

ESAC

European Space Astronomy Centre
Camino Bajo del Castillo s/n
Urb. Villafranca del Castillo
28692 Villanueva de la Cañada
Madrid
Spain
Tel. (34) 91 813 1100
Fax (34) 91 813 1139
www.esa.int

PSA PDS3 ARCHIVING GUIDE

Prepared by	Diego Fraga Agudo, Maud Barthelemy, Daniela Coia,
Emmanuel Grotheer	
Reference	ESDC-PSA-TN-0008
Issue / Revision	1 / 0
Date of Issue	10/12/2018
Status	ISSUED
Document Type	TN
Distribution	-

Original - <https://issues.cosmos.esa.int/socciwiki/display/PSA/ESDC-PSA-TN-0008+PSA+PDS3+Archiving+Guide>

European Space Agency
Agence spatiale européenne

APPROVAL

PSA PDS3 Archiving Guide	
Issue 1	Revision 0
Author Diego Fraga Agudo, Maud Barthelemy, Daniela Coia, Emmanuel Grotheer	Date 10-12-2018
Approved by	Date
B. Merin	18-12-2018
S. Besse	18-12-2018
I. Barbarisi	18-12-2018

CHANGE LOG

Reason for change	Issue	Revision	Date
First version	1	0	10/12/2018

CHANGE RECORD

Issue 1	Revision 0		
Reason for change	Date	Pages	Paragraph(s)
First version	10/12/2018	All	All

Table of contents:

1. Introduction	4
1.1 Purpose and Scope	4
1.2 Applicable Documents	4
1.3 Reference Documents	4
1.4 Abbreviations and Acronyms	4
2. PDS3 Guidelines	5
2.1 The standards and the Dictionary	5
2.2 Examples	5
2.3 PDS3 Labels	5
2.4 File Naming Restrictions	6
2.5 Directory Structure and Contents	6
2.5.1 Data Set	6
2.5.2 Directory Structure	7
2.5.3 Top Level Directories	8
2.5.4 Subdirectories	9
2.5.5 The AAREADME.TXT	9
2.5.6 The VOLDESC.CAT file	10
2.5.7 BROWSE Directory	10
2.5.8 CATALOG Directory	13
2.5.9 DATA Directory	16
2.5.10 DOCUMENT Directory	18
2.5.11 Other directories	19
2.6 Validation	19
3. Available Tools	20
4. Appendix A: PDS3 Catalogue templates	21
4.1 DATASET.CAT	21
4.2 INST.CAT	22
4.3 INSTHOST.CAT	23
4.4 MISSION.CAT	24
4.5 PERSONNEL.CAT	25
4.6 REFERENCE.CAT	26
4.7 SOFTWARE.CAT	26
4.8 TARGET.CAT	27

1. INTRODUCTION

1.1 Purpose and Scope

The current document aims at being a guideline to the preparation of data to be delivered to the Planetary Science Archive (PSA) by potential external data providers not familiar with the PSA.

Within this document “**external data provider**” means a provider of scientific products to the PSA not directly involved in an instrument team of any missions archived in the PSA. For providers directly involved in European Space Agency’s (ESA) missions, parts of this document are not applicable.

This document is an introduction to the PDS3 standard and in addition contains guidelines specific for the PSA archive within the PDS framework. The PDS4 standard is not covered by this document. Readers interested in PDS4 within the PSA must read AD2.

It is recommended to read first [AD3] *PSA Archiving Guide for External Data Providers* which is a very brief introduction to the PSA, the PDS standards and the life cycle of the data sets from preparation, interactions with the PSA, and public release.

1.2 Applicable Documents

Ref.	Document	Version
AD1	PDS3 Planetary Data System Standards Reference , JPL D-7669, Part 2	V 3.8, 27 (February 2009)
AD2	PSA PDS4 Archiving Guide, ESDC-PSA-TN-0002	V 2.0, 2 (August 2017) or latest
AD3	PSA Archiving Guide for External Data Providers, ESDC-PSA-TN-0009	1.0 or latest version

Table 1 Applicable documents

1.3 Reference Documents

Ref.	Document	Version
RD1	Abbreviations and Acronyms, ESDC-PSA-TN-0003	1.0 or latest version
RD2	Data User Guide Template, ESDC-PSA-TN-0008	1.0 or latest version

Table 2 Applicable documents

1.4 Abbreviations and Acronyms

See document [RD1].

2. PDS3 GUIDELINES

2.1 The standards and the Dictionary

PDS3 data sets and PDS3 products must follow the PDS3 standard and be compliant with the PDS3 dictionary (extended by the local ESA PDS3 dictionary).

The standards can be found in document [AD1] “PDS3 Planetary Data System Standards Reference, JPL D-7669, Part 2” which can be found on the PDS web page at: <https://pds.nasa.gov/tools/standards-reference.shtml> .

Among other rules the standard defines a language for the metadata. In its simplest form the metadata consists on "keywords" that take a value like in the following example:

```
MISSION_NAME = "MARS EXPRESS"
```

Keywords are in some circumstances grouped to define "objects".

