservation of digital data. Since then the NSSDC AIP has evolved to a form where data content is stored in a POSIX-compliant tar file and associated preservation description information is maintained in an XML object. In 2011 staff provided software and interfaces to migrate data from legacy magnetic tapes into the current form AIP, permitting the original AIP generation software to be retired. They also developed software to convert older format AIPs to the current format and provided software for reading legacy tapes and reading "images" of legacy tapes provided by a contracted data recovery company. The database used for managing the digital data ingest process was improved, with the augmentation of process history tracking and additional checks to ensure relational integrity. New software automated much of the process for ingesting data from Planetary Data Systems. Also, NSSDC upgraded to LTO-5, the most recent generation of Linear Tape-Open technology, for AIP storage.

3. DATA MANAGED AT NSSDC

There are several ways to characterize the multi-disciplinary NSSDC archive and we use most of them herein to give a true sense of the NSSDC, i.e. byte counts, numbers of distinct datasets, and numbers of media volumes managed, as well as the diversity of datasets and of media types. For the remainder of this section we will present this variety of statistics in tables, similar in format

to prior years' reports, though recognizing that the content of some tables will not be comparable to those for years earlier than 2005, when our databases were revamped.

Table 1. Counts of NSSDC Datasets on December 31, 2011

Discipline	Digital	Analog	Total
Astrophysics	229	76	305
Space/Solar Physics	1,319	665	1,984
Planetary Science	1,353	761	2,114
Earth Science	107	129	236
Other	118	440	558
Total	3,126	2,071	5,197

By the end of 2011 NSSDC was managing 5,197 distinct datasets and accompanying documentation packages. Table 1 indicates the disciplines from which these datasets come and whether the datasets are digital or analog. Space Physics had been the dominant discipline for digital datasets, counting nearly half of NSSDC's holdings, reflecting NASA's early years with a preponderance of space physics missions whose spacecraft typically carry more independent experiments than do other missions. In recent years Planetary has approached parity and this year surpassed the Space Physics counts with 237 new datasets. There were only minor changes in other totals for 2011, each plus/minus one or two datasets. One apparent major addition was a bookkeeping update, correcting entries overlooked in a database transition a few years ago, with the effect of increasing the number digital datasets in Space/Solar Physics by more than 100.

NSSDC manages almost as many analog (e.g. film, microfilm and microfiche) datasets as digital datasets, though in recent years newly arriving data has been all digital. NSSDC has generated digital versions for some of its film archive, often in response to requests.

Table 2 is a different characterization of the NSSDC archive, showing byte counts for the entire digital archive. Some of the byte counts, particularly for older media, are estimates involving assumptions about the mean numbers of bytes on various media types for some datasets. The totals in Table 2 are also adjusted for any data removed from NSSDC, e.g. the 1.65 TB of Earth Science data which were transferred to a NASA Earth Science archive. The greatest change in a discipline was Planetary, whose holdings increased by more than 50%, adding 6 TB in 2011.

Table 2. State of the NSSDC Archive December 31, 2011

All Digital Data (TB)	
Astrophysics	64.67
Space Physics	29.55
Planetary	15.50
Earth Science	1.63
Other (includes Legacy)	23.31
Total	134.66

Data are also being moved from NSSDC's traditional offline archive to a near line archive based on DLT, SDLT, and LTO-4 jukeboxes attached to unix and linux servers, respectively. Data are

newly archived in Archive Information Packages (AIPs), which hold data files and companion attribute files and are media-independent and platform-independent. These are defined as per the AIP concept of the ISO/CCSDS Open Archival Information System reference model. Table 3 shows the volumes of data ingested to this portion of the archive for 2009-11; the data stored as AIPs has reached a total of 19 TB.

Table 3. Data Ingested to Nearline Permanent Archive

	2009		2010		2011	
	AIPs	GB	AIPs	GB	AIPs	GB
ALOUETTE	16281	9.43				
APOLLO			2	0.02		
GP-B			532	82.41		
ISIS	22168	10.47				
LEGACY DATA	1073	14.97	1282	50.51	2404	55.43
PDS DATA	171	247.66	2180	2492.14	2400	5976.04
RHESSI	7233	577.52	6178	424.74	10988	842.81
SPDF			2	0.06		
TWINS					2566	634.49
WIND	367	0.30	275	0.22	397	0.23
TOTALS	47,293	860.35	10,451	3050.09	18,755	7509.01

In Table 3 “Legacy Data” refers to data currently archived on aging magnetic tapes, generally 9-track tapes or 3480 cartridges. These are being read and ingested as AIPs in an on-going effort.

