# MARS SCIENCE LABORATORY (MSL) PROJECT SAMPLE ANALYSIS AT MARS (SAM)

# REDUCED DATA RECORD (RDR) ARCHIVE VOLUME SOFTWARE INTERFACE SPECIFICATION

Version 1.0 March 20, 2013 Heather Franz, GSFC/UMBC Susan Slavney, PDS Geosciences Node

# TABLE OF CONTENTS

1. Introduction
1.1. Purpose and Scope1
1.2. Content Overview
1.3. Applicable Documents and Constraints
1.4. Relationships with Other Interfaces
2. Archive Contents
2.1. Root Directory Contents
2.2. Data Directory Contents and Naming
2.3. Index Directory Contents
2.4. Document Directory Contents
2.5. Catalog Directory Contents
2.6. Calib Directory Contents (optional)
3. Archive Format
3.1. Document File Formats
3.2. Table File Formats
3.3. PDS Label Format
3.4. Catalog File Format
3.5. Science Data File Formats
4. Archive Generation
4.1. Data Transfer and Validation Methods10
4.2. Data Product Sizes and Delivery Rates10
4.3. Interface Media Characteristics
4.4. Backup and Duplicates10
4.5. Labeling and Identification
5. Support Staff and Cognizant Persons

## TABLES AND FIGURES

Table 1-1. SAM Standard RDR Data Products	1
Table 1-2. Processing Levels for Science Data Sets	2
Table 2-1. Index Table Contents	5
Table 4-1. Standard Product Sizes and Delivery Rates	0

# DOCUMENT CHANGE LOG

Change	Date	Affected Portions
Initial draft by S. Slavney	4/6/12	All
Edits by H. Franz	5/9/12	All
Edits for Release 1 by S. Slavney	3/4/13	All

# **TBD ITEMS**

Section	Description

# ACRONYMS AND ABBREVIATIONS

ASCII	American Standard Code for Information Interchange
CD-ROM	Compact Disk - Read-Only Memory
CD-WO	Write-Once Compact Disk
EDR	Experiment Data Record
GC	Gas Chromatograph
HTML	HyperText Markup Language
ISO	International Standards Organization
JPL	Jet Propulsion Laboratory
MSL	Mars Science Laboratory
NSSDC	National Space Science Data Center
OPGS	Operations Product Generation Subsystem
PDF	Portable Document Format
PDS	Planetary Data System
PSG	Project Science Group
QMS	Quadropole Mass Spectrometer
RDR	Reduced Data Record
SAM	Sample Analysis at Mars
SIS	Software Interface Specification
TBD	To Be Determined
TIFF	Tagged Image File Format
TLS	Tunable Laser Spectrometer

### GLOSSARY

**Archive** – An archive consists of one or more data sets along with all the documentation and ancillary information needed to understand and use the data. An archive is a logical construct independent of the medium on which it is stored.

Archive Volume, Archive Volume Set – A volume is a unit of media on which data products are stored; for example, one CD-ROM or DVD-ROM. An *archive volume* is a volume containing all or part of an archive; that is, data products plus documentation and ancillary files. When an archive spans multiple volumes, they are called an *archive volume set*. Usually the documentation and some ancillary files are repeated on each volume of the set, so that a single volume can be used alone.

**Catalog Information** – Descriptive information about a data set (e.g. mission description, spacecraft description, instrument description), expressed in Object Description Language (ODL) which is suitable for loading into a PDS catalog.

**Data Product** – A labeled grouping of data resulting from a scientific observation, usually stored in one file. A product label identifies, describes, and defines the structure of the data. An example of a data product is a planetary image, a spectrum table, or a time series table.

**Data Set** – An accumulation of data products. A data set together with supporting documentation and ancillary files is an archive.

**Standard Data Product** – A data product generated in a predefined way using wellunderstood procedures, processed in "pipeline" fashion. Data products that are generated in a nonstandard way are sometimes called *special data products*.

vi

## 1. Introduction

#### 1.1. Purpose and Scope

This Software Interface Specification is intended to be used by those who wish to understand the format and content of the Mars Science Laboratory (MSL) Sample Analysis at Mars (SAM) Reduced Data Record (RDR) Archive. Typically, these individuals would be software engineers, data analysts, or planetary scientists.

