2009 ANNUAL STATISTICS AND HIGHLIGHTS REPORT FOR

THE

NATIONAL SPACE SCIENCE DATA CENTER

Edwin J. Grayzeck

National Space Science Data Center

Greenbelt, Maryland 20771

2010-Sept-27

Table of Contents

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 astrophysics and space physics 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 2009. 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. This report covers only the NSSDC. Reports before 2003 covered both the NSSDC and the Space Physics Data Facility (SPDF), who were organizational peers within the GSFC Space Sciences Data Operations Office (SSDOO). Following a NASA reorganization the two entities are now in separate GSFC divisions, NSSDC within the Solar System Exploration Division and SPDF within the Heliophysics Science Division. Note also that many statistics in this report are only comparable to those from the 2005 and later, since our central NIMS database was revamped in 2005.

NSSDC is pleased to issue this 2009 Annual Report describing the 2009 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 2009, with a focus on the 2009 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 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.

2. HIGHLIGHTS

The center of this report is the 13 Tables which summarize NSSDC activities in 2009. In most cases these numbers speak for themselves, though it is irresistible to address a few highlights.

The most important result of NSSDC's 2009 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 88 TB of space science data and an additional 3.3 TB of Earth science data. During 2009, 14.5 TB of data were added to the NSSDC.

In 2009 we defined the first dataset for distribution from the Gravity Probe-B total project archive, i.e. the Level 2 data, out of the full 0.54 TB which was delivered via RAID in 2007. That data is included in the year's statistics. The entirety of GP-B data are not yet included, but will be ingested soon now that updates for data sets delivered via RAID have been completed.

In 2001 NSSDC began using a reengineered data management approach, which stores data as Archive Information Packages (AIPs; bundles of data files and companion attribute files as prescribed by the ISO/CCSDS Open Archival Information Systems reference model) written to DLTs (to SuperDLTs since 2006). The X-man software, developed at NSSDC, allows data providers to produce a manifest with the most pertinent metadata. The manifest is submitted with the data to NSSDC and packaging into AIPs is done locally. This streamlined method of data archiving has been in use over a year and has lived up to all our hopes. In 2009 several improvements were made, focused on ingesting data received via a variety of delivery methods.

3. DATA MANAGED AT NSSDC AND 2009 INFLOW AND OUTFLOW

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, 2009

Discipline	Digital	Non-Digital	Total
Astronomy	229	76	305
Space/Solar Phys	1212	667	1879
Planetary	701	761	1462
Earth	110	131	241
Other (incl Ephemeris)	98	440	538
Total	**2350**	**2075**	**4425**

By the end of 2009 NSSDC was managing 4,425 distinct datasets and accompanying documentation packages. Table 1 indicates the disciplines from which these datasets come and whether the datasets are digital or non-digital. Space physics is the dominant discipline by dataset counts, accounting for nearly half of NSSDC's holdings. This reflects that in its early years NASA launched a preponderance of space physics missions and also that space physics spacecraft typically carry more independent experiments than do astrophysics missions. The total number of datasets shows a slight decrease from last year, which is an artifact of our on-going efforts to move data from old magnetic media (e.g. 9-track tapes). Part of the process is to review documentation, disentangle merged datasets and remove duplicate entries; there has been no actual decrease in datasets, but actually a small increase.

NSSDC manages almost as many non-digital (e.g. film, microfilm and microfiche) datasets as digital datasets, though in recent years newly arriving data has been all digital. NSSDC also 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. We foresee a shift in coming years with the expected arrival of large planetary datasets. We have received our first large data deliveries from nodes of the Planetary Data System, much of which has been ingested as AIPs.

Table 2. State of the NSSDC Archive December 31, 2009

	All Digital Data (TB)
Astrophysics	47.24
Space Physics	27.03
Planetary	6.99
Earth Science	3.28
Other	7.20
Total	**91.75**

Data are also being moved from NSSDC's traditional offline archive to a near line archive based on DLT and SDLT 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 2007-9; the total of data stored as AIPs has reached 8.6 TB.

Table 3. Data Ingested to Nearline Permanent Archive

	2007		2008		2009	
	AIPs	GB	*AIPs*	GB	*AIPs*	GB
ALOUETTE	17151	9.79	17208	9.42	16281	9.43
ISIS	133581	67.34	114775	40.40	22168	10.47
LEGACY DATA	310	2.48	871	6.40	1073	14.97
PDS DATA	5	62.26	88	288.95	171	247.66
RHESSI	8042	574.88	8216	587.03	7233	577.52
WIND	332	0.26	346	0.29	367	0.30
Totals	159,422	717.25	141,504	932.50	47,293	860.35

Table 4. Space Physics Data FTP Accessible from NSSDC on December 31, 2009

Top Level Dir > 1 GB	GB
ACE	48.518
AE	1.279
APOLLO	7.160
CRRES	34.061
DE	186.642
DIRBE	19.963
EXPLORER	3.695
GEOTAIL	2.924
HELIOS	2.538
IMAGE	260.995
IMP	105.761
ISEE	19.336
ISIS	206.721
MAGSAT	1.869
OGO	1.889
OMNI	28.349
PIONEER	3.169
SAMPEX	54.070
ULYSSES	276.636
VOYAGER	29.890
WIND	29.711
Others*	8.534
TOTAL	1333.710

About half of the data stored in AIPs are made network-accessible on nssdcftp for the convenience of some portions of the user community. Table 4 (below) lists NSSDC's network-accessible Space Physics data as of 31 December 2009 for projects with > 1GB. Spacecraft with less data -- including Aeros, Alouette, ARCAD, Cassini, DMR, Galileo, Genesis, Hinotori, Mariner 10, Prognoz 6,7, and 9, San Marco, SNOE and miscellaneous others – are omitted from the table. It must be noted that nssdcftp has a continually evolving directory structure, so the table represents only a snapshot.

