

2013 ANNUAL STATISTICS AND HIGHLIGHTS
for
THE NATIONAL SPACE SCIENCE DATA CENTER

Edwin J. Grayzeck

National Space Science Data Center

Greenbelt, Maryland 20771

2014-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

4.5. Other Staff Projects

Glossary

PREFACE

The National Space Science Data Center (NSSDC) serves as the deep 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 2013. 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 2013 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 2013, with a focus on the 2013 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 12 Tables which summarize NSSDC activities in 2013. In most cases these numbers speak for themselves, though here we address a few highlights.

The most important result of NSSDC's 2013 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 grown to 233 TB of space science data. Over 80 TB of data were added to the NSSDC archive in 2013; the largest data contributors were 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 2013 the main providers of data to be stored as AIPs in the near line archive were the nodes and subnodes of the Planetary Data System (PDS). Our major development efforts this year have been preparing to receive and archive data in the new PDS4 form, which the PDS nodes expect to begin using in late 2014. NSSDC also provides off-site backup service for data submitters' original media. Levels of archive services provided by NSSDC are discussed in more detail at http://nssdc.gsfc.nasa.gov/nssdc/submitting_data.html.

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 2013 NSSDC was managing 5,582 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 2013 nearly all the new digital data collections were Planetary, supplemented by a few "Other".

NSSDC manages a number of analog data collections (e.g. film, microfilm and microfiche) comparable to that of the digital archive. In recent years new data collections have been all digital, while the number of analog collections have declined as we reviewed and re-inventoried a portion of those holdings.

Table 1. Counts of NSSDC Data Collections on December 31, 2013

DISCIPLINE	DIGITAL	ANALOG	TOTAL
Astronomy	229	76	305
Heliophysics	1318	656	1974
Planetary Science	1749	761	2510
Earth Science	106	123	229
Other	124	440	564
TOTAL	3526	2056	5582

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; 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 Earth Science data which were transferred to the NASA ESDIS or data which are replaced by new submissions. In 2013 the greatest changes were in the disciplines of Astrophysics and Planetary, whose holdings increased by more than 77 TB and 3 TB, respectively. The large Astrophysics contribution was an anomaly, as HEASARC delivered a new baseline of the entirety of their holdings. “Other” includes mainly Legacy data and other multi-disciplinary missions such as Gravity Probe-B.

Table 2. State of the NSSDC Archive December 31, 2013

All Digital Data (TB)	
Astrophysics	180.70
Planetary	22.62
Space Physics	28.88
Earth Science	0.00
Other	0.89
Total	233.09

Digital data are stored in a near-line archive based on DLT, SDLT, and LTO jukeboxes attached to 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 2011-13; the data stored as AIPs has reached a total over 26 TB. Legacy Data refers to data on aging magnetic tapes such as 9-tracks or 3480 cartridges, which are being read and migrated to AIPs in an on-going effort. The data called out as NSSDC Data Conversion is just that – datasets converted to newer, more universal formats (e.g. ascii) from earlier submissions in antique formats (e.g. VMS backup). PDS data are predominant among the 2013 newly produced AIPs.

Table 3. Data Ingested to Nearline Permanent Archive

	2011		2012		2013	
	<i>AIPs</i>	GB	<i>AIPs</i>	GB	<i>AIPs</i>	GB
LEGACY DATA	2404	55.43	3715	75.82	1224	19.17
NSSDC Data Conversion					209	0.65
PDS DATA	2400	5976.04	1012	3096.60	1834	3093.75
RHESSI	10988	842.81	7787	586.82	7302	591.61
TWINS	2566	634.49	688	241.18		
WIND	397	0.23	367	0.21	366	0.21
TOTAL	18,755	7509.01	13,569	4000.63	11,791	3705.40

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 2013 for projects with > 1GB in the archive. 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, 2013

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	PIONEER	5.461
AMPTE	5.014	IBEX	9.068	SAMPEX	54.070
APOLLO	14.849	IMP	116.661	SOLAR_MAX	20.892
CRRES	34.061	ISEE	36.454	ULYSSES	297.328
DE	186.642	ISIS	207.126	VOYAGER	31.015
DIRBE	19.963	MAGSAT	1.869	WIND	67.423
EXPLORER	4.761			Others*	9.185
		TOTAL	1191.120		

* 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 within NSSDC, not including backup copies. It should be noted that most volumes are replicable and have one backup volume. The remaining 40 tapes are special cases, which are being transferred as they are resolved. Note that they represent a negligible amount of data, so we entered 0.0 GB in Table 2 above. Data from over 1894 legacy tapes are in progress to be ingested as AIPs. For now those legacy tapes are included in the counts, but they will be released from the archive and no longer counted in Table 5 in a future step of the legacy data recovery.