The specifications in this document apply to all SAM RDR standard product archives that are generated by the SAM Science Team.

#### **1.2. Content Overview**

This Software Interface Specification (SIS) describes the format, content, and generation of the SAM RDR Archive. Section 2, Archive Contents, describes the structure of the archive and the contents of each file. Section 3, Archive Format, describes the file formats used in the archive. Section 4, Archive Generation, describes the procedure for assembling the archive and delivering it to the Planetary Data System (PDS). Finally, Section 5, Support Staff and Cognizant Persons, lists the individuals responsible for generating the archive.

The SAM RDR archive includes data sets at various levels of processing, from NASA Level 0 products, which are essentially raw data, through NASA Level 2 products, which are results of data interpretation. Four data sets in all are included in the SAM RDR archive, as shown in Table 1-1. The standard processing levels for science data sets are described in Table 1-2 with their commonly used NASA Level and CODMAC Level identifiers. (CODMAC is the Committee on Data Management and Computation; see Applicable Document 6 below.)

Table 1-1. SAM Stanuaru KDK Data Houucis			
Product Name	NASA Level	Description	PDS Data Set ID
SAM L0	0	Unpacking of telemetry into data numbers as raw ADC values or counts and verification of data integrity	MSL-M-SAM-2-RDR-L0-V1.0
SAM L1A	1A	Conversion of raw ADC values or counts to science units	MSL-M-SAM-3-RDR-L1A-V1.0
SAM L1B	1B	Application of corrections to data, e.g., detector dead time, TCD temperature, noise removal, corrections for saturation and instrument response function. Instrument- specific data may include the following types of information: QMS: time, m/z, signal; oven temperature for EGA runs; GC column temperature for GCMS runs	MSL-M-SAM-4-RDR-L1B-V1.0

Table 1-1. SAM Standard RDR Data Products

		GC: TCD signal vs. retention time; pressure; temperature; column used TLS: direct and harmonic spectra	
SAM L2	2	Results of data interpretation completed by the SAM science team. Instrument-specific data may include the following types of information:	MSL-M-SAM-5-RDR-L2-V1.0
		QMS: gas composition; isotope ratios; gas composition vs. sample temperature for EGA runs	
		GC/GCMS: species; relative abundance; identification of pyrolysis products and derivatized compounds	
		TLS: abundance and isotope ratios	

# Table 1-2. Processing Levels for Science Data Sets

NASA	CODMAC	Description
Packet data	Raw - Level 1	Telemetry data stream as received at the ground station, with science and engineering data embedded.
Level-0	Edited - Level 2	Instrument science data (e.g., raw voltages, counts) at full resolution, time ordered, with duplicates and transmission errors removed.
Level 1A	Calibrated - Level 3	Level 0 data that have been located in space and may have been transformed (e.g., calibrated, rearranged) in a reversible manner and packaged with needed ancillary and auxiliary data (e.g., radiances with the calibration equations applied).
Level 1B	Resampled - Level 4	Irreversibly transformed (e.g., resampled, remapped, calibrated) values of the instrument measurements (e.g., radiances, magnetic field strength).
Level 1C	Derived - Level 5	Level 1A or 1B data that have been resampled and mapped onto uniform space-time grids. The data are calibrated (i.e., radiometrically corrected) and may have additional corrections applied (e.g., terrain correction).
Level 2	Derived - Level 5	Geophysical parameters, generally derived from Level 1 data, and located in space and time commensurate with instrument location, pointing, and sampling.
Level 3	Derived - Level 5	Geophysical parameters mapped onto uniform space-time grids.

There is no SAM EDR archive in PDS. The products known as SAM EDRs consist of raw telemetry at the packet level. These are delivered to the SAM Science Team by the MSL Operations Products Generation Subsystem (OPGS). The SAM Team creates the lowest-level RDR product by unpacking the EDR and removing extraneous engineering data. The science content of the EDR is not changed in this process; hence the NASA Level 0 SAM RDR is essentially the raw instrument data.