Not all objects and keywords are fully defined in the mentioned document. The exact definition of the objects and keywords are in the PDS3 dictionary. The object definition includes: the components, whether they are optional or mandatory, and if there can be more than one occurrence. The keyword definition includes: the data type, whether they accept numbers, free text, or restricted text. The dictionary is available as an online search interface at <https://pds.jpl.nasa.gov/tools/dd-search/>

Finally, PSA extends the PDS dictionary by adding new keywords or new values to existing keywords. For the moment, the only way to consult the PSA dictionary is the PVV validation tool presented in section 2.6 and 3. Once installed this tool it is possible to know all information about a given keyword by typing:

```
pvv help object_or_keyword_name
```

2.2 Examples

The best way to obtain examples is to see real data sets in the PSA archive itself, which can be found at psa.esa.int. From there it is possible to browse the archive holdings by clicking on the "FTP Access" icon or link. In the FTP area it is possible to see examples, of individual data products and also full data sets, including the directory structure and accompanying documents and files.

[AD1] also contains many examples of data objects and keyword usage.

2.3 PDS3 Labels

All files in a data set must come with a “PDS3 label” with the only exceptions being those files that are placed in the EXTRAS directory, though the use of labels is still recommended for this extra content as well. The labels contain metadata about the data or document.

According to the PDS3 standard, labels can be:

- **Attached:** The PDS data product label is included at the beginning of the file. There is one label attached to each file.

- **Detached:** The PDS data product label is detached from the data and resides in a separate file that contains a pointer to the data product file. There is one detached label file for each data product file. The label file should have the same file name as its associated data product file, but with the extension ".LBL".
- **Combined Detached:** A single PDS detached label file is used to describe the contents of more than one data product file. This approach is not recommended by PSA.

The recommendation by PSA is to use individual detached labels for all data files. For documentation files attached or detached labels can be used.

PDS3 labels are written in ODL (Object Description Language) language (unlike PDS4 labels, which follow XML language). For a detailed description of the ODL language see [AD1].

For examples of real PDS3 labels look for data in the PSA archive.

To learn more about the availability of tools to help generate and validate labels, see section 3.

2.4 File Naming Restrictions

All file names, including data files, documents, or any other file, may only contain characters from the following set: the upper case alphanumeric characters (A- Z, 0-9) and the underscore ("_").

Characters from outside the English alphabet are not permitted. A dot is used to separate the file extension. The file name without the extension cannot be longer than 27 characters and the extension has to be 3 characters long (preferred) or shorter. In total the file name with extension cannot be longer than 31 characters (which includes the dot for the file extension).

2.5 Directory Structure and Contents

2.5.1 Data Set

PDS3 defines the concept of a data set as:

A data set is an aggregation of data products with a common origin, history, or application. A data set includes primary (observational) data plus the ancillary data, software, and documentation needed to understand and use the observations. Files in a data set share a unique data set name, share a unique data set identifier [...].

The data provider will have to arrange their data products and documents into one or more data sets. The data set is also the parent directory where other directories for the data, documents, etc. are located. The data set is self-contained in the sense that it includes all documentation and files needed to interpret the data. If the data is split into several data sets (e.g. by acquisition epoch or by instrument type) the documents that are applicable to all data sets must also be included in all data sets.

PDS3 also defines the concept of volume. Originally a volume was linked to physical storage, such as a CD for instance. There could be several data sets in one volume or one data set could be split into several volumes. Currently data is distributed online and thus this term has partially lost its original meaning. PSA recommends defining one and only one volume per data set.

2.5.2 Directory Structure

The following diagram describes the directory structure of a data set and the standard required files.

Top-level directory (Data Set Directory)

```

|
|- AAREADME.TXT      A quick overview of the data set and its structure
|
|- ERRATA.TXT       Description of known anomalies and errors
|                   present in this data set
|
|- VOLDESC.CAT      Data set volume description
|
|
|- [BROWSE]         Browse products (aka postcards)
| |
| | - BROWINFO.TXT  Description of the BROWSE directory contents and browser
| | information.
|- [CATALOG]
| |
| | - CATINFO.TXT   Description of files in the CATALOG
| |                 directory
| |
| | - DATASET.CAT   Description of the data set's content
| |
| |
| | - PERSONNEL.CAT Personnel involved in the data set production
| |
| |
| | - REFERENCE.CAT List of cited references (aka bibliography)
| |
| |
| | - INSTHOST.CAT  Spacecraft or host overview
| |
| |
| |

```

	- MISSION.CAT	Mission description
	- INSTRUMENT.CAT	Instrument description
	- TARGET.CAT	Target information
	- SOFTWARE.CAT	Software description
	- [DATA]	Directory containing the data products
	- [CALIB]	Directory containing calibration files
	- CALINFO.TXT	Description of files in the CALIB directory
	- [DOCUMENT]	Directory containing document files
	- DOCINFO.TXT	Description of files in the DOCUMENT directory

2.5.3 Top Level Directories

Table 3 describes the data set directories' content, for top level directories only. Some of these directories are mandatory and others are optional (see column comments).