Table 5. Counts of Volumes* at NSSDC Archive on Dec 31, 2013

	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	4795	1401	33	6760
3480 Cartridges	474	1983	1156	7	3620
DLT	84	101	2	0	187
LTO	131	0	0	0	131
CD	519	23183	2717	0	26419
DVD	1014	1047	213	0	2274
12-in Worm	0	4	0	0	4
M-O Disk	0	0	0	0	0
TOTAL	4152	31406	5569	40	41167

* 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 2012, but we have updated those for Microfilm and Microfiche based on an on-going inventory of analog holdings. We expect no new additions to Table 6 in the foreseeable future, only updates from further inventory and review of our current analog holdings.

Table 6. Analog Data Products at NSSDC by Discipline

Discipline	Microfilm (reels)	Microfiche (cards)	Film (ft)	Film (frames)	Movie (reels)	Maps	Slides
Astrophysics	2382	5963	100	11971	0	0	62
Earth Science	198	983	13067	64572	4	0	0
Planetary Science	2095	13492	409081	333218	181	1773	768
Space Physics	34915	30507	580	6173	2	0	36893
Other	3061	3620	3785	6206	20	0	2112
Totals	42,651	54,565	426,613	422,140	207	1773	39,835

3.1 Data Inflow for 2013

Table 7 reports on data arriving on media by type and by discipline. In recent years LTOs have been the dominant media type received. The CD was Space Physics data, all other media received were backup data archived for Astrophysics groups. Note that each year we do receive additional media such as data bricks, i.e. portable hard drives, which are just transfer media and so not included in Table 7.

Table 7. Media Arriving in 2013

Type	Count
DVDs	2
CDs	1
LTOs	93
TOTAL	96

Table 8 lists by project the data volumes that NSSDC received in 2013, totaling over 80 TB of data via a combination of electronic deliveries and those on media. Dominating the statistics are data from Astrophysics and the PDS nodes. The large HEASARC delivery was a complete backup of their entire archive, a new baseline for further incremental deliveries. By the terms of the MOU with HEASARC, their previous deliveries were released from NSSDC and returned to them.

3.2 Data Dissemination for 2013

NSSDC provides access to chosen collections with network-accessible data. 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. Off-line digital and analog data volumes are made available via ftp by request through a user support infrastructure.

In May 2013 nssdcftp was moved to a GSFC cloud as part of a GSFC consolidation effort. There was no impact to the NSSDC user community. One effect of the move was a break in tracking the monthly ftp access statistics. Rather than report incomplete statistics, we have opted to omit the table "Number of Files Downloaded via ftp" reported in recent annual reports.

NSSDC responded to 30 distinct requests for ftp data products; Table 9 characterizes the user community of requestors, including domestic and international researchers, the general public, educational enterprises, publishers, etc.

Table 9. NSSDC User Community (Off-line Requests Only) for CY 2013

Affiliation Category	Total Requests
No Affiliation (General Public)	3
Non-US	8
US Academic Institutions	7
US Private Industry	2
NASA/GSFC	5
NASA Centers, excluding GSFC	4
Other Government Agencies	1
Total	30

Table 8. Data Arriving at NSSDC During 2013

	GB	Sub-Totals by Discipline
<i>Astrophysics</i>		76962.44
FERMI	11200.00	
GEOTAIL	0.72	
HEASARC	65761.72	
<i>Planetary</i>		3093.76
PDS_ATM	497.57	
PDS_GEO	1463.84	
PDS_NAI	385.07	
PDS_PPI	305.91	
PDS_PSI	89.97	
PDS_SBN	351.40	
<i>Space Physics</i>		591.88
RHESSI	591.61	
SPDF	0.06	
WINDWAV	0.21	
<i>Miscellaneous</i>		19.82
Legacy Data	19.17	
Data Conversions	0.65	
GRAND TOTAL		80667.90

Table 10. Number of Requests for Off-line Data by Discipline

DISCIPLINE	Data Collection Requests
Planetary Science	9
Space & Solar Physics	12
TOTAL	21

Table 10 gives the counts of requests for off-line data collections from various disciplines in 2013. 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 9 because some requests are for data/items related to more than one discipline, so are double counted in Table 9. Also, some requests are for documents or other forms that are not counted as data collections, thus are not included in Table 10.