The SAM Science Team is responsible for creating all the RDR products in this archive, for assembling the archive, and for delivering the archive to the PDS Geosciences Node according to the delivery scheduled determined by PDS and the MSL Project.

#### **1.3. Applicable Documents and Constraints**

This Archive Volume SIS is intended to be consistent with the following documents:

- 1. *Mars Exploration Program Data Management Plan,* R. E. Arvidson et al., Rev. 4.0, June 15, 2011.
- 2. Mars Science Laboratory Archive Generation, Validation, and Transfer Plan, J. Crisp, MSL-214-1333, JPL D-35281, May 28, 2010.
- 3. Mars Science Laboratory (MSL) Software Interface Specification SAM Reduced Data Record (RDR), H. Franz, SIS-SCI020-MSL, JPL D-38123, Version 2.0, February 13, 2013.
- 4. Planetary Data System Archive Preparation Guide, April 1, 2010, Version 1.4, JPL D-31224.
- 5. *Planetary Data System Standards Reference*, February 27, 2009, Version 3.8. JPL D-7669, Part 2.
- 6. Issues and Recommendations Associated with Distributed Computation and Data Management Systems for the Space Sciences, Committee on Data Management and Computation (CODMAC), National Academy of Sciences, National Academy Press, 111p, 1986.
- 7. Mars Science Laboratory Sample Analysis at Mars (SAM) Science Team and PDS Geosciences Node Interface Control Document (ICD), S. Slavney, Version 3.0, April 5, 2011.

#### **1.4. Relationships with Other Interfaces**

This Archive Volume SIS could be affected by changes to the design of the SAM RDR standard data products (Applicable Document 3).

### 2. Archive Contents

This section describes the contents of the SAM RDR Archive, including the directory names, file names, file contents, file types, and organization responsible for providing the files.

#### 2.1. Root Directory Contents

Files in the Root Directory include an overview of the archive, a description of the volume for the PDS Catalog, and a list of errata or comments about the archive. The following files are contained in the Root Directory.

File Name	File Contents	File Provided By
AAREADME.TXT	Volume content and format information	Geosciences Node
ERRATA.TXT	A cumulative listing of comments and updates concerning all archive volumes published to date	SAM Team
VOLDESC.CAT	A description of the contents of this volume in a PDS format readable by both humans and computers	Geosciences Node

#### 2.2. Data Directory Contents and Naming

The Data Directory contains one subdirectory for each SAM experiment, using names in the form EIDnnnn, where nnnnn is a 5-digit SAM experiment identifier. Within each experiment, subdirectories are named LEVEL0, LEVEL1A, LEVEL1B, and LEVEL2 and correspond to the RDR data sets identified in Table 1-1.

See the SAM RDR SIS, Applicable Document 3, for the naming scheme for individual SAM RDR data files.

#### 2.3. Index Directory Contents

Files in the Index Directory are provided to help the user locate products in the archive. There is one index table for each data set. If it ever becomes necessary for PDS to subdivide the archive into multiple volumes, cumulative index tables will be provided that cover all the volumes. The following files are contained in the Index Directory.

File Name	File Contents	File Provided By
INDXINFO.TXT	A description of the contents of this directory	Geosciences Node
L0_INDEX.TAB	A table listing all Level 0 RDR data products on this volume	SAM Team
L0_INDEX.LBL	A PDS detached label that describes L0_INDEX.TAB	SAM Team
L1A_INDEX.TAB	A table listing all Level 1A data products on this volume	SAM Team
L1A_INDEX.LBL	A PDS detached label that describes L1A_INDEX.TAB	SAM Team
L1B_INDEX.TAB	A table listing all Level 1B data products on this volume	SAM Team
L1B_INDEX.LBL	A PDS detached label that describes L1B_INDEX.TAB	SAM Team
L2_INDEX.TAB	A table listing all Level2 data products on this volume	SAM Team
L2_INDEX.LBL	A PDS detached label that describes L2_INDEX.TAB	SAM Team
L0_CUMINDEX.TAB	A table listing all Level 0 data products on all SAM RDR volumes, if the archive contains more than one volume	Geosciences Node
L0_CUMINDEX.LBL	A PDS detached label that describes L0_CUMINDEX.TAB	Geosciences Node