Directory name	Content according to PDS definition	Comments
BROWSE	Reduced resolution versions of data products.	Optional and highly encouraged by PSA. The images will be used as small browse images (postcards) in the PSA user interface.
CALIB	Calibration files used in the original processing of the data, or needed to use the data.	Optional.
CATALOG	PDS catalog files.	Mandatory by PDS standard.

DATA	Contains the data products. The DATA directory may contain subdirectories to improve data product classification (e.g., year, month, type, etc.).	Mandatory by PDS standard.
DOCUMENT	Documentation, supplementary and ancillary information to assist in understanding and using the data products.	Mandatory by PDS standard.
EXTRAS	“Value added” elements included by the data preparer, but outside the scope of the PDS archive requirements.	Optional. Files in this directory do not have to conform to any specific PDS rule.
GAZETTEER	Tables of information about the geological features of a target.	Optional.
GEOMETRY	Files describing the observational geometry (e.g., SEDRs, SPICE kernels).	Optional.
INDEX	Indices to assist in locating data of interest.	Mandatory by PDS standard. The PSA's PVV tool can be used to generate the INDEX, and if applicable also the BROWSE_INDEX files (.TAB and its matching label with .LBL).
LABEL	“Include” files with FMT extension that describe specific aspects of the data format and organization.	Optional. The FMT files allow the provider to avoid repetition in the labels. See [AD1] for further information.
SOFTWARE	Utilities, application programs, or subroutines used to access or process the data.	Optional. It is not recommended to provide software in the SOFTWARE directory because it cannot be maintained. Current policy for software both at PDS and PSA is to include it either in the EXTRAS or in the DOCUMENTATION directory.

Table 3 Top level directories within a data set

2.5.4 Subdirectories

Within each top-level directory there can be other directories to better arrange the different files. The structure and naming of those directories is free. For example, if there are many files in the DATA directory, they are commonly grouped in subdirectories either by product type, instrument or by acquisition time intervals (e.g. one subdirectory per month).

2.5.5 The AAREADME.TXT

This file must describe the directory structure and the content of the data set. It also includes the provider's contact details.

The AAREADME.TXT file must be located in the root directory, i.e. inside the data set directory and outside any other directory.

A table of expected contents and an example can be found in [AD1] appendix D1.

2.5.6 The VOLDESC.CAT file

This file contains the VOLUME object, which provides a high-level description of the contents of an archive volume.

The VOLDESC.CAT file must be located in the root directory, i.e. inside the data set directory and outside any other directory.

Examples of this file can be found in the data sets of the PSA itself.

2.5.7 BROWSE Directory

2.5.7.1 General Considerations

Browse images of the data products are highly recommended by PSA, though it is not a strict requirement by PDS3 standards. When browse images are included, they must be in a directory called BROWSE.

Browse images are used in the PSA user interface to show a preview of the product without the need of downloading the product. If a browse product is not defined for a given data product then a preview image cannot be shown.

Browse products are recommended not only for image data, but also for any other data types that can be represented graphically, such as spectra or time series.

The recommended formats for the browse images in PSA are JPEG and PNG. JPEG is recommended for pictures and PNG for plots. The browse image size is flexible and must ensure that enough details are visible. The minimum recommended size is 128x128 and the maximum recommended size is 1024x1024 pixels. Note that browse images in JPEG format should use the file extension “.JPG”.

Within the BROWSE directory, the subdirectories (if any) should be structured to mirror the way the DATA directory is organized (see section 2.5.4).

2.5.7.2 Browse Image Label Files

Each browse image must come with a corresponding label file. Figure 1 shows a template of the label file for browse images in PNG format with the minimum keywords required by PSA and Figure 2 shows the equivalent template for JPEG browse images. “***” is a placeholder for a text string, which can be longer than three characters.

PDS_VERSION_ID = PDS3
RECORD_TYPE = UNDEFINED

```

^PNG_DOCUMENT      = "****.PNG"
DATA_SET_ID        = "****"
PRODUCT_ID         = "****"
SOURCE_PRODUCT_ID  = "****"

OBJECT             = PNG_DOCUMENT
DOCUMENT_NAME      = "****"
PUBLICATION_DATE   = yyyy-mm-dd
DOCUMENT_TOPIC_TYPE = "BROWSE IMAGE"
INTERCHANGE_FORMAT = "BINARY"
DOCUMENT_FORMAT    = "PNG"
DESCRIPTION        = "****"
END_OBJECT         = PNG_DOCUMENT

END

```

Figure 1 PNG browse image template

```

PDS_VERSION_ID     = PDS3
RECORD_TYPE        = UNDEFINED

^JPG_DOCUMENT      = "****.JPG"
DATA_SET_ID        = "****"
PRODUCT_ID         = "****"
SOURCE_PRODUCT_ID  = "****"

OBJECT             = JPG_DOCUMENT
DOCUMENT_NAME      = "****"
PUBLICATION_DATE   = yyyy-mm-dd
DOCUMENT_TOPIC_TYPE = "BROWSE IMAGE"
INTERCHANGE_FORMAT = "BINARY"
DOCUMENT_FORMAT    = "JPG"
DESCRIPTION        = "****"
END_OBJECT         = JPG_DOCUMENT

END

```

Figure 2 JPG browse image template

Please bear in mind the following rules:

- The ^PNG/^JPG_DOCUMENT pointer must indicate the file name of the browse image, including the file extension.
- The PRODUCT_ID is a product identification assigned to the browse image product, not to the source data file.

