**Process for Submission of PDS3 Data to NSSDCA**

**V2.2**

**2022-02-01**

Table of Contents

[1 Introduction 2](#_Toc42260027)

[1.1 Purpose 2](#_Toc42260028)

[2 NSSDCA Provided Software 2](#_Toc42260029)

[3 Assigning the NSSDCA Data Collection ID 2](#_Toc42260030)

[4 Transfer Objects and Submission Information Packages 3](#_Toc42260031)

[5 Data Delivery Process 3](#_Toc42260032)

[5.1 Delivery Process Overview 3](#_Toc42260033)

[5.2 Generate SIP manifests 3](#_Toc42260034)

[5.3 Create delivery packages 3](#_Toc42260035)

[5.4 Validate delivery packages 4](#_Toc42260036)

[5.5 Assign a delivery identifier and generate the delivery manifest 5](#_Toc42260037)

[5.6 Generate the delivery manifest 5](#_Toc42260038)

[5.7 Notify NSSDCA of PDS data delivery 5](#_Toc42260039)

[6 Notifications 6](#_Toc42260040)

[6.1 PDS data delivery 6](#_Toc42260041)

[6.2 NSSDCA delivery received notification 6](#_Toc42260042)

[6.3 NSSDCA archiving complete provider notification 6](#_Toc42260043)

[7 Online Status Check 6](#_Toc42260044)

[8 Data Retrieval 7](#_Toc42260045)

[9 Glossary 7](#_Toc42260046)

# Introduction

## Purpose

This document describes the standard process for the delivery of PDS3 data from the Planetary Data System (PDS) to the NASA Space Science Data Coordinated Archive (NSSDCA).

# NSSDCA Provided Software

NSSDCA provides software to generate the manifest of the Submission Information Package (SIP) for a PDS3 data volume and software to perform data integrity checks. The *xman* software is run at the data provider site to create an XML manifest document that lists the contents of the SIP with information required to identify the data and verify successful transfer of the data. The *xcheck* is run at the data provider site to validate the contents of the delivered SIP against the information in the provider-generated manifest.

*Xman* and *xcheck* are available for Windows machines as executable binaries in a standard installation kit. For Linux and Macintosh operating systems a distribution kit is provided with a document describing how to build the utilities. Basic operating instructions are contained in the supplied Man pages. Additional installation and configuration information is supplied in the Xman and Xcheck Operations Guide (PDS Version):

<https://nssdc.gsfc.nasa.gov/pds/PDS3_Xman_operations_guide_20210716.docx>

# Assigning the NSSDCA Data Collection ID

NSSDCA groups data into collections for ingest and access purposes. A data volume must be assigned an NSSDCA Data Collection ID before it can be delivered to NSSDCA.

To begin the delivery and ingest process, a PDS node submitting data must first use the PDS Submission Interface (PSI) to enter information required by the NSSDCA (PDS volume URL, node, and transfer contact):

<https://nssdc.gsfc.nasa.gov/psi/>

After the PDS node “submits” the volumes via the PSI, the NSSDCA will register the volumes, assign new or existing NSSDCA Data Collection IDs, and then send notification that an updated id-map.tsv file, also known as the *xman* configuration file, is ready at:

<https://nssdc.gsfc.nasa.gov/pds/configuration_files/id-map.tsv>

Once the updated *xman* configuration file has been downloaded and installed, the PDS node can run *xman* to generate manifests supporting delivery of their PDS volumes.

After the NSSDCA Data Collection ID is assigned, the entry can be viewed using the online NSSDCA Master Catalog:

https://nssdc.gsfc.nasa.gov/nmc/.

# Transfer Objects and Submission Information Packages

NSSDCA archives transfer objects, the transfer object is defined as the PDS volume. NSSDCA data ingest, data dissemination, and data replacements occur at the transfer object level. Transfer objects are grouped into Submission Information Packages (SIPs) for delivery to NSSDCA. In the PDS case a SIP contains a single transfer object (PDS volume).

