

**NASA Space  
Science Data  
Coordinated  
Archive**

Mail Code 690.1

NASA Goddard Space Flight Center

Greenbelt, MD20771

Web: <http://nssdc.gsfc.nasa.gov/>

01 March 2015



**2014 ANNUAL STATISTICS  
AND HIGHLIGHTS**

## Table of Contents

Preface.....1

Introduction ..... 2

Highlights ..... 2

Data Managed at NSSDCA..... 2

    State of the Archive ..... 2

    Data Inflow ..... 5

    Data Dissemination ..... 6

Additional Services and Activities ..... 7

    NSSDCA Information Management System (NIMS)..... 7

    Modifications to the NSSDCA Information Model to Support PDS4 ..... 7

    SPASE Data Model and Virtual Observatories (VOs)..... 9

    Data Recovery Support..... 9

    Consultative Committee for Space Science Data Systems (CCSDS)..... 9

    Solar System Exploration and Integration..... 9

    Other Staff Projects ..... 9

Glossary ..... 10

Cover Photo Credit: LROC image of Apollo 17 landing site.

## Preface

The NASA Space Science Data Coordinated Archive (NSSDCA) (formerly known as the National Space Science Data Center, or 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. NSSDCA also provides current data access, complementary to the efforts of other NASA active archives, in support of NASA and international research enterprises. Finally, NSSDCA is a conduit for the general public and education community to acquire NASA space science data that may be of interest to them.

Herein we report on the activities of the NSSDCA for the calendar year 2014. 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 NSSDCA evolves, some statistical tables have been updated to better reflect current operations. These are noted in the text.

NSSDCA is pleased to issue this 2014 Annual Report describing the growth and evolution of the data archive and other tools and services, as well as the access to those data and services by NSSDCA'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 NSSDCA services.

Edwin J. Grayzeck  
Head, NASA Space Science Data Coordinated Archive

## Introduction

This report characterizes NSSDCA's data holdings, metadata holdings, access pathways and services at the end of 2014, with a focus on the 2014 activities leading to that end-of-year state. In addition this report characterizes the nature and access to NSSDCA's data and services by its many users from various communities. It is assumed the reader will have a general familiarity with NSSDCA and its mission. For more information see the top level NSSDCA web page at <http://nssdc.gsfc.nasa.gov/>.

## Highlights

The most important result of NSSDCA's 2014 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. The archive has grown to 273 TB of space science data, a growth of around 40TB. The largest data contributor was the Planetary Data System (PDS).

In 2014, as in 2013, the major development effort has been preparing to receive and archive data in the new PDS4 format, which some PDS nodes began using operationally in late 2014 starting with the LADEE mission. As shown later in the report, this format has required the development of an entirely new model for how PDS data are ingested into the archive.

NSSDCA also continues to provide off-site backup service for selected 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](http://nssdc.gsfc.nasa.gov/nssdc/submitting_data.html).

## Data Managed at NSSDCA

### State of the Archive

There are several ways to characterize the multi-disciplinary NSSDCA archive and we use most of them herein to give a true sense of the archive, *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.

Discipline	Digital	Analog	Total
Astrophysics	230	76	306
Heliophysics	1,318	656	1,974
Planetary Science	1,894	761	2,655
Earth Science	106	127	233
Other	125	440	565
<b>Total</b>	<b>3,673</b>	<b>2,060</b>	<b>5,733</b>

Table 1: Counts of data collections at NSSDCA on 31 December 2014.

Table 1 shows a summary of the variety of data collections held in the archive. Shown in the table are those science discipline areas covered by those collections as well as whether or not the collections are digital or analog. Planetary science continues to be the only discipline in which significant growth in terms of digital data occurs.

Some small adjustments were made to the number of analog collections. This was not due to new analog data arriving, but was due to a small number of Apollo Earth photography collections, initially turned over to the Earth Science Data and Information System (ESDIS) at Goddard as part of an on-going effort to consolidate Earth Science data in one location, being returned to the archive.

Discipline	Digital Data (TB)
Astrophysics	179.74
Heliophysics	29.50
Planetary Science	62.24
Earth Science	0.00
Other	1.22
<b>Total</b>	<b>272.70</b>

Table 2: Total volume of the NSSDCA digital archive on 31 December 2014.

