

2012 ANNUAL STATISTICS AND HIGHLIGHTS
for
THE NATIONAL SPACE SCIENCE DATA CENTER

Edwin J. Grayzeck

National Space Science Data Center

Greenbelt, Maryland 20771

2013-Mar-01

Table of Contents

PREFACE

1. INTRODUCTION

2. HIGHLIGHTS

3. DATA MANAGED AT NSSDC

3.1. Data Inflow for 2012

3.2. Data Outflow for 2012

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 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 be of interest them.

Herein we report on the activities of the NSSDC for the calendar year 2012. 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. These are noted in the text.

NSSDC is pleased to issue this 2012 Annual Report describing the growth and evolution of NSSDC's data archives 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 web 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 services at the end of 2012, with a focus on the 2012 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/> .

2. HIGHLIGHTS

The center of this report is the 13 Tables which summarize NSSDC activities in 2012. In most cases these numbers speak for themselves, though here we address a few highlights.

The most important result of NSSDC's 2012 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 152 TB of space science data, even after we transferred the last 1.6 TB of NSSDC's Earth Science digital data to the Earth Sciences Data Information and Services Center (DISC) during 2012. Over 20 TB of data were added to the NSSDC archive; the largest contributors were data from Astrophysics missions, HEASARC and the Planetary Data System (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 2012 NSSDC replaced the database used to manage the migration of legacy tape data into AIPs. "Legacy Data" refers to data on aging magnetic tapes, such as 9-tracks or 3480 cartridges, which are being read and ingested as AIPs in an on-going effort. The new database provides a comprehensive inventory of legacy media, captures metadata essential to long-term preservation of the legacy data, and tracks data reclaimed from legacy media by a contracted data recovery services provider.

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 data collections, and numbers of media volumes managed, as well as the diversity of data collections 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 as much as possible.

By the end of 2012 NSSDC was managing 5,261 distinct data collections and accompanying documentation packages. Table 1 indicates the disciplines from which these data collections come and whether the data collections are digital or analog. Space Physics had been the dominant discipline for digital data collections, 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 surpassed the Space Physics counts and in 2012 nearly all the new digital data collections were Planetary.

NSSDC manages a number of analog data collections (e.g. film, microfilm and microfiche) comparable to that of the digital archive, though in recent years newly arriving data have been all digital. NSSDC has generated digital versions for some of its film archive, often in response to requests.

Table 1. Counts of NSSDC Data Collections on December 31, 2012

Discipline	Digital	Analog	Total
Astrophysics	229	76	305
Space/Solar Physics	1318	665	1984
Planetary Science	1419	761	2180
Earth Science	106	129	235
Other	118	440	558
Total	3,190	2,071	5,261

Table 2 is a different characterization of the NSSDC archive, showing byte counts for the entire digital archive. Some of the byte counts, especially for older media, are estimates assuming the mean numbers of bytes on various media types for some data collections, though each year any uncertainty from such estimates has less effect due to accurate accounting for larger new data collections. The totals in Table 2 are also adjusted for any data removed from NSSDC, e.g. the 1.63 TB of Earth Science data which were transferred to the Goddard DISC, completing the transfer of our digital Earth Science holdings. In 2012 the greatest changes were in the disciplines of Astrophysics and Planetary, whose holdings increased by more than 5 TB and 3 TB, respectively. "Other" includes mainly Legacy data and other multi-disciplinary missions such as Gravity Probe-B. In recent years some data had been mistakenly included in "Other" and this year have been corrected, thus discipline subtotals may not be comparable to those in recent reports, but "Total" is correct. The main shift was for ~23TB of FERMI data previously been included as "Other" which were corrected to "Astrophysics" in this report.

Table 2. State of the NSSDC Archive December 31, 2012

All Digital Data (TB)	
Astrophysics	104.03
Earth Science	0.00
Space Physics	28.29
Planetary	19.53
Other (Legacy, etc)	0.87
Total	152.71

Data are also being moved from NSSDC's traditional off-line archive to a near-line archive based on DLT, SDLT, and LTO jukeboxes attached to Unix or Linux servers. Data are newly archived in Archive Information Packages (AIPs), which hold both 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 the years 2010-12; the data stored as AIPs has reached a total of 23 TB. Legacy Data refers to data on aging magnetic tapes such as 9-tracks or 3480 cartridges, which are being read and ingested as AIPs in an on-going effort.