- The SOURCE_PRODUCT_ID must match the PRODUCT_ID value given in the source data file's own label.
- The DESCRIPTION keyword allows for long form text to describe the browse product. It may be necessary to spread the text over several lines.
- For all the lines in the browse image label, the maximum line length is 80 characters, including the carriage return and line feed terminators (leaving 78 usable character spaces per line).
- The values of the keywords inside of the PNG/JPG_DOCUMENT object are descriptive of the browse image itself and not of the source data file.

Figure 3 shows a realistic example of a browse image label.

```

PDS_VERSION_ID      = PDS3
RECORD_TYPE         = UNDEFINED

^PNG_DOCUMENT       = "SPIM_oBR_07726A01_S_05_QL.PNG"
DATA_SET_ID         = "MEX-M-SPI-2-IREDR-RAWXMARS-EXT3-V2.0"
PRODUCT_ID         = "SPIM_oBR_07726A01_S_05_QL.PNG"
SOURCE_PRODUCT_ID   = "SPIM_oBR_07726A01_S_05.DAT"

OBJECT              = PNG_DOCUMENT
DOCUMENT_NAME       = "SPICAM IR EDRAW OBSERVATION QUICK LOOK"
PUBLICATION_DATE    = 2013-01-31
DOCUMENT_TOPIC_TYPE = "BROWSE IMAGE"
INTERCHANGE_FORMAT  = "BINARY"
DOCUMENT_FORMAT     = "PNG"
DESCRIPTION         = "Quick look views of Mars Express SPICAM IR EDRAW observations after early
processing of raw data"
END_OBJECT          = PNG_DOCUMENT

END

```

Figure 3 Browse image label example

If you are creating data sets linked to a specific mission there could be additional requirements. For example:

- if your data is related to Mars Express or Venus Express, it is important to know that these missions use the "Release and Revision concept". Thus, even if your data set will only ever include 1 release (and no further revisions), you must still add the RELEASE_ID and REVISION_ID keywords. For the first delivery, the values will be RELEASE_ID = 0001 and REVISION_ID = 0000. Subsequent deliveries of more data for inclusion within the same data set would require incrementing the release ID, while updates to already published data would require incrementing the revision ID.
- if your data is related to Rosetta there are no other required keywords.

The relation between source data product and browse image does not need to be one to one. It is possible to link one source data product to several browse images or several source data products to one browse image.

2.5.7.3 Linking One Source Data Product to Multiple Browse Images

In this case the only extra requirement is to add, within each browse image label, the keyword `BROWSE_USAGE_TYPE` (outside the `PNG_DOCUMENT` or `JPEG_DOCUMENT` objects). This keyword should be given either the value `PRIMARY` or `SECONDARY`, to indicate which is the primary browse image and which are the secondary browse images. There can only be one primary browse image per source data product, but there can be multiple secondary browse images.

2.5.7.4 Linking Multiple Data Products to One Browse Image

In this case the `SOURCE_PRODUCT_ID` of the browse image label will contain multiple values, as shown in the following example:

```
SOURCE_PRODUCT_ID = {
"N0346AE02_RR.TAB"
"N0346AE02_RRTH.TAB"
"N0347AE02_RR.TAB"
, "N0347AE02_RRTH.TAB" }
```

2.5.8 CATALOG Directory

The CATALOG directory contains catalogue files. The catalogue files are a set of standardized files that summarize important information relevant to the whole data set. They include information about the data set content, the instrument used to collect the data, the instrument host (e.g. spacecraft or ground facility), the mission, the personnel involved, references to published papers, software provided in the data set and targets of the data (i.e. the astronomical body observed). All the catalogue files use a set of objects and keywords to make this information readable by software so that current and future archive systems can classify this information.

Though PDS3 standards leave some flexibility on the naming convention of the catalogue files, for external data providers for the PSA, the recommended naming convention is:

```
CATINFO.TXT
DATASET.CAT
INST.CAT
INSTHOST.CAT
MISSION.CAT
PERSONNEL.CAT
REFERENCE.CAT
SOFTWARE.CAT
TARGET.CAT
```

Nevertheless, if the provided data sets are linked to a specific mission that follows a different naming convention that naming convention can also be used.

All catalogue files must start with the following keywords:

- PDS_VERSION_ID
- LABEL_REVISION_NOTE
- RECORD_TYPE

The following subsections will discuss one by one these catalogue files. Templates can be found in appendix A. The full definition of the objects in the catalogue can be found in the dictionary and real examples can be found in the holdings of the PSA archive.