Table 2 presents a different view of the archive. It shows the total volume (in TB) for each discipline. Byte counts for older data collections (and some collections held on media from projects such as HEASARC) are estimates assuming that the media on which they are held have a mean number of bytes depending on the medium. The totals shown in the table have also been adjusted to take into account any data released due to replacement (e.g., PDS and HEASARC) or any data transferred to ESDIS. For 2014 there was significant growth (almost 40 TB) in planetary science resulting from data submitted by PDS. A slight drop in astrophysical data resulted

from a complete replacement of data held for HEASARC where the number of replacement tapes was less than the replaced tapes. Figure 1 shows the same information as Table 2, but includes information from 2005-2013 as well.

Digital data are stored in a near-line archive based on SDLT and LTO jukeboxes attached to Linux servers. Data are archived in Archive Information Packages (AIPs) that hold both data files and accompanying attribute files and a media- and platform-independent. These AIPs are defined as per the ISO/CCSDS Open Archival Information System (OAIS) reference model.

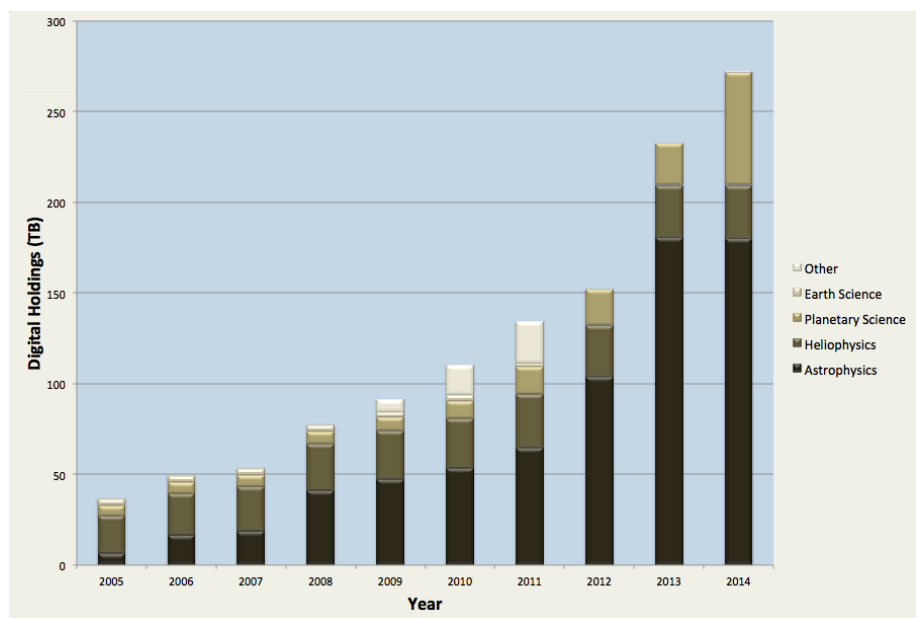


Figure 1: Total volume of the NSSDCA digital archive (2005-2014).

	2012		2013		2014	
	AIPs	GB	AIPs	GB	AIPs	GB
Legacy Data	3,715	75.82	1224	19.17	1,086	15.58
Data Conversion	0	0	209	0.65	0	0
Gravity Probe-B	0	0	0	0	10	326.84
PDS	1,012	3,096.60	1,834	3,093.75	286	50,558.88
RHESSI	7,787	586.82	7,302	591.61	8,094	617.01
TWINS	688	241.18	0	0	0	0
Wind	367	0.21	366	0.21	355	0.20
<b>Total</b>	<b>13,569</b>	<b>4,000.63</b>	<b>10,935</b>	<b>3,705.39</b>	<b>9,831</b>	<b>51,518.51</b>

Table 4: Data ingested to near-line permanent archive.

Top-level Directory	GB
ACE	58.561
AE	1.279
AMPTE	5.014
Apollo	14.849
COBE	41.610
CRRES	34.061
DE	186.642
Explorer	4.761
Geotail	2.945
Helios	2.689
IBEX	9.068
IMP	116.661
IRAS	92.747
ISEE	37.109
ISIS	207.126
Magsat	1.869
OGO	3.804
Pioneer	5.461
SAMPEX	54.070
Solar Maximum	20.892
SWAS	1.353
Ulysses	297.328
Voyager	31.015
WIND	67.858
Others	9.319
<b>TOTAL</b>	<b>1,308.091</b>

Table 3: Space science data accessible via FTP from NSSDCA on 31 December 2014.