Table 5 (below) 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 majority of the Earth Science tapes are in transition to NASA's Earth Science archive, thus the 2009 summary shows a reduction of about 7000 9-track tapes in the Earth Science discipline.

Table 5. Counts of Volumes* at NSSDC Archive on Dec 31, 2009

Media	Astro Physics	Space Physics	Planetary Science	Earth Science	Total
4-mm Tape	587	92	3	97	779
8-mm Tape	189	503	74	0	766
9-Track Tape	531	4743	3721	10153	19148
3480 Cartridges	491	1940	1126	2688	6245
DLT	65	31	2	0	98
LTO4	48	0	0	0	48
CD	853	23158	5346	54	29411
DVD	987	1210	255	0	2452
12 Worm	0	4	0	0	4
M-O Disk	274	0	0	0	274
TOTAL	4025	31681	10527	12992	59225

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

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

Table 6 lists NSSDC's analog archive holdings by disciplines and by form factor. The film and movie counts are based on the inventory of the photo materials started in 2007. Movie reels had not been called out separately in previous reports, now they are. Film (feet) are reels of uncut film containing sequences of still photos. We removed from the inventory any backup copies of

images, which had notably increased previously reported numbers. In 2008 we completed an inventory of the microfilm and microfiche portions of the analog archive, increasing both. Apparently the earlier estimates were conservative. This table is unchanged from 2008 and not comparable to versions from earlier years. The inventories are continuing, e.g. sorting through duplicate sets of media, and, when complete, total counts will be updated to reflect those changes.

3.1 Data Inflow

Table 7. Media Arriving at NSSDC During 2009*

	Astrophysics	Space Physics	Planetary	Total
4-mm Tapes	89	0	0	89
SDLTs	0	-6	0	-6
CDs	3	384	1	388
DVDs	19	110	0	129
LTO-4	46	0	0	46
Totals	157	488	1	646

* Ephemeris and Other data not included.

Table 7 characterizes the in-flowing media types by discipline. As in recent years, CDs continued as the dominate input media type received. The reduction in SDLTs reflects data which was redelivered on other media.

Table 8 (below) shows by project the data volumes that NSSDC received in 2009, approximately 14 TB of new data via a combination of electronic deliveries and on media. Dominating the statistics are HEASARC and FERMI in Astrophysics and plus FAST, RHESSI and TWINS within Space Physics. The actual HEASARC delivery was many terabytes larger, but by the terms of the MOU data sets delivered earlier are removed from NSSDC and archived on the newly delivered media. In coming years we expect the largest data deliveries from Planetary missions.

3.2 Data Outflow

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 non-digital 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.

Table 8. Data Arriving at NSSDC During 2009

	GB	Sub-Totals by Discipline
Astrophysics		
FERMI	5600.02	
GALEX	1468.15	
HEASARC	5668.94	
		12737.12
Planetary		
PDS	247.66	
		247.66
Space Physics		
ACE	7.46	
Alouette	9.43	
FAST	251.22	
ISIS	10.47	
RHESSI	577.52	
TWINS	237.22	
Ulysses	1.14	
Wind	0.30	
		1094.76
Other		
Gravity Probe-B	380.48	
Earth Science	-22.95	
Legacy Data	14.97	
		372.49
GRAND TOTAL		**14452.02**

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 is and remains a shared resource for now and that is reflected in the reported statistics.

Table 9. Number of Files Downloaded via FTP

	2005	2006	2007	2008	2009
Photo Gallery	1,190,555	936,039	720,213	481,089	193,577
Spacecraft Data	1,154,900	802,438	689,961	1,856,362	1,564,930
All others on nssdcftp	1,263,719	998,388	930,342	634,070	615,242
TOTAL	3,609,174	2,736,865	2,340,516	2,971,521	2,373,749

Table 9 gives the annual counts of files downloaded, both overall and for selected directories with high activity. The Photo Gallery continued to decline, possibly because of the proliferation of websites with space photos. Researchers downloading data files via ftp from the spacecraft_data subdirectory remained at roughly the same level as previous years, showing the high interest in and great value of these services provided by NSSDC and SPDF on this shared resource.