Table 4. Space Science Data FTP Accessible from NSSDC on December 31, 2011

Top Level Dir	GB 2011	Top Level Dir	GB 2011	Top Level Dir	GB 2011
ACE	55.694	GEOTAIL	2.945	OGO	3.774
AE	1.279	HELIOS	2.689	OMNI	29.971
AMPTE	5.014	IBEX	9.067	PIONEER	5.461
APOLLO	14.842	IMAGE	260.995	SAMPEX	54.070
CRRES	34.061	IMP	116.285	SOLAR_MAX	20.892
DE	186.642	ISEE	36.400	ULYSSES	297.339
DIRBE	19.963	ISIS	207.126	VOYAGER	30.987
EXPLORER	4.761	MAGSAT	1.869	WIND	58.237
				Others*	9.179
		TOTAL	1469.5428		

* Others are spacecraft with <1Gb data each, including Aeros, Alouette, ARCAD, Cassini, DMR, Galileo, Genesis, Hinotori, Mariner 10, Prognoz 6,7, and 9, San Marco, SNOE, Viking and miscellaneous additional spacecraft.

About half of the data stored in AIPs are made network-accessible on nssdcftp for the convenience of some of the user community. Table 4 lists NSSDC's network-accessible Space Physics data as of 31 December 2011 for projects with > 1GB. Spacecraft with < 1GB are summed as part of “Other” at the end of list. It must be noted that nssdcftp has a continually evolving directory structure, so Table 4 represents only a snapshot.

Table 5 characterizes the digital media types managed at NSSDC, not including backup copies. It should be noted that most volumes are replicable and have one backup volume. The transfer of the NSSDC Earth Science tapes to NASA's Earth Science archive is continuing, thus the 2011 summary shows a further reduction of over 4800 tapes, mostly 9-track and 3480 cartridges, in the Earth Science discipline.

Table 5. Counts of Volumes* at NSSDC Archive on Dec 31, 2011

	Astro- physics	Space Physics	Planetary Science	Earth Science	TOTAL
4-mm Tape	721	94	3	0	818
8-mm Tape	205	199	77	0	481
9-Track Tape	531	4,794	1,679	6,660	13,664
3480 Cartridges	474	1,984	1,157	195	3,810
DLT	84	114	2	0	200
LTO-4	86	0	0	0	86
CD	855	23,182	5,346	9	29,392
DVD	1,016	1,432	255	0	2,703
12-in Worm	0	4	0	0	4
M-O Disk	274	0	0	0	274
TOTAL	4,246	31,803	8,519	6,864	51,432

* Backup volumes and those not attributable to these 4 disciplines are not included.

Table 6 lists NSSDC's analog archive holdings by disciplines and by form factor. Film (feet) are reels of uncut film containing sequences of still photos while Movies are counted by reels. No backup copies of images are counted here. This table is unchanged from 2008, when we completed an inventory of the microfilm and microfiche portions of the analog archive, and is not comparable to versions from earlier years. The continuing inventories of other categories -- e.g. sorting through duplicate sets of media -- have been suspended due to budgetary considerations. We expect no additions to Table 6 in the foreseeable future. However, in 2011 we did discuss with NARA possibly transferring some portions of the Analog Archive to them; these discussions continued into 2012.