Table 4 shows the volume of data ingested to this portion of the archive for the years 2012-2014. The total volume of data ingested as AIPs in 2014 was more than an order of magnitude larger than either of the preceding years, this despite the smaller absolute number of AIPs ingested. This is due to the fact that much of the data taken into the archive during 2014 was on very large data volumes from PDS. Specifically, the largest single volume (LROC data) from PDS in 2014 that was delivered to NSSDCA was 12.5 TB.

Selected data are made network-accessible on a server (<http://nssdcftp.gsfc.nasa.gov/>) for the convenience of the user community. Table 3 provides a list of the data being made available in this manner as of 31 December 2014. The data on this server can be accessed via both http and FTP protocols and continues to be a popular way to access some of the data archived by NSSDCA. Spacecraft included in the "Others" category are those 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 others. It should be noted that the directory structure on this server is continually evolving, so that information in Table 3 is only a snapshot.

Media Type	Astrophysics	Heliophysics	Planetary Science	Earth Science	Total
4-mm Tape	1,194	94	3	0	1,291
8-mm Tape	205	199	77	0	481
9-track Tape	531	4,795	1,401	33	6,760
3480 Cartridges	474	1,983	1,155	10	3,622
DLT	84	101	2	0	187
LTO	147	0	0	0	147
CD	519	23,183	2,533	0	26,235
DVD	1,012	1,047	206	0	2,265
12-in WORM	0	4	0	0	4
<b>Total</b>	<b>4,166</b>	<b>31,406</b>	<b>5,377</b>	<b>43</b>	<b>40,992</b>

Table 5: Counts of media volumes at NSSDCA on 31 December 2014. Backup volumes and those not attributable to the four disciplines listed are not included.

Table 5 presents a characterization of the various off-line media that are managed within NSSDCA. These are data that are not a part of the near-line system. These data are replicable. Data on tape media generally have one backup volume. The items listed under Earth Science represent a negligible amount of the overall data holdings, so 0.0 GB was entered into Table 2 above. A decrease from 2013 in the number of Planetary CD-ROMs resulted from the release of those volumes that PDS submitted replacement data electronically and was archived in AIP form. Legacy data on tape continues to be migrated into AIP form as well and will result in a decrease in media from this table in the future when volumes are released.

Table 6 lists the analog holdings at NSSDCA by discipline and by form. Reels of uncut sequences of still photos are listed as Film (feet). Photographic stills are also held as cut frames. No images are double-counted nor are backup or extra copies counted. Most entries are unchanged from previous years, but significant changes have been made as Earth Science data has been transitioned to ESDIS.

Media Type	Astrophysics	Heliophysics	Planetary Science	Earth Science	Other	Total
Microfilm (reels)	2,382	34,915	2,095	38	3,061	42,651
Microfiche (cards)	5,963	30,507	13,492	548	3,620	54,565
Film (feet)	100	580	409,081	0	3,785	426,613
Film (frames)	11,971	6,173	333,218	31,368	6,206	422,140
Movie (reels)	0	2	181	4	20	207
Maps	0	0	1,773	0	0	1,773
Slides	62	36,893	768	0	2,112	39,835

Table 6: Counts of analog products by form at NSSDCA on 31 December 2014.

## Data Inflow

In previous years we reported on data arriving on various media types. With the sole exception of those entities for which NSSDCA is holding data as a backup function (e.g., HEASARC), all data are now being put in the near-line archive as AIPs and any media on which they arrive is treated as transfer media. In 2014 only 108 tapes (HEASARC) were received by NSSDCA for archive.

Table 7 lists by project the volume of data received by NSSDCA in 2014, nearly 198 TB of data via a combination of electronic deliveries as well as on media. As in 2013, deliveries in Astrophysics and Planetary Science dominate new data. The HEASARC delivery was a full replacement of previously delivered data, a complete backup of their entire archive. As specified in the MOU with HEASARC all previously delivered media were released from NSSDCA and returned to them.

### Data Dissemination

Access to selected collections of data is provided via the server [nssdcftp](#). A snapshot of these data is provided in Table 3 above. In addition, however, NSSDCA also responds to *ad hoc* requests for data. The vast majority of these are now provided in digital form (including digital forms of previously analog materials) and staged for retrieval on [nssdcftp](#).