2.5.8.1 Data Set Catalogue (DATASET.CAT)

The goal of this document is to describe the content of a data set.

It contains the following object:

- DATA_SET

An important part of the data set catalogue is the content of the DATA_SET_DESC keyword within the DATA_SET object. The value of this keyword must be a textual description of the data set. The goal of this description is to be a starting point or introduction for a new user. It must summarize the content of the data set, with what instrument it was acquired, when it was acquired and if there were any issues during acquisition, how it was processed, what type of products can be found, etc. The extent of this text might vary but typically its length is on the order of one page of text. A table of contents and/or underlined sections can be included, especially if the description is long. Additionally, this catalogue shall include recommendations on what documents to read to start using the data.

2.5.8.2 Instrument Catalogue (INST.CAT)

The goal of this document is to describe the instrument that was used to acquire the data.

This file must contain the following object:

- INSTRUMENT

The main content of this file is the value of the keyword INSTRUMENT_DESC within INSTRUMENT. This is a comprehensive description of the instrument used to obtain the data. If necessary, a reference to a document in the DOCUMENT directory covering this topic can be included here, but a summary of such a document must still be included in this catalogue.

2.5.8.3 Instrument Host Catalogue (INSTHOST.CAT)

The goal of this document is to describe the host of the instrument. "Host" is usually understood to mean the spacecraft carrying the instrument, or the ground facility hosting the instrument.

This file must contain the following object:

- INSTRUMENT_INFORMATION

The main content of this file is in the keyword INSTRUMENT_HOST_DESC. Here a comprehensive description of the host of the instrument shall be given. If necessary, a reference to a document in the DOCUMENT directory covering this topic can be added but a summary must be provided in this file.

2.5.8.4 Mission Catalogue (MISSION.CAT)

The goal of this document is to describe the mission in which the data was obtained. For data obtained from a spacecraft, the term "mission" is understood to indicate the whole project, including not only the spacecraft but also the ground facilities as well as characteristics such as orbit descriptions or phases during the mission. For ground observations, the concept of mission must be interpreted as the context in which the data was collected, such as for example the case in which they belong to a broader research program.

This file must contain the following object:

- MISSION_INFORMATION

2.5.8.5 Personnel Catalogue (PERSONNEL.CAT)

The goal of this document is to list the people involved in the production of the data and their contact information.

This file must contain at least one instance of the following object

- PERSONNEL

2.5.8.6 Reference Catalogue (REFERENCE.CAT)

The goal of this document is to list peer reviewed publications related to the data set or to similar data, such as other data sets acquired with the same instrument.

This file must contain at least one instance of the following object:

- REFERENCE

2.5.8.7 Software Catalogue (SOFTWARE.CAT)

The goal of this document is to list and describe the software provided in the data set if any.

This file must contain the following object:

- SOFTWARE

It is acceptable to set some of the keywords to N/A (not applicable) as illustrated in the template in appendix A. As explained, PDS and PSA no longer support the archiving of software because it cannot be maintained. However, this catalog file is required for legacy reasons: the catalog was required when software could be archived and the PSA architecture expects its presence.

2.5.8.8 Target Catalogue (TARGET.CAT)

The goal of this document is to list the target(s) or astronomical body or bodies that are linked to the data set.

This file must contain at least one instance of the following object:

- TARGET

2.5.9 DATA Directory

2.5.9.1 Contents of the Data Directory

The DATA directory contains data products or more precisely in PDS terminology “Observational Products”.

2.5.9.2 Format of PDS3 Data Products

An observational product is composed of a label file and a data file. It is possible to merge both files in one but this is discouraged. The label file and the data file must have the same name with different extensions, and the extension of the label must be “LBL”.

PDS3 is not strictly speaking a data format. Any data file in any format can potentially be a PDS3 data product as long as the format can be declared unambiguously using the PDS3 objects, keywords, and rules and as long as other required information is provided in the label. Simple formats containing for example ASCII fixed-width tables, comma separated tables, or multidimensional arrays will be very likely describable using PDS3 and therefore can become a PDS3 product by just adding its label file.

Both ASCII data and binary data are supported although ASCII data is preferred due to its easier readability despite its larger size.

For example, in PDS3 images can be stored as simple binary two dimensional arrays and its label file will have to define its exact format indicating the number of columns, rows and bits per pixel. The object to use is called “IMAGE” and the details on how to define it can be found in the dictionary.

The widely spread FITS (Flexible Image Transport System) format used in astronomy and other fields can be a PDS3 compliant product by just adding a label file defining its internal structure. Examples can be found in the PSA archive by the ALICE instrument of the International Rosetta Mission.

Another common format for the data are ASCII tables with fixed length for each column of the table. This is described using the TABLE object, which holds several COLUMN objects (one for each column in the table).

There are many other objects available and many options for the objects by using different keywords. Objects can be combined in one single file and a relatively complex structure can be created. For a full explanation see [AD1] and the dictionary.

2.5.9.3 Label of a Data Product

All data products must come with a label (metadata following PDS3 specification). PSA recommends the use of detached labels, i.e. the label is in its own file and not part of the same file as the data.