# Data Delivery Process

## Delivery Process Overview

There are a number of distinct steps in the delivery process:

1. Generate a SIP manifest for each volume
2. Create the delivery package
3. Validate delivery package
4. Assign a delivery identifier
5. Generate the delivery manifest
6. Notify NSSDCA and transmit the delivery

Details of the process steps follow.

## Generate SIP manifests

The delivering PDS node uses the *xman* software to generate an XML manifest for each PDS volume being delivered. A SIP manifest lists the directory structure of an on-line PDS volume and provides fixity information for each file. NSSDCA uses SIP manifests to perform a network transfer of the volumes from the PDS node to NSSDCA. Once a volume has been delivered to NSSDCA, the manifest is used to validate the volume. The manifest is also an input to NSSDCA’s archival process. The manifests will be placed in the delivery package when that package is prepared.

## Create delivery packages

Fully assembled PDS data delivery packages contain one or more SIPs (PDS volumes), a SIP manifest for each SIP in the delivery, and a delivery manifest.

The directory structure consists of a top level directory which contains a subdirectory for each SIP being delivered, each of these sub-directories being the root level of a PDS volume. The top level directory of the delivery also contains a ‘Manifest’ subdirectory; this is where the delivery manifest is located. SIP manifests may be located there as well.

A sample delivery directory structure:

./

vg\_2001

Manifest

./vg\_2001

aareadme.txt

catalog

errata.txt

jupiter

neptune

saturn

uranus

voldesc.cat

./Manifest

PDS\_ATM\_2008-12-15T144000.tab *[delivery manifest]*

Sip-manifest-VG\_2001.xml *[SIP manifest]*

Delivery mechanism supported:

* Network delivery. NSSDCA transfers data identified in SIP manifests from data nodes to NSSDCA.

Requirements for network delivery are:

* + The delivering node must be running an HTTP or FTP server capable of serving the files in the delivery.
	+ The data delivery must be posted in a directory accessible to the NSSDCA
	+ Data will be in a directory tree.
	+ The delivering node will provide NSSDCA a delivery manifest to retrieve the delivery (see section 5.6). If needed, user name and password will be provided.
	+ The delivering node will provide NSSDCA the sip\_manifest(s) or the URLs to the sip\_manifest(s) (see section 5.6).
	+ The data delivery must remain available to NSSDCA until the providing data node has been notified that archiving of the data is complete (see section 6.3).

## Validate delivery packages

Volumes ready for delivery shall be validated by the submitting node using *xcheck* software.

## Assign a delivery identifier and generate the delivery manifest

Each delivery is assigned a unique identifier in the form *CuratingNodeId*\_*SubNodeId*\_*yy-mm-dd*T*hhmmssc* (e.g., PDS\_RNG\_2010-10-29T145000).

## Generate the delivery manifest

Delivery manifests are tab delimited files providing the locations of each SIP and its related SIP manifest in the delivery package. Delivery manifest file names are the delivery identifier with extension “.txt”.

These locations are paths specifications relative to the top of the SIP delivery package:

URL to SIP\_directory*<tab>URL to* SIP\_manifest

The following example is based on delivery PDS\_ATM\_2013-08-20T090000

Manifest file name: PDS\_ATM\_2013-08-20T090000.txt

Manifest contents:

https://pds-atmospheres.nmsu.edu/PDS/data/jnogrv\_0001/ http://atmos.nmsu.edu/~jojohnso/Sip-manifest-JNOGRV\_0001-VERSION\_3.xml

https://pds-atmospheres.nmsu.edu/PDS/data/couvis\_0007/ http://atmos.nmsu.edu/~jojohnso/Sip-manifest-COUVIS\_0007-VERSION\_3.xml

https://pds-atmospheres.nmsu.edu/PDS/data/couvis\_0009/ http://atmos.nmsu.edu/~jojohnso/Sip-manifest-COUVIS\_0009-VERSION\_3.xml

https://pds-atmospheres.nmsu.edu/PDS/data/couvis\_0010/ http://atmos.nmsu.edu/~jojohnso/Sip-manifest-COUVIS\_0010-VERSION\_3.xml

The delivery manifest is placed in the ./Manifest directory of the delivery package or sent via email to NSSDCA (see section 5.7).