Table 6. Analog Data Products at NSSDC by Discipline

Discipline	Microfilm (reels)	Microfiche (cards)	Film (ft)	Film (frames)	Movie (reels)	Maps	Slides
Astrophysics	2357	5963	100	11975	0	0	62
Earth Science	55	773	13067	64610	4	0	0
Planetary Science	2188	13492	409081	335053	181	1773	768
Space Physics	34589	30507	580	6173	2	0	36893
Other	3224	3624	3785	6206	20	0	2112
Totals	42,413	54,359	426,613	424,017	207	1773	39,835

3.1 Data Inflow for 2011

Table 7 reports arriving media types by discipline. The numbers are net, i.e. some providers redeliver data on other media and we return/subtract their originals, so as not to double count. In recent years the dominant media type received has moved from CDs to DVDs. However, with their larger capacity, the much smaller number of LTO-4 tapes delivered a larger volume

Table 7. Media Arriving at NSSDC in 2011

	Astrophysics	Space Physics	Total
CDs	1	0	1
DVDs	1	79	80
LTO-4	2	0	2
Totals	4	79	83

Table 8. Data Arriving at NSSDC During 2011

	GB	Sub-Totals by Discipline
Astrophysics		
FERMI	6400.01	17,681.84
HEASARC	11281.83	
Planetary		
PDS_ATM	926.89	6,022.85
PDS_GEO	1040.24	
PDS_IMG	276.05	
PDS_PPI	3613.29	
PDS_PSI	8.27	
PDS_SBN	158.10	
Space Physics		
RHESSI	846.96	1,142.38
SPDF	44.99	
TWINS	250.19	
Wind	0.23	
Miscellaneous		
Legacy Data	55.43	55.43
TOTAL		24,902.50

of data than the discs. We receive additional data each year via electronic delivery or on data bricks, i.e. portable hard drives, which are just transfer media and so not included in Table 7.

Table 8 shows by project the data volumes that NSSDC received in 2011, totaling nearly 25 TB of new data via a combination of electronic deliveries and media. Dominating the statistics are data from FERMI and HEASARC in Astrophysics and PDS (reported by node) in Planetary. The actual HEASARC delivery was much larger, but we report a net value; by the terms of the MOU datasets delivered earlier are removed. We note PDS delivered more than double the data volume of 2010 (6.0 vs 2.5 TB).

3.2 Data Outflow for 2011

NSSDC provides access to its data holdings with network-accessible data for chosen datasets and, in addition, through a user support infrastructure for the mailing of offline digital and analog data volumes. Most electronic interfaces are accessible through NSSDC's WWW home page and include special WWW-based interfaces to specific datasets or groups thereof and ftp pathways to a range of

data files maintained permanently on NSSDC disks at nssdcftp.

Much of the data outflow discussed in NSSDC Annual Reports before 2003 was activity within SPDF, which maintains the Active Archive for NASA Space Physics missions; the activities of CDAWeb, etc, now are covered in SPDF reports elsewhere. Because NSSDC and SPDF have been, and are still, co-located at NASA's Goddard Space Flight Center, nssdcftp has remained a shared resource for now and that is reflected in the reported statistics.

Table 9 gives the annual counts of files downloaded from nssdcftp for 2011 and four prior years and singles out those with high activity, the Photo Gallery and Spacecraft Data subdirectories. The Photo Gallery downloads continued to decline, probably because of the proliferation of websites with space photos. Researchers downloading data files via ftp from the Spacecraft Data subdirectory had a notable increase from recent years, showing the high interest in and great value of services provided by NSSDC and SPDF on this shared resource.

Table 9. Number of Files Downloaded via FTP					
	2007	2008	2009	2010	2011
Photo Gallery	720,213	481,089	193,577	152,533	129,161
Spacecraft Data	689,961	1,856,362	1,564,930	2,314,095	2,436,245
All others on nssdcftp	930,342	634,070	615,242	789,413	468,459
TOTAL	2,340,516	2,971,521	2,373,749	3,256,041	3,033,865

NSSDC responded to 35 distinct requests for ftp data and for data products and Table 10 characterizes the user community of requestors. The requests are from the U.S. and the international researchers, the general public, educational enterprises, publishers, etc.

Table 10. NSSDC User Community (Offline Requests Only) for CY 2011

Affiliation Category	Total Requests
No Affiliation [General Public]	5
Non_US	7
US Academic Institutions	6
US Private Industry	2
NASA/GSFC	12
NASA Centers, excluding GSFC	3
Total	35

Table 11 gives the counts of requests for offline datasets from various disciplines in 2011. It again shows the dominance of planetary data requests, largely associated with lunar and planetary image data that are widely requested by the general public. The number of requests is different than in Table 10 because some requests are for data/items related to more than one

discipline, so are double counted in Table 10; some requests are for documents or other form of information that are not counted as a datasets, thus not included in Table 11.