Table 3. Data Ingested to Nearline Permanent Archive

	2010		2011		2012	
	AIPs	GB	AIPs	GB	AIPs	GB
ALOUETTE						
APOLLO	2	0.02				
GP-B	532	82.41				
ISIS						
LEGACY DATA	1282	50.51	2404	55.43	3715	75.82
PDS DATA	2180	2492.14	2400	5976.04	1012	3096.60
RHESSI	6178	424.74	10988	842.81	7787	586.82
SPDF	2	0.06				
TWINS			2566	634.49	688	241.18
WIND	275	0.22	397	0.23	367	0.21
TOTALS	10,451	3050.09	18,755	7509.01	13,569	4000.63

Selected data collections stored in AIPs are made network-accessible on nssdcftp for the convenience of the user community. Table 4 lists NSSDC's network-accessible Space Physics data as of 31 December 2012 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 4. Space Science Data FTP Accessible from NSSDC on December 31, 2012

Top Level Dir	GB 2012	Top Level Dir	GB 2012	Top Level Dir	GB 2012
ACE	58.561	GEOTAIL	2.945	OGO	3.804
AE	1.279	HELIOS	2.689	OMNI	31.165
AMPTE	5.014	IBEX	9.068	PIONEER	5.461
APOLLO	14.849	IMAGE	260.995	SAMPEX	54.070
CRRES	34.061	IMP	116.661	SOLAR_MAX	20.892
DE	186.642	ISEE	36.452	ULYSSES	297.339
DIRBE	19.963	ISIS	207.126	VOYAGER	31.015
EXPLORER	9.363	MAGSAT	1.869	WIND	67.423
				Others*	9.185
		TOTAL	1487.891		

* 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.

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. Not included are tapes stored as a courtesy to our GSFC colleagues, e.g. 252 4-mm tapes of ISEE data for SPDF. The transfer of the NSSDC Earth Science tapes to the Goddard DISC is nearly complete, so the

2012 counts show a reduction of more than 6700 tapes in the Earth Science discipline. The remaining 150 tapes and CDs are special cases, which are being transferred as they are resolved. Note that, at most, they represent 0.03 GB of data total, so we entered 0.0 GB in Table 2 above. Data from over 1100 legacy tapes are in progress to be ingested as AIPs; when that occurs, those tapes will be released from the archive and no longer counted in Table 5.

Table 5. Counts of Volumes* at NSSDC Archive on Dec 31, 2012

	Astro- physics	Space Physics	Planetary Science	Earth Science	TOTAL
4-mm Tape	1194	94	3	0	1291
8-mm Tape	205	199	77	0	481
9-Track Tape	531	4794	1679	134	7138
3480 Cartridges	474	1984	1156	7	3621
DLT	84	114	1	0	199
LTO	71	0	0	0	71
CD	855	23182	5346	9	29392
DVD	1018	1297	255	0	2570
12-in Worm	0	4	0	0	4
M-O Disk	274	0	0	0	274
TOTAL	4,706	31,668	8,517	150	45,041

* Backup volumes and those not attributable to these four 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. Film(frames) are exclusive of Film (feet); no images are double counted, neither are backup or extra copies counted here. Most entries in this table are unchanged from 2008, when we completed an inventory of the microfilm and microfiche portions of the analog archive and we expect no additions to Table 6 in the foreseeable future. However, for 2012 we have updated the Film(frames) column to reflect a request from NARA for which we have processed a transfer of 1877 NASA Press Release photos to them.

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	11971	0	0	62
Earth Science	55	773	13067	64572	4	0	0
Planetary Science	2188	13492	409081	333218	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	422,140	207	1773	39,835

The continuing inventories of other categories -- e.g. sorting through duplicate sets of media -- were suspended in 2011 due to budgetary considerations, though routine inventory review and

updates continue. Note that additional NSSDC analog materials beyond those in Table 6 are in long-term storage off-site and their updated inventory remains a background project.