Discipline/Project	GBTotals by Discipline	
<b>Astrophysics</b>		<b>146,200.00</b>
Fermi	11,200.00	
HEASARC	135,000.00	
<b>Planetary Science</b>		<b>50,558.88</b>
PDS_ATM	865.71	
PDS_IMG	45,056.47	
PDS_PPI	2,359.17	
PDS_PSI	1,763.99	
PDS_SBN	513.55	
<b>Heliophysics</b>		<b>617.21</b>
RHESSI	617.01	
Wind/Wave	0.20	
<b>Miscellaneous</b>		<b>342.42</b>
Gravity Probe-B	326.84	
Legacy Data	15.58	
<b>Grand Total</b>		<b>197,718.51</b>

Table 7: Data arriving at NSSDCA during 2014.

Affiliation Category	Total Requests
No Affiliation (General Public)	1
Non-US	4
US Academic Institutions	9
US Private Industry	1
NASA/GSFC	621
NASA Centers (excluding GSFC)	1
Other Government Agencies	0

Table 8: NSSDCA off-line request user community for 2014.

were a total of 23 requests for off-line data in 2014.

Requests for off-line data are also tracked by science discipline. Table 9 presents the number of requests by discipline for 2014. The total number of requests by discipline (676) exceeds the requests by affiliation (637) because some requests include data associated with more than one discipline. In addition, some requests are requests for documentation or other forms that are not included in Table 9. Finally, as with Table 8, the large number of requests supporting the transfer of data to SPDF greatly reduces the number of requests,

Discipline	Total Requests
Astrophysics	21
Earth Science	7
Heliophysics	622
Planetary Science	24
Ephemeris	2

Table 9: Number of off-line requests for data by discipline for 2014.

Requests for data come from a wide user community. Table 8 shows the number of requests for 2014. The large number of requests for NASA/GSFC is the result of transferring a large number of Heliophysics data collections to the Space Physics Data Facility (SPDF) at GSFC. Of the 621 requests originating at Goddard, 618 of them were from this transfer of data. Excluding these, there



not only in Heliophysics, but in the other disciplines as well. Excluding these requests reduces the overall total discipline requests to 19, eliminating all but two each in Astrophysics and Heliophysics and 15 in Planetary Science.

In 2014 all data distribution was via electronic distribution. Table 10 presents the distribution of requested data by discipline for recent years. The information provided does not include any repeat requests for data previously made available electronically, but only data that were newly posted in response to a request. It is noteworthy that an error was made in the 2013 annual report in that the number of files transferred for Heliophysics and Planetary Science were accidentally transposed. This has been corrected in Table 10.

Discipline	2010 (items)	2011 (items)	2012 (files)	2013 (files)	2014 (files)
Astrophysics	27	124	48	0	8,674
Heliophysics	6,319	57	10,517	11,512	15,643
Planetary Science	932	31,891	453	32,387	5,216
Other	13	463	1,376	0	34
<b>Total</b>	<b>7,291</b>	<b>32,535</b>	<b>12,394</b>	<b>43,899</b>	<b>29,567</b>

Table 10: NSSDCA off-line data dissemination statistics, 2010-2014. Note that for years prior to 2012 that “items” were reported, which included not only files, but also hard media (e.g., CD-ROMs, etc.). For the years 2012 and after data distribution has been of electronic files.

## Additional Services and Activities

In addition to its archive of scientific data, NSSDC offers a number of additional services that are described below.

### NSSDCA Information Management System (NIMS)

The NSSDCA Information Management System (NIMS) encompasses most of the separate databases that have been used to track data and information through the years. Table 11 summarizes pertinent statistics for NIMS for 2014. NIMS identifies virtually all launched spacecraft, the experiments carried by many of these spacecraft, and data collections from these spacecraft, primarily as archived at NSSDCA. This portion of the database is the source of information for many of NSSDCA’s web pages. The NSSDCA 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 also 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 do not exactly predict the totals from year to year.

Partition	Records	Records Added
Spacecraft	7,238	175
Experiment	5,474	20
Data Collection	6,598	130
<b>Total</b>	<b>19,310</b>	<b>325</b>

Table 11: NIMS database statistics for 2014. Also of note: Spacecraft with experiments (1,093); Experiments with collections (1,670); Data collections associated with only spacecraft, not experiments (817); and, Data collections not associated with spacecraft/experiments (535).

### Modifications to the NSSDCA Information Model to Support PDS4

With the advent of the PDS4 Data Model, an opportunity arose to rethink the way in which PDS data were processed and stored by NSSDCA. The PDS4 Data Model largely divorces itself from the idea of data stored on media and

introduces the concept of products. Each product in PDS4 consists of a single descriptor file (in XML) and zero or more files associated with the descriptor. A logical identifier (LID) and a version (VID) are assigned to each product such that it is uniquely defined within the PDS domain. In the PDS4 model there is no requirement for related data items to be co-located on the same medium (e.g., a disk) or even co-located on the same server or within the same domain.