Table 11. Number of Requests for Offline Data by Discipline

DISCIPLINE	Data Set Requests 2011
Astrophysics	6
Earth Science	2
Planetary Science	11
Space & Solar Physics	9
TOTAL	28

In 2011 data distribution via ftp was preferred due to its immediacy and to NSSDC budget limitations, though some media items were distributed. Table 12a below shows the distribution of data served within these categories for 2011 and the previous two years; Table 12b gives the distribution of the items by discipline for the same years. Both Tables 12a and 12b show ftp as the dominant mode of distribution. Note that the statistics in Table 12 include only ftp data that were newly posted in response to a request. Table 12 does not include any repeat requests for data previously made available via ftp.

Tables 12a,b. NSSDC Offline Data Dissemination Statistics 2007-2011

Table 12a					
ITEMS	2007	2008	2009	2010	2011
DISCs	441	186	196	209	205
PRINTED	177	109	36	21	0
OTHER	9	4	11	2	4
FTP	20,907	20,877	20,719	7,059	32,326
TOTAL	21,534	21,176	20,962	7,291	32,535
Table 12b					
DISCIPLINE	2007	2008	2009	2010	2011
Astrophysics	277	95	10,737	27	124
Planetary	2,311	20,979	740	932	31,891
Space Physics	18,946	102	9,470	6,319	57
Other	0	0	15	13	463
TOTAL	21,534	21,176	20,962	7,291	32,535

NOTES:

DISCs include CDs & DVDs

PRINTED materials include Photos, Posters, Maps, Documents

OTHER media include Microfilm, Microfiche, Tapes, Videos, Reels

FTP include Data, Documents, & Photos posted for FTP download, not files already posted

4. ADDITIONAL NSSDC SERVICES

In addition to its archive of scientific data and the variety of data interfaces characterized in the preceding sections, NSSDC offers a number of additional services, which are described below.

4.1 NSSDC Information Management System (NIMS)

The NSSDC Information Management System (NIMS) encompasses most of the separate databases that NSSDC has used to track data and information through the years. The NSSDC has a long term goal of incorporating its off-line data inventory system into NIMS and a major effort for this is underway.

Table 13. NIMS/JEDS Database Statistics for CY 2011

Subpartition	Number of Records as of 12/31/2011	Number Added in 2011
Spacecraft	6794	122
Experiment	5428	15
Dataset	6014	155
TOTAL	18,236	292

Number of spacecraft with experiment records - 1,085

Number of experiments with datasets at NSSDC - 1,640

Additional datasets associated only with spacecraft, not experiments - 749

Additional datasets that are not associated with spacecraft/experiment - 476

Table 13 details the pertinent NIMS database statistics for 2011. NIMS identifies virtually all launched spacecraft, the experiments carried by many of these spacecraft, and datasets from these spacecraft primarily as archived at NSSDC. This portion of the database is the source of information for many of NSSDC's WWW information pages. The NSSDC Master Catalog (NMC) dynamically generates WWW pages so that the latest information is presented to the user. A number of discipline and project pages are based on information derived from NIMS or utilize the NMC to generate such information.

4.2 SPASE and Virtual Observatories (VOs)

With over 30 years experience in managing and preserving digital information comprising thousands of datasets, NSSDC is acutely aware of the need to acquire and preserve data along with adequate documentation to ensure they are independently understandable and usable for current and future researchers. In this era of increasingly seamless access to archived data, the NSSDC mission also includes support for the Virtual Observatories, particularly as part of the development of the Space Physics Archive Search & Exchange (SPASE), the dictionary which will be the common language among space physics archives.

Version (2.2.2) of the SPASE Data Model was released on 26 September 2011. More information can be found at <http://www.spase-group.org/> which also has a link from NSSDC through its VO Portal.