3.1 Data Inflow for 2012

Table 7 reports on data arriving on media by media type and 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 LTOs have replaced CDs and DVDs as the dominant media type received. With their larger capacity, the much smaller number of LTO

Table 7. Media Arriving at NSSDC in 2012

	Astrophysics	Space Physics	Total
4-mm tapes	473	0	473
DVDs	2	28	30
LTOs	40	0	40
Totals	515	28	543

Table 8. Data Arriving at NSSDC During 2012		
	GB	Sub-Totals by Discipline
Astrophysics		
FERMI	4000.00	16,829.79
GALEX	1718.00	
HEASARC	3407.77	
WISE	7704.02	
Planetary		
PDS_ATM	859.64	3,096.60
PDS_GEO	1210.45	
PDS_IMG	234.24	
PDS_NAI	301.45	
PDS_PSI	26.09	
PDS_SBN	464.73	
Space Physics		
RHESSI	586.82	660.56
TWINS	73.54	
Wind	0.21	
Miscellaneous		
Legacy Data	75.82	75.82
TOTAL		20,662.77

tapes delivered a larger volume 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 2012, totaling over 20 TB of new data via a combination of electronic deliveries and media. Dominating the statistics are data from Astrophysics and Planetary (reported by PDS node). The actual HEASARC delivery was much larger, but here again we report a net value; by the terms of the MOU data collections delivered earlier are released from NSSDC.

3.2 Data Outflow for 2012

NSSDC provides access to its data holdings with network-accessible data for chosen data collections and, in addition, through a user support infrastructure for the mailing of off-line digital and analog data volumes. Most electronic interfaces are accessible through NSSDC's home page and include special web-based interfaces to specific data collections or groups thereof and ftp pathways to a range of data files maintained permanently on nssdcftp.

Much of the data outflow discussed in NSSDC Annual Reports before 2003 was activity within the Space Physics Data Facility (SPDF), which maintains the Active Archive for NASA Space Physics missions; their activities are now reported elsewhere. Because NSSDC and SPDF are

co-located at NASA's Goddard Space Flight Center, nssdcftp remained a shared resource until September 2012, when SPDF moved data services to its own site (see spdf.gsfc.nasa.gov). Thus the nssdcftp 2012 ftp statistics are pre- and post-transition and not comparable to other years.

Table 9 gives the annual counts of files downloaded from nssdcftp for 2012 and four prior years; it singles out those subdirectories with high activity, Photo Gallery and Spacecraft Data. The Photo Gallery downloads are comparable to recent years. Researchers downloading data files via ftp from the Spacecraft Data subdirectory had a notable decrease, perhaps because of the NSSDC and SPDF changes. The coming years will show any lasting effects of the change.

Table 9. Number of Files Downloaded via FTP

	2008	2009	2010	2011	2012
Photo Gallery	481,089	193,577	152,533	129,161	171,970
Spacecraft Data	1,856,362	1,564,930	2,314,095	2,436,245	923,808
All others on nssdcftp	634,070	615,242	789,413	468,459	361,598
TOTAL	2,971,521	2,373,749	3,256,041	3,033,865	1,457,376

NSSDC responded to 18 distinct requests for ftp data and for data products; 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 (Off-line Requests Only) for CY 2012

Affiliation Category	Total Requests
No Affiliation [General Public]	4
Non_US	1
US Academic Institutions	2
US Private Industry	1
NASA/GSFC	10
Total	18

Table 11. Number of Requests for Off-line Data by Discipline

DISCIPLINE	Data Collection Requests
Astrophysics	2
Earth Science	1
Planetary Science	3
Space & Solar Physics	6
TOTAL	12

Table 11 gives the counts of requests for off-line data collections from various disciplines in 2012. The planetary data requests are largely associated with lunar and planetary image data that are widely requested. 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. Also, some requests are for documents or other forms that are not counted as data collections, thus are not included in Table 11.

In 2012 data distribution via ftp was so dominant that it is no longer informative to include Table 12a, which previously counted distributions of CDs, maps, etc. Table 12, comparable to 12b in previous reports, gives the distribution of requested data by discipline for recent years. Note that

the statistics in Table 12 do not include any repeat requests for data previously made available via ftp, but only data that were newly posted in response to a request.

Tables 12. NSSDC Off-line Data Dissemination Statistics 2008-2012

DISCIPLINE	2008 Items*	2009 Items*	2010 Items*	2011 Items*	2012 Files*
Astrophysics	95	10,737	27	124	48
Planetary	20,979	740	932	31,891	453
Space Physics	102	9,470	6,319	57	10,517
Other	0	15	13	463	1,376
TOTAL	21,176	20,962	7,291	32,535	12,394

** files counts are not comparable to earlier item counts*

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, though this effort is not yet underway.