L1A_CUMINDEX.TAB	A table listing all Level 1A data products on all SAM RDR volumes, if the archive contains more than one volume	Geosciences Node
L1A_CUMINDEX.LBL	A PDS detached label that describes L1A_CUMINDEX.TAB	Geosciences Node
L1B_CUMINDEX.TAB	A table listing all Level 1B data products on all SAM RDR volumes, if the archive contains more than one volume	Geosciences Node
L1B_CUMINDEX.LBL	A PDS detached label that describes L1B_CUMINDEX.TAB	Geosciences Node
L2_CUMINDEX.TAB	A table listing all Level 2 data products on all SAM RDR volumes, if the archive contains more than one volume	Geosciences Node
L2_CUMINDEX.LBL	A PDS detached label that describes L2_CUMINDEX.TAB	Geosciences Node

An index table contains one row for each product in the data set. Table 2-1 lists the columns of the index table. If cumulative index tables are present, they have the same structure as the index tables.

Column Name	Description	Source
VOLUME_ID	PDS archive volume on which a data product is stored	Specified at time of index table creation
PATH_NAME	Directory path to data product, relative to volume root	Determined by location of PDS label file
FILE_NAME	Name of <i>label</i> file for data product	PDS label file name
PRODUCT_ID	Unique identifier for data product	PDS label
PRODUCT_VERSION_ID	Version number of data product	PDS label
PRODUCT_CREATION_TIME	Date and time of data product creation	PDS label
RELEASE_ID	ID of scheduled data release to PDS in which this product was included	PDS label
START_TIME	UTC date and time of start of observation	PDS label
STOP_TIME	UTC date and time of end of observation	PDS label
SPACECRAFT_CLOCK_START_COUNT	Spacecraft clock count at start of observation	PDS label
SPACECRAFT_CLOCK_STOP_COUNT	Spacecraft clock count at end of observation	PDS label
PLANET_DAY_NUMBER	Sol number	PDS label
OBSERVATION_ID	Unique identifier for a science observation in the data set	PDS label
SOLAR_LONGITUDE	Angle between Mars-Sun line at vernal equinox and at time of observation; a measure of season	PDS label
MSL:SAM_EXPERIMENT_ID	A number that uniquely identifies the SAM experiment that produced the data	PDS label
MSL:SAM_GC_COLUMN_NUMBER	The SAM GC column number used in the experiment	PDS label

**Table 2-1. Index Table Contents** 

#### 2.4. Document Directory Contents

The Document Directory contains documentation to help the user understand and use the archive data. The following files are contained in the Document Directory.

File Name	File Contents	File Provided By
DOCINFO.TXT	A description of the contents of this directory	Geosciences Node
RDRDPSIS.HTM	The Data Product SIS as hypertext	SAM Team or Geosciences Node
RDRDPSIS.PDF	The Data Product SIS as a PDF file	SAM Team
RDRDPSIS.LBL	A PDS detached label that describes both DPSIS.TXT(HTM) and DPSIS.PDF	SAM Team
RDRARCHSIS.HTM	The Archive Volume SIS (this document) as hypertext	SAM Team or Geosciences Node
RDRARCHSIS.PDF	The Archive Volume SIS (this document) as a PDF file	SAM Team
RDRARCHSIS.LBL	A PDS detached label that describes both ARCHSIS.HTM and ARCHSIS.PDF.	SAM Team
MSL_LDD.FUL, .LBL	MSL local data dictionary and PDS label	PDS Engineering Node
PDSDD.FUL, .LBL	PDS data dictionary and label	PDS Engineering Node
RDRDPSIS_FILES	Subdirectory containing image files referenced in RDRDPSIS.HTM	SAM Team