4.3 Consultative Committee for Space Science Data Systems (CCSDS)

The NSSDC continues to lead the Data Archive Ingest (DAI) Working Group (WG) within the Consultative Committee for Space Data Systems (CCSDS) and also provides the deputy chair of the Repository Audit and Certification (RAC) WG. For this report we provide a summary of key projects in progress including:

- *The Producer-Archive Ingest Specification*: the revised version was prepared after some iterations and has been provided to the CCSDS Technical Editor for agency review.
- *Reference Model for Open Archival Information System*: a draft update to the standard was provided to the CCSDS Technical Editor as an updated CCSDS and ISO Standard.
- *Audit and Certification of Trustworthy Digital Repositories*: has become a CCSDS Standard, passed international review and the ISO standard proof is complete.
- *Guidelines for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories*: is now a CCSDS standard, passed international review, has been approved and the proof is in progress for publication as an ISO standard.

NSSDC agreed to participate as a subject in a test audit using the aforementioned *Guidelines*. The audit took place in June 2011 and preliminary results were presented at the closing session. There were no surprises in the final report, which was issued in early 2012 and has been distributed internally. As preparation for the test audit, NSSDC had updated documentation and process plans; the auditors encouraged the data center to follow up and do more of the same. NSSDC will do so as time allows.

The reader is referred to <http://www.ccsds.org/> for specifics on CCSDS activities.

4.4 Sun-Earth Day

The NASA Sun-Earth Day team, with major NSSDC participation, prepared for and orchestrated Sun-Earth Day 2011 with the theme “Ancient Mysteries, Future Discoveries” for the main event on March 19. Many thousands of packets of information were sent to teachers, scientists and others for Sun-Earth Day programs, reaching hundreds of thousands of people with live webcasts and podcasts. Throughout the year the team also sponsored a number of workshops and teacher professional development events reaching thousands of teachers, amateur astronomers, and the general public in partnership with Heliophysics missions, museums, science centers, and planetariums, as well as science and educational professional societies. The same team of scientists and educators has since been preparing to support Sun-Earth Day for 2012, a double observance united under the theme “Shadows of the Sun.” The annual equinox celebration will be March 19, 2012; the Transit of Venus will be the highlight on June 5, 2012.

For more information on Sun-Earth Day, the reader is referred to <http://sunearthday.nasa.gov/> .

Glossary

ACE	Advanced Composition Explorer
AE	Atmospheric Explorer
AIP	Archive Information Package
ARCAD	Arc Aurorale et Densite
CCSDS	Consultative Committee for Space Data Systems
CDAWeb	Coordinated Data Analysis Web
CRRES	Chemical Release and Radiation Effects Satellite
DE	Dynamics Explorer
DIRBE	Diffuse Infrared Background Experiment
DLT	Digital Linear Tape
DMR	Differential Microwave Radiometers
DVD	Digital Versatile Disk (originally, V = video)
FTP	File Transfer Protocol
GB	Gigabyte
GP-B	Gravity Probe-B
GSFC	Goddard Space Flight Center
HEASARC	High Energy Astrophysics Science Archive Research Center
IBEX	Interstellar Boundary Explorer
IMAGE	Imager for Magnetopause-to-Aurora Global Exploration
IMP	Interplanetary Monitoring Platform
ISEE	International Sun-Earth Explorer
ISIS	International Satellite for Ionosphere Studies
ISO	International Organization for Standardization
JEDS	Java Experiments, Datasets, Spacecraft
LTO-4	Linear Tape-Open [i.e. open standard], 4 th generation
MAGSAT	MAGnetic field SATellite
M-O	Magneto-optic
NARA	National Archives and Records Administration
NIMS	NSSDC Information Management System
NMC	NSSDC Master Catalog
NSSDC	National Space Science Data Center
NUG	NSSDC User Group
OAIS	Open Archival Information System
OGO	Orbiting Geophysical Observatories
OMNI	Interplanetary Medium Data (not an acronym)
PAIMAS	Producer-Archive Ingest Methodology Abstract Standard
PDS	Planetary Data System
RHESSI	Reuven Ramaty High Energy Solar Spectroscopic Imager
SAMPEX	Solar Anomalous and Magnetospheric Particle Explorer
SDLT	Super DLT (see above)
SNOE	Student Nitrogen Oxide Explorer
SPASE	Space Physics Archive Search & Exchange
SPDF	Space Physics Data Facility
TB	Terabyte
TWINS	Two Wide-angle Imaging Neutral-atom Spectrometers
VO	Virtual Observatory
WORM	Write-Once, Read-Many