Labels must contain at least the keywords shown in Figure 4. The label file might include other keywords to document the data. For further information on the expected content of each keyword see the PDS3 dictionary.

PDS_VERSION_ID
 LABEL_REVISION_NOTE

DATA_SET_ID
 PRODUCT_ID
 INSTRUMENT_HOST_NAME
 INSTRUMENT_NAME
 TARGET_NAME
 START_TIME
 STOP_TIME
 SPACECRAFT_CLOCK_START_COUNT *
 SPACECRAFT_CLOCK_STOP_COUNT *
 PRODUCT_CREATION_TIME

* Data acquired from ground shall not include keywords SPACECRAFT_CLOCK_START_COUNT and SPACECRAFT_CLOCK_STOP_COUNT

Figure 4 Label of a data product

In addition, objects shall be defined to declare the format of the data. See the PDS3 standards [AD1] and the dictionary for further information.

2.5.9.4 Approach to Generate PDS3 Compliant Data Products

To our knowledge there is no generic writer software applicable to all types of PDS3 data products. The data provider is likely to face one of two situations:

(1) The data provider does not yet have the data files and can decide the format:

It is recommended to get familiar with the PDS3 objects by reading [AD1] and by looking at examples of similar data found in the PSA archive. Then the provider shall design a data structure that is as simple as possible and which can be described with PDS3 objects. Finally, the data file has to be generated according to the designed format, with a label file describing said format, and including at least the keywords described in section 2.5.9.3. This usually involves programming some code to generate the data and label files.

(2) The data provider already has the data files in a certain format:

One option here is not to reformat the data and simply describe the format in the label file. This is possible in many, but not all, cases. Alternatively, if the format is not describable via a PDS3 label or if it is very complex to find the right representation of the format, it will be necessary to reformat the data following the same approach as in option (1) above.

As mentioned above, usually the generation of the products requires programming. An alternative way to produce the PDS3 labels, when there are not many needed, is to edit them manually with any text editor. In this case it can be useful to prepare a template and then fill in the values that are different for the individual files.

For validation of the products see section 2.6.

2.5.10 DOCUMENT Directory

All relevant documentation to use and understand the data and its context shall be placed in this folder.

The two accepted formats for the documents are:

- Plain text files: These files must be encoded with 7 bit ASCII. This means that non-English characters such as for example á, è, ö etc. are not permitted. Line endings must be indicated in a Windows-like format, i.e. by using a carriage return followed by a line feed. This ensures proper visualization in Windows, Linux and other platforms. The file extension must be ASC if the label is detached and TXT if the label is attached.
- PDF/A files: PDF files are also accepted as long as they are compliant with the PDF/A specifications. PDF/A files, unlike normal PDFs, are self-contained and designed for long-term readability. Some text editors, such as recent versions of Word, have the option to save a document as an PDF/A compliant document.

All documents have to be accompanied by a label.

The producer can include all documentation that they consider relevant, including documents describing the calibration or processing method, peer reviewed scientific publications, etc. The goal is that the data shall be understandable and usable without requiring external information. The producer must bear in mind that this should still be possible in the long term (for example more than 50 years from now) when other external sources of information might no longer be available.

Some data providers might be familiar with the so-called EAICD document that comes with most data sets in the PSA. For external data providers it is **not** required to provide this document and a partially equivalent document must be produced instead: a Data User Guide.

Data User Guide

All external data providers shall include a Data User Guide document to describe the data set, its contents, and production. Where needed, the Data User Guide can reference other documents in the DOCUMENT folder if needed instead of repeating information.

A template with guidelines on the expected content of this document is available [RD2].

2.5.11 Other directories

Directories EXTRAS, GAZETTEER, GEOMETRY and LABEL are not required but the provider is invited to use them when applicable. For further information about these directories and their content see the standards [AD1].

The INDEX directory and its content will be automatically generated by ESA if it is not provided, though the contents for this directory can also be created by the provider using the PVV tool.

2.6 Validation

There are two types of validation: a technical one and a scientific one:

- **Technical validation:** The goal is to guarantee that the PDS standards and the guidelines given in this document are followed. Prior to submitting the data set to the PSA, it is mandatory for the data provider to successfully run (i.e. no error messages are produced; some warning messages are acceptable) the validation tool PVV (see chapter 3). This tool checks a number of rules of the PDS standards and consults the PSA dictionary to guarantee a minimum level of quality, including the values assigned to the PDS3 keywords in the label files. Note that PVV is not exhaustive in its checks, so care must be put in the production of the data set and its data products. PSA will also review the data set and will ask for corrections if necessary.
- **Scientific validation:** In order to provide confidence in the scientific quality of the data, it is highly recommended to add references to peer reviewed papers that use or describe the data. Peer-reviewed publications about the calibration or processing applied to the data are of special interest here. If there are no publications available then PSA might consider, depending on the available resources, to perform a peer review of the data set involving recognised experts in the specific field.