Table 13. NIMS Database Statistics for CY 2012

Subpartition	Number of Records 12/31/2012	Number Added in 2012
Spacecraft	6,920	129
Experiment	5,438	6
Data collection	6,165	151
TOTAL	18,523	286

Number of spacecraft with experiment records - 1,089
 Number of experiments with data collections at NSSDC - 1,641
 Additional data collections associated only with spacecraft, not experiments - 771
 Additional data collections that are not associated with spacecraft/experiment - 506

Table 13 details the pertinent NIMS database statistics for 2012. NIMS identifies virtually all launched spacecraft, the experiments carried by many of these spacecraft, and data collections from these spacecraft primarily as archived at NSSDC. This portion of the database is the source

of information for many of NSSDC's web pages. The NSSDC Master Catalog (NMC) dynamically generates web 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. The counts of new records are also detailed but, because of routine database maintenance, those do not exactly predict the totals from year to year.

4.2 SPASE and Virtual Observatories (VOs)

NSSDC staff have been involved in the Heliophysics Data and Model Consortium for many years. This organization, composed of a wide representation of the Heliophysics data holding and archiving community, is attempting to unify and facilitate access to Heliophysics data for the research community. A key aspect of this work is the development of a common data description model named the Space Physics Archive Search and Extract (SPASE) Data Model. NSSDC staff have been involved in this project since it was first proposed in 1998. Coordination of the development of the SPASE Data Model continues to be provided by the NSSDC.

Version (2.2.2) of the SPASE Data Model was released 16 October 2012. SPASE has been used to describe the majority of NASA-funded experimental data holdings and is spreading to usage within the ground-based Heliophysics data community. The next step in the SPASE Data Model effort will be to incorporate support for modeling and simulation data. The Data Model usage is increasing in the Heliophysics community and SPASE will be an assumed standard for the Cooperative Agreement Notice for Heliophysics Data Services to be released in 2013. For more information see <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* was approved as a CCSDS Draft Standard and released for CCSDS Agency Review in February 2012. Subsequently prototyping efforts were slowed by budget cuts. ESA since has been adapting NASA's implementation and, along with the CNES prototype, is identifying desired updates to the standard.
- *Reference Model for Open Archival Information System*: the world's most recognized digital archiving standard was updated and became an approved CCSDS Standard in June 2012. It was submitted to ISO and published as a new version of the ISO Standard in August.
- *Audit and Certification of Trustworthy Digital Repositories*: had become a CCSDS Standard in September 2011 and was published as an ISO standard in February 2012.

NSSDC agreed to participate as a subject in a test audit using the latter two items from the list above. 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 reviewed 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 is doing 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 NSSDC participation, prepared for and orchestrated Sun-Earth Day for 2012, a double observance united under the theme “Shadows of the Sun.” First was the annual equinox celebration on March 19. The subsequent Transit of Venus was an international highlight on June 5/6. It was the most successful program to date, with over 600 million web hits and a TV viewership of nearly 8 million, over 570,000 at its peak. Many thousands of advance packets of information were sent to teachers, scientists and others for the Sun-Earth Day programs. 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 the next Sun-Earth Day on March 22, 2013 using the theme “Solar Max, Storm Warning”. This will feature a webcast by the NASA EDGE team (Langley-based group) highlighting the launches and work done by the Wallops Flight Facility, especially in support of Heliophysics science. For more information the reader is referred to <http://sunearthday.nasa.gov>.

NSSDC hosts a website with information about the Moon Tree Project as well. The Moon Tree project tracks the many trees that were planted around the world with seeds taken to the Moon by Apollo 14 and returned for planting. In 2012 NSSDC hosted a visit by Erica Botkin, a writer and photographer, who is working on a photo project and book on the Moon Trees. She has visited and photographed many of the trees, and sent the images to NSSDC for posting on its "Moon Trees" web site. We have also been in contact with and gotten information from Glenn Kovar, former director of the Moon Trees project with the Forest Service.

NSSDC staff also are involved in the coordination of the Radio Jove program which enables students, informal education groups, and other interested individuals to learn about radio astronomy by building their own radio telescope for observing radio emissions from the Sun, Jupiter, the galaxy and the Earth itself. In 2012 the number of kits that have been distributed by the project worldwide surpassed 1600 and the amateur radio observers continue to contribute the results of their observations through email to the general distribution list as well as a general archive of observations. In 2013 the project will continue to expand and inspire international education and general audiences about astronomy that can be done without an optical telescope.

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	Linear Tape-Open [i.e. open standard]
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 & Extract
SPDF	Space Physics Data Facility
TB	Terabyte
TWINS	Two Wide-angle Imaging Neutral-atom Spectrometers
VO	Virtual Observatory
WORM	Write-Once, Read-Many