In 2013 virtually all data distribution was via ftp. Table 11 gives the distribution of requested data by discipline for recent years. Note that the statistics in Table 11 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. Note also that for 2012-2013 we have reported the number of files, easily counted for ftp deliveries, and not comparable to counts of “Items” used in earlier reports.

Table 11. NSSDC Off-line Data Dissemination Statistics 2009-2013

DISCIPLINE	2009	2010	2011	2012	2013
	<i>Items*</i>	<i>Items*</i>	<i>Items*</i>	<i>Files*</i>	<i>Files*</i>
Astrophysics	10,737	27	124	48	0
Planetary	740	932	31,891	453	11,512
Space Physics	9,470	6,319	57	10,517	32,387
Other	15	13	463	1,376	0
TOTAL	20,962	7,291	32,535	12,394	43,899

** 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 12 details the pertinent NIMS database statistics for 2013. 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.

Table 12. NIMS Database Statistics for CY 2013

Subpartition	Number of Records 12/31/2013	Number Added in 2013
Spacecraft	6,920	148
Experiment	5,438	22
Data collection	6,165	303
TOTAL	18,988	473

Number of spacecraft with experiment records - 1,092

Number of experiments with data collections at NSSDC - 1,654

Additional data collections associated only with spacecraft, not experiments - 831

Additional data collections that are not associated with spacecraft/experiment - 523

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 with it since it was proposed in 1998, providing coordination of the development of the SPASE Data Model.

The latest version (2.2.2) of the SPASE Data Model was released 16 October 2012. Work on SPASE continues. It 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. Discussions continue on how to incorporate support for modeling and simulation data. 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)

NSSDC previously provided the lead for the Data Archive Ingest (DAI) Working Group (WG) within the Consultative Committee for Space Data Systems (CCSDS) and also the deputy chair of the Repository Audit and Certification (RAC) WG, but due to funding limitations NSSDC participation in these committees has been reduced and transferred to an outside consultant. The remaining NSSDC effort is concentrated within the Data Archive Ingest group, which has been finalizing the Producer Archive Interface Specification (PAIS). PAIS is of interest to both NSSDC and PDS. It provides a way to capture the logical model and internal associations for a data collection using Collection and Transfer Object descriptors, which specify occurrences and sizes of transfer objects. The reader is referred to <http://www.ccsds.org/> for more specifics on the CCSDS activities.

4.4 Sun-Earth Day

The NASA Sun-Earth Day team, with NSSDC participation, prepared for and orchestrated Sun-Earth Day for 2013, observed March 22nd under the theme “Solar Max: Storm Warning.”

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 Sun-Earth Day team also provided expert communications surrounding the September 6th launch from Wallop’s Flight Facility of the NASA Lunar Atmosphere and Dust Environment Explorer (LADEE) and supported the Planetary Science Division campaign to observe Comet ISON (see <http://cometison.gsfc.nasa.gov/>).

Though Sun-Earth Day is enduring a funding gap after 2013, the same team of scientists and educators has since been moving on, supporting science communication related to NASA planetary missions. For more information the reader is referred to <http://sunearthday.nasa.gov>.

4.5 Other Staff Projects

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. More than 1800 kits that have been distributed by the project worldwide 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 Radio Jove announced a joint conference with the Society of Amateur Radio Astronomers (SARA) at the National Radio Astronomy Observatory in Green Bank, West Virginia for June 29 - July 4, 2014.

Glossary

ACE	Advanced Composition Explorer
AE	Atmospheric Explorer
AIP	Archive Information Package
ARCAD	Arc Aurorale et Densite
CCSDS	Consultative Committee for Space Data Systems
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)
ESDIS	Earth Science Data and Information System
FTP	File Transfer Protocol
GB	Gigabyte
GSFC	Goddard Space Flight Center
HEASARC	High Energy Astrophysics Science Archive Research Center
IBEX	Interstellar Boundary Explorer
IMP	Interplanetary Monitoring Platform
ISEE	International Sun-Earth Explorer
ISIS	International Satellite for Ionosphere Studies
ISO	International Organization for Standardization
ISON	International Scientific Optical Network
LTO	Linear Tape-Open [i.e. open standard]
MAGSAT	MAGnetic field SATellite
M-O	Magneto-optic
NIMS	NSSDC Information Management System
NMC	NSSDC Master Catalog
NSSDC	National Space Science Data Center
OGO	Orbiting Geophysical Observatories
PAIS	Producer Archive Interface Specification
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