3. AVAILABLE TOOLS

There are a number of tools written in different languages that are useful for writing, reading, validating or visualizing PDS3 products or PDS3 data sets. Some are standalone programs whereas others are libraries to be used by programmers.

The PSA information about available tools (including PVV) can be found at:

<https://www.cosmos.esa.int/web/psa/tools>

The PDS information about available tools can be found at:

<https://pds.nasa.gov/tools/tool-registry/>

Finally SBN (the Small Bodies Node which is part of PDS) maintains a webpage with some tools designed by them. These tools are, unless otherwise specified in their page or documentation, applicable to any planetary data and not only to small bodies:

<https://pdssbn.astro.umd.edu/tools/software.shtml>

An important tool provided by PSA is PVV (PSA Volume Verifier). This is a command-line software used for validation of PDS3 datasets. All PDS3 data sets must be verified with this tool before submission to the PSA.

4. APPENDIX A: PDS3 CATALOGUE TEMPLATES

This appendix shows basic templates for the catalogue files with the minimum objects and keywords. Other keywords can be added but the ones shown here are the minimum that all catalogue files must include.

To see the description of each object or keyword and its expected or allowed values see the dictionary.

Some objects admit more than one occurrence in the file (for example, several PERSONNEL objects in PERSONNEL.CAT if there is more than one person listed in the file). To know which objects admit several occurrences see the dictionary.

Real examples can be found in the PSA archive itself at <ftp://psa.esac.esa.int/pub/mirror/>

4.1 DATASET.CAT

PDS_VERSION_ID = PDS3
 LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description (optional)"
 RECORD_TYPE = STREAM

OBJECT = DATA_SET
 DATA_SET_ID = "XXXXX"

OBJECT = DATA_SET_INFORMATION

DATA_SET_NAME = "XXXXX"

DATA_SET_TERSE_DESC = "XXXXX"

ABSTRACT_DESC = "XXXXX"

CITATION_DESC = "XXXXX"

DATA_SET_COLLECTION_MEMBER_FLG = "N"
 DETAILED_CATALOG_FLAG = "N"
 DATA_OBJECT_TYPE = "XXXXX"
 DATA_SET_RELEASE_DATE = yyyy-mm-dd
 PRODUCER_FULL_NAME = "XXXXX"

START_TIME = yyyy-mm-ddThh:mm:ss.xxx

STOP_TIME = yyyy-mm-ddThh:mm:ss.xxx

```
DATA_SET_DESC = "  
XXXXXXXXXXXXXX  
"
```

```
CONFIDENCE_LEVEL_NOTE = "  
XXXXXXXXXXXXXX  
"
```

```
END_OBJECT = DATA_SET_INFORMATION
```

```
OBJECT      = DATA_SET_MISSION  
MISSION_NAME = "XXXXX"  
END_OBJECT = DATA_SET_MISSION
```

```
OBJECT      = DATA_SET_TARGET  
TARGET_NAME = "XXXXX"  
END_OBJECT = DATA_SET_TARGET
```

```
OBJECT      = DATA_SET_HOST  
INSTRUMENT_HOST_ID = "XXXXX"  
INSTRUMENT_ID      = "XXXXX"  
END_OBJECT = DATA_SET_HOST
```

```
OBJECT      = DATA_SET_REFERENCE_INFORMATION  
REFERENCE_KEY_ID = "XXXXX"  
END_OBJECT =  
DATA_SET_REFERENCE_INFORMATION
```

```
END_OBJECT = DATA_SET
```

```
END
```

4.2 INST.CAT

```
PDS_VERSION_ID      = PDS3  
LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description  
(optional)"  
RECORD_TYPE         = STREAM
```

```

OBJECT          = INSTRUMENT
INSTRUMENT_HOST_ID = XXXXX
INSTRUMENT_ID     = XXXXX

OBJECT          = INSTRUMENT_INFORMATION

INSTRUMENT_NAME  = "XXXXXX"

INSTRUMENT_TYPE  = XXXXX
INSTRUMENT_DESC  = "
XXXXXXXX
"

END_OBJECT      = INSTRUMENT_INFORMATION

OBJECT          = INSTRUMENT_REFERENCE_INFO
REFERENCE_KEY_ID = "XXXXXX"
END_OBJECT      = INSTRUMENT_REFERENCE_INFO

END_OBJECT      = INSTRUMENT

END

```

4.3 INSTHOST.CAT

```

PDS_VERSION_ID      = PDS3
LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description (optional)"
RECORD_TYPE         = STREAM

```

```

OBJECT          = INSTRUMENT_HOST
INSTRUMENT_HOST_ID = XXXXX

OBJECT          = INSTRUMENT_HOST_INFORMATION
INSTRUMENT_HOST_NAME = "XXXXXX"
INSTRUMENT_HOST_TYPE = XXXXXX
INSTRUMENT_HOST_DESC = "XXXXXX"
END_OBJECT      = INSTRUMENT_HOST_INFORMATION

OBJECT          = INSTRUMENT_HOST_REFERENCE_INFO
REFERENCE_KEY_ID = "XXXXXX"