#### 2.5. Catalog Directory Contents

The files in the Catalog Directory provide a top-level understanding of the mission, spacecraft, instruments, and data sets. The files in this directory are coordinated with the PDS data engineer, who is response

File Name	File Contents	File Provided By
CATINFO.TXT	A description of the contents of this directory	Geosciences Node
SAMRDRL0_DS.CAT	Data set information for the PDS catalog	SAM Team
SAMRDRL1A_DS.CAT	Data set information for the PDS catalog	SAM Team
SAMRDRL1B_DS.CAT	Data set information for the PDS catalog	SAM Team
SAMRDRL2_DS.CAT	Data set information for the PDS catalog	SAM Team
MSL_INSTHOST.CAT	Instrument host (i.e., spacecraft) information for the PDS catalog	MSL Project
SAM_INST.CAT	Instrument information for the PDS catalog	SAM Team
MSL_MISSION.CAT	Mission information for the PDS catalog	MSL Project
PERSON.CAT	Personnel information for the PDS catalog (Team and PDS personnel responsible for generating the archive)	SAM Team
MSL_REF.CAT	References mentioned in MSL*.CAT files	MSL Project
SAM_REF.CAT	References mentioned in SAM*.CAT files and any other pertinent references	SAM Team

#### 2.6. Calib Directory Contents

The Calib Directory contains calibration files used to process the data products, or calibration data needed to use the data products. The following files are contained in the Calib Directory.

File Name	File Contents	File Provided By
CALINFO.TXT	A description of the contents of this directory	Geosciences Node
SAM_HK_CONVERSIONS. CSV, .LBL	Conversion factors for science and engineering units	SAM Team

### 3. Archive Format

This section describes file formats used in the SAM RDR archive. Data that comprise the archive will be formatted in accordance with Planetary Data System specifications [Applicable Documents 4 and 5].

#### 3.1. Document File Formats

Document files with the .TXT suffix exist in all directories. They are ASCII files which may have embedded PDS labels. Lines in a .TXT file end with a carriage return character (ASCII 13) and a line feed character (ASCII 10). This allows the files to be readable under various operating systems.

Documents in the Document directory may contain formatting and figures that cannot be rendered as ASCII text. Therefore each document is given in two formats, hypertext and PDF. The hypertext file contains ASCII text plus hypertext markup language (HTML) commands that enable it to be viewed in a Web browser such as Netscape Navigator or Microsoft Internet Explorer. The hypertext file may be accompanied by ancillary files such as images and style sheets that are incorporated into the document by the Web browser. The second format is PDF (Portable Document Format) from Adobe Systems Incorporated. PDF is an International Standards Organization (ISO) open standard format (ISO 32000) for which free reader software is widely available.

#### 3.2. Table File Formats

Two kinds of table formats are used in the SAM RDR archive, TAB and CSV. Both are ASCII files. Files with the .TAB suffix have fixed-width columns so that all rows in the table are the same length, and the columns are vertically aligned when viewed in a text editor. Columns are separated by commas. Character values are enclosed in double quotation marks, padded with spaces to keep the closing quotation mark at the end of the column. The label that describes a TAB file gives the start byte and number of bytes for each column, along with the column's name, description, and data type. The last two bytes of each row are the ASCII carriage return and line feed characters.

Files with the CSV suffix are Comma-Separated-Value tables. Instead of columns, a row in a CSV table has fields. The fields are separated by commas, but fields are not necessarily the same width from one row to the next and therefore they may not align vertically. The rows in a CSV file may not be all the same length. If there is no data value for a given field, it may be omitted,

indicated by two successive commas with no space between them. This format is useful for saving storage space in sparsely populated tables. As in TAB files, rows in CSV files end in the ASCII carriage return and line feed characters.

#### 3.3. PDS Label Format

Every data file in the archive is described by a detached PDS label; that is, a label in a separate file with the same base name but the extension LBL. Many of the non-data files in the archive are also accompanied by PDS