With this model in mind, it became clear that the old definition by which PDS3 data ingest took place (defining a transfer object as a volume) was impractical. The data as organized by the new model required a new definition of the transfer object. This resulted in the notion that, for PDS4 data, each product (or LIDVID) is the logical definition for the archival object at NSSDCA.

Shown in Figure 2 is a high-level model of how PDS3 data ingest currently takes place. Each of the PDS discipline/data nodes notifies NSSDCA directly of their desire to submit a data volume to the deep archive. This results in several pre-processing steps in which NSSDCA identifies whether or not the data have been previously submitted and if it is an extension to an existing data collection. A configuration file for the manifest generation software (provided by NSSDCA) is given back to the submitting node so that they can generate a delivery manifest. The node then delivers (via a number of mechanisms) the data volume(s) and manifest(s) to NSSDCA for archive. These are processed into AIPs and, once processing has been

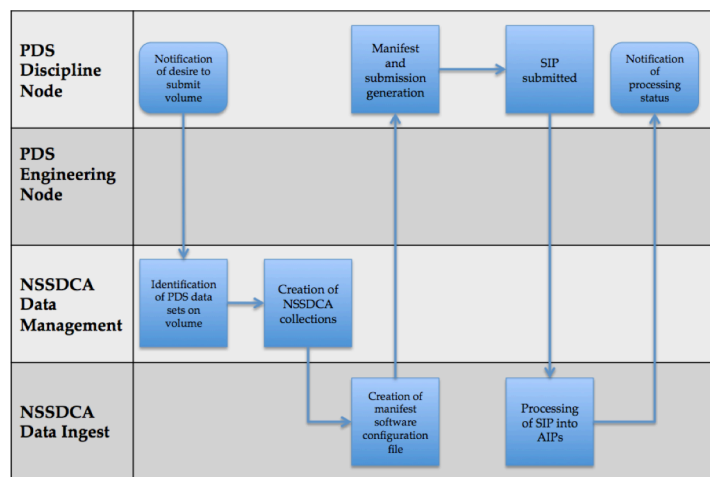


Figure 2: This diagram illustrates the process by which PDS3 data volumes are submitted to and processed by the NSSDCA.

completed, notification is sent back to the node that the data have been archived.

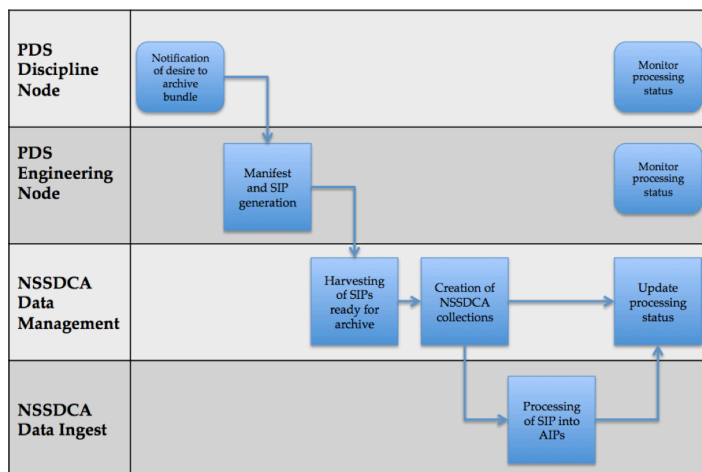


Figure 3: This diagram illustrates the process by which PDS4 data are expected to be submitted to and processed by the NSSDCA.

Figure 3 shows how the ingest model for PDS4 data is likely to take place. Notification of a data submission and generation of a delivery manifest and the SIP are now intended to go through the PDS Engineering node. NSSDCA will periodically query the registry at the Engineering node to ascertain the presence of new submissions and then proceed to process any data so found automatically, generating new collections as required. As a data submission is processed, the status of the individual transfer objects in the submission will be updated continuously.

## SPASE Data Model and Virtual Observatories (VOs)

NSSDCA staff has 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. NSSDCA staff has been involved with it since it was proposed in 1998, providing coordination of the development of the SPASE Data Model.

The latest version (2.2.3) of the SPASE Data Model was released 22 May 2014. 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/>.

## Data Recovery Support

NSSDCA supported efforts of two external organizations to recover data from obsolete magnetic tapes.