```

END_OBJECT = INSTRUMENT_HOST_REFERENCE_INFO

END_OBJECT = INSTRUMENT_HOST

END

4.4 MISSION.CAT

PDS_VERSION_ID = PDS3

LABEL_REVISION_NOTE = " yyyy-mm-dd Author Change description
(optional)"

RECORD_TYPE = STREAM

OBJECT = MISSION

MISSION_NAME = "XXXXXXX"

OBJECT = MISSION_INFORMATION

MISSION_ALIAS_NAME = XXXXX

MISSION_START_DATE = yyyy-mm-dd

MISSION_STOP_DATE = "yyyy-mm-dd"

MISSION_OBJECTIVES_SUMMARY = "

XXXXXX

"

MISSION_DESC = "

XXXXXX

"

END_OBJECT = MISSION_INFORMATION

OBJECT = MISSION_HOST

INSTRUMENT_HOST_ID = XXXXX

OBJECT = MISSION_TARGET

TARGET_NAME = "XXXXXXX"

END_OBJECT = MISSION_TARGET

END_OBJECT = MISSION_HOST

OBJECT = MISSION_REFERENCE_INFORMATION

REFERENCE_KEY_ID = "XXXXXXX"

END_OBJECT = MISSION_REFERENCE_INFORMATION

END_OBJECT = MISSION

END

4.5 PERSONNEL.CAT

PDS_VERSION_ID = PDS3

LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description (optional)"

RECORD_TYPE = STREAM

OBJECT = PERSONNEL

PDS_USER_ID = XXXXX

OBJECT = PERSONNEL_INFORMATION

FULL_NAME = "XXXXX"

LAST_NAME = "XXXXX"

TELEPHONE_NUMBER = "XXXXX"

ALTERNATE_TELEPHONE_NUMBER = "XXXXX"

FAX_NUMBER = "XXXXXX"

INSTITUTION_NAME = "XXXXX"

NODE_ID = "N/A"

PDS_AFFILIATION = "N/A"

PDS_ADDRESS_BOOK_FLAG = N

REGISTRATION_DATE = yyyy-mm-dd

ADDRESS_TEXT = "XXXXX"

END_OBJECT = PERSONNEL_INFORMATION

OBJECT = PERSONNEL_ELECTRONIC_MAIL

ELECTRONIC_MAIL_ID = "XXXXX"

ELECTRONIC_MAIL_TYPE = "INTERNET"

PREFERENCE_ID = 1

END_OBJECT = PERSONNEL_ELECTRONIC_MAIL

END_OBJECT = PERSONNEL

END

4.6 REFERENCE.CAT

```

PDS_VERSION_ID      = PDS3
LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description
(optional)"
RECORD_TYPE         = "STREAM"

OBJECT              = REFERENCE
REFERENCE_KEY_ID    = "XXXXXX"
REFERENCE_DESC      = "
XXXXXX
"
END_OBJECT          = REFERENCE

```

END

4.7 SOFTWARE.CAT

```

PDS_VERSION_ID      = PDS3
LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description
(optional)"
RECORD_TYPE         = STREAM

DESCRIPTION          = "XXXXXX"

OBJECT               = SOFTWARE
SOFTWARE_ID           = "N/A"
SOFTWARE_VERSION_ID  = "1.0"

```

```

OBJECT               = SOFTWARE_INFORMATION
DATA_FORMAT          = "N/A"
NODE_ID              = "N/A"
PDS_USER_ID          = "N/A"
REQUIRED_STORAGE_BYTES = "N/A"
SOFTWARE_DESC         = "N/A"
SOFTWARE_LICENSE_TYPE = "N/A"
SOFTWARE_NAME         = "N/A"
TECHNICAL_SUPPORT_TYPE = "N/A"
END_OBJECT           = SOFTWARE_INFORMATION

```

```

OBJECT               = SOFTWARE_ONLINE
NODE_ID              = "N/A"
ON_LINE_IDENTIFICATION = "N/A"
ON_LINE_NAME         = "N/A"

```

PLATFORM = "N/A"
 PROTOCOL_TYPE = "N/A"
 END_OBJECT = SOFTWARE_ONLINE

OBJECT = SOFTWARE_PURPOSE
 SOFTWARE_PURPOSE = DOCUMENTATION
 END_OBJECT = SOFTWARE_PURPOSE

END_OBJECT = SOFTWARE

END

4.8 TARGET.CAT

PDS_VERSION_ID = PDS3
 LABEL_REVISION_NOTE = "yyyy-mm-dd Author Change description
 (optional)"

RECORD_TYPE = STREAM

OBJECT = TARGET
 TARGET_NAME = "XXXXXX"
 OBJECT = TARGET_INFORMATION
 TARGET_TYPE = "XXXXXX"
 PRIMARY_BODY_NAME = "XXXXXX"
 ORBIT_DIRECTION = "XXXXXX"
 ROTATION_DIRECTION = "XXXXXX"
 TARGET_DESC = "XXXXXX"
 END_OBJECT = TARGET_INFORMATION

END_OBJECT = TARGET

END