## Notify NSSDCA of PDS data delivery

The delivering PDS node notifies the NSSDCA operations coordinators of the delivery (see section 6.1). The node provides NSSDCA via email:

* + the delivery’s identifier (e.g., IMAGING\_USGS\_2008-06-23T120000)
	+ the top level URL for the network delivery
	+ delivery manifest “.txt” file

All notifications of a delivery should be sent to NSSDCA’s operation coordinators:

Larry Martin

NASA/Goddard Space Flight Center

Code 690.1 Bldg 28 Room N189

Greenbelt, MD 20771

lawrence.e.martin@nasa.gov

Barbara Pope

NASA/Goddard Space Flight Center

Code 690.1 Bldg 28 Room N189

Greenbelt, MD 20771

barbara.pope-1@nasa.gov

Upon receipt of a delivery notification, the NSSDCA will begin downloading and ingesting data from the delivering PDS node.

# Notifications

## PDS data delivery

The delivering PDS node notifies the NSSDCA operation coordinators of the delivery. The node provides NSSDCA the information identified in section 5.7.

## NSSDCA delivery received notification

NSSDCA informs the delivering PDS data node and the PDS Engineering node when network data deliveries are downloaded. The message will provide the Delivery\_Identifier and all SIP IDs of the received data.

Data node recipient: sender of delivery notification

Engineering node recipient: Vivian Tang

vivian.tang@jpl.nasa.gov>

## NSSDCA archiving complete provider notification

NSSDCA informs the delivering PDS data node when all PDS volumes in deliveries have been archived at NSSDCA. The message will provide the delivery ID, the SIP IDs and the IDs of all transfer objects packaged into AIPs and written to archive media. The NSSDCA may send an additional notification of accepting archival responsibility, this notification is letting the delivery node know that the NSSDCA has processed the data and it is waiting to be written to archival media and will notify the delivery node when the data has been written to archival media.

Data node recipient: sender of delivery notification

Engineering node recipient: Vivian Tang

vivian.tang@jpl.nasa.gov>

# Online Status Check

In addition to the notifications listed above, the delivering PDS data node can use the PSI to check the status of a submission (e.g., NSSDCA ID assigned, failed submission) and the status of a delivery (e.g., ingested, archived):

<https://nssdc.gsfc.nasa.gov/psi/>

Select the PDS3 Reporting tab in the left navigation bar then enter one or more of the available criteria for searching, such as volume ID, data set ID, or Provider Site.

# Data Retrieval

Upon request, NSSDCA will retrieve and return data to the providing PDS node. The data returned will be the transfer objects submitted to NSSDCA, i.e., whole PDS volumes. The providing node should contact the NSSDCA acquisition scientist or the NSSDCA Coordinated Request and User Support Office (CRUSO). Contact information for CRUSO is available at:

<https://nssdc.gsfc.nasa.gov/about/about_cruso.html>

NSSDCA can retrieve PDS volumes on the basis of:

* PDS Volume ID
* SIP ID – from element <SIPID> of the SIP manifest
* Delivery ID (could include multiple volumes)
* NSSDCA collection ID (could include multiple volumes)
* Transfer Object ID – from element <TransferObjectID> of the SIP manifest

The NSSCA will stage the retrieved data online for network transfer. CRUSO will email instructions for downloading to the PDS node.

# Glossary

**Submission Information Package (SIP)**

A digital information package delivered to the NSSDCA.

**Transfer Object**

A digital information object within a SIP to be preserved in an AIP. Transfer objects are the basic unit archiving unit. Each transfer object will be preserved in a single AIP and any data replacements will occur as replacements complete transfer objects. For PDS3, a transfer object contains the directory structure and all files from a single PDS volume.

**SIP Manifest**

For PDS3, a manifest of a SIP in XML format included in data submissions to NSSDCA.

**Delivery Manifest**

For PDS3, a tab-delimited file providing the locations of each SIP and its related SIP manifest in the delivery package.