The Smithsonian Institution requested the retrieval of radargrams from Arecibo Observatory from ten 8mm tapes, all written in 1988. Nearly all of the imagery was successfully retrieved and made available on NSSDCA's FTP site.

A GSFC request to retrieve data from 9-track system backup tapes created by the Cosmic Background Explorer (COBE) mission was also received. The data of interest were processed time-ordered data from the Diffuse Infrared Background Experiment (DIRBE) instrument from the period after the COBE satellite's cryogen supply was exhausted. The data were generated in 1996. NSSDCA read over 200 tapes and concluded the requested data were not present. The requestor is seeking additional COBE tapes that could possibly contain the data. Should the requestor's search be successful, NSSDCA will resume efforts to locate and retrieve the data.

## Consultative Committee for Space Science Data Systems (CCSDS)

NSSDCA 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 NSSDCA participation in these committees has been reduced and transferred to an outside consultant. The remaining effort is concentrated within the Data Archive Ingest group, which has been finalizing the Producer Archive Interface Specification (PAIS). PAIS 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.

## Solar System Exploration and Integration

NSSDCA staff supported a number of solar system celestial events including the arrival of Comets ISON and PanSTARRS, Comet Siding Spring's close approach to Mars, as well as mission events for Rosetta/67P and New Horizons/Pluto. Support was provided through observing challenges, educator professional development workshops, public talks, K-12 activity development, social media postings, conference presentations, and workshop coordination.

## Other Staff Projects

NSSDCA staff are also involved in the coordination of the Radio Jove program that 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. Nearly 2000 kits have been distributed by the project worldwide and amateur radio observers continue to contribute the results of their observations through email to a general distribution list as well as a general archive of observations. A joint conference was held with the Society of Amateur Radio Astronomers (SARA) at the National Radio Astronomy Observatory (NRAO) in Green Bank, West Virginia, from June 29-July 04, 2014 and resulted in a record attendance for the annual SARA conference. Lectures by two Nobel Laureates were among the highlights of the meeting. Radio Jove kit distributor, Dr. Charles Higgins of Middle Tennessee State University, spent a six-month sabbatical at Goddard studying Jupiter radio emission observations by the Long Wavelength Array instrument and how they can be complemented by Radio Jove amateur observations.

## Glossary

ACE	Advanced Composition Explorer
AE	Atmospheric Explorer
AIP	Archive Information Package
AMPTE	Active Magnetospheric Particle Tracer Explorers
ARCAD	Arc Aurorale et Densite
CCSDS	Consultative Committee for Space Data Systems
COBE	Cosmic Background Explorer
CRRES	Chemical Release and Radiation Effects Satellite
DAI	Data Archive Ingest
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
IRAS	Infrared Astronomical Satellite
ISEE	International Sun-Earth Explorer
ISIS	International Satellite for Ionosphere Studies
ISO	International Organization for Standardization
ISON	International Scientific Optical Network
LADEE	Lunar Atmosphere and Dust Environment Explorer
LID	Logical Identifier
LIDVID	Logical Identifier/Version Identifier
LTO	Linear Tape-Open [i.e. open standard]
MAGSAT	MAGnetic field SATellite

M-O	Magneto-optic
MOU	Memorandum of Understanding
NIMS	NSSDCA Information Management System
NMC	NSSDCA Master Catalog
NRAO	National Radio Astronomy Observatory
NSSDC	National Space Science Data Center
NSSDCA	NASA Space Science Data Coordinated Archive
OAIS	Open Archival Information System
OGO	Orbiting Geophysical Observatories
PAIS	Producer Archive Interface Specification
PanSTARRS	Panoramic Survey Telescope and Rapid Response System
PDS	Planetary Data System
RAC	Repository Audit and Certification
RHESSI	Reuven Ramaty High Energy Solar Spectroscopic Imager
SAMPEX	Solar Anomalous and Magnetospheric Particle Explorer
SARA	Society of Amateur Radio Astronomers
SIP	Submission Information Package
SDLT	Super DLT (see above)
SNOE	Student Nitrogen Oxide Explorer
SPASE	Space Physics Archive Search & Extract
SPDF	Space Physics Data Facility
SWAS	Submillimeter Wave Astronomy Satellite
TB	Terabyte
TWINS	Two Wide-angle Imaging Neutral-atom Spectrometers
VID	Version Identifier
VO	Virtual Observatory
WG	Working Group
WORM	Write-Once, Read-Many