Table 10. NSSDC User Community (Offline Requests Only) for CY 2009

Affiliation Category	Total Requests	Percent
No Affiliation [General Public]	16	20%
Non_US	17	21%
US Academic Institutions	18	22%
US Private Industry	2	2%
NASA/GSFC	21	26%
NASA Centers, Excluding GSFC	4	5%
Other Government Agencies	2	2%
Miscellaneous	2	2%
Total	**82**	**100.00%**

NSSDC responded to 82 distinct requests for ftp data and for data products. Table 10 above characterizes the user community of requestors. To a very large extent it is the U.S. and international general public, the education enterprise, publishers, etc. and their desire for NASA imagery.

Table 11 (below) gives the counts of requests for offline datasets from various disciplines in 2009 and as integrated over NSSDC's history. Note particularly the dominance of planetary data over both time scales. This is 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.

In 2009 data distribution via ftp was preferred due to its immediacy and to NSSDC budget limitations, though some media was distributed. Table 12a shows the distribution of data served within these categories for 2009 and the previous two years; Table 12b shows the distribution of the items by discipline for the same years. Both Tables 12a and 12b (below) show ftp as the dominant mode of distribution, even though these statistics include only ftp data that were newly posted in response to a request. Table 12 does not include all data previously available via ftp.

Table 11. Number of Requests for Offline Data by Discipline

DISCIPLINE	Data Set Requests 1968 - 2009	Data Set Requests 2009
Astrophysics	11461	6
Earth Science	7156	4
Planetary Science	47594	38
Space & Solar Physics	9184	19
Ephemeris	97	0
Other	45	2
TOTAL	75537	69

Tables 12a,b. NSSDC Offline Data Dissemination Statistics 2007-2009

Table 12a

ITEMS	2007	2008	2009
DISCs	441	186	196
PRINTED	177	109	36
OTHER	9	4	11
FTP	20907	20877	20719
TOTAL	21534	21176	20962

Table 12b

DISCIPLINE	2007	2008	2009
Astrophysics	277	95	10737
Planetary	2311	20979	740
Space Physics	18946	102	9470
Other	0	0	15
TOTAL	21534	21176	20962

NOTES:

DISCs include CDs & DVDs

PRINTED materials include Photos, Posters, Maps, Documents

OTHER media include Microfilm, Microfiche, Tapes, Videos

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.

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. Table 13 details the pertinent NIMS database statistics for 2009.

Table 13. NIMS/JEDS Database Statistics for CY 2009

Subpartition	Number of Records as of 12/31/09	Number Added in 2009
Spacecraft	6557	119
Experiment	5398	26
Dataset	5380	32
Totals	17,335	177

Number of spacecraft with experiment records - 1,079
Number of experiments with datasets at NSSDC - 1,573
Additional datasets associated only with spacecraft, not experiments - 641
Additional datasets that are not associated with spacecraft/experiment - 237

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 and 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 (2.0.0) of the SPASE Data Model was released on 29 April 2009. Accompanying the release of the new version of the SPASE Data Model was a new version of NSSDC's SPASE registry for observatories (spacecraft) and instruments (experiments). The new registry interface now produces human-readable output in addition to the raw XML format and provides a more friendly query interface. In addition it allows the generation of output that is compatible with V1.3 of SPASE as well as V2.0. 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 within the Consultative Committee for Space Data Systems (CCSDS) for the widespread adoption of the Reference Model for an Open Archival Information System (OAIS). The reader is referred to http://www.ccsds.org/ for specifics.

During this past year, an updated OAIS Reference Model CCSDS Standard was produced and is now being reviewed as a new version of the ISO Standard. Also the first additional archiving standard from the OAIS work is the Producer-Archive Ingest Methodology Abstract Standard (PAIMAS), essentially a checklist of activities for ingest to archives. In 2009 PAIMAS was reconfirmed unchanged as an ISO Standard.

The Repository Audit and Certification Working Group within CCSDS is working on two new standards - Audit and Certification of Trustworthy Digital Repositories, which provides best practices to maintain information over the long term, and Guidelines for Auditors of Trusted Digital Repositories, which provides procedures for ISO auditors to certify Archives.

4.4 Sun-Earth Day

In 2009 the NASA Sun-Earth Connection Education Forum (SECEF) team, with major NSSDC participation, prepared for and orchestrated Sun-Earth Day 2009 with the theme "Our Sun -- Yours to Discover" for the main event on March 20. 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. In late 2009 SECEF was shutdown, but the same team of scientists and educators are working on funded activities to support the next Sun-Earth Day (see http://sunearth.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
CDF        Common Data Format
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)
FAST       Fast Auroral SnapshoT
FTP        File Transfer Protocol
GB         Gigabyte
GP-B       Gravity Probe-B
GSFC       Goddard Space Flight Center
HEASARC    High Energy Astrophysics Science Archive Research Center
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
MAGSAT     MAGnetic field SATellite
M-O        Magneto-optic
NDADS      NSSDC Data Archive and Distribution System
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)
OSS        Office of Space Science
RAID       Redundant Array of Independent Disks (or I = "Inexpensive")
RHESSI     Reuven Ramaty High Energy Solar Spectroscopic Imager
SAMPEX     Solar Anomalous and Magnetospheric Particle Explorer
SDLT       Super DLT (see above)
SECEF      Sun Earth Connection Education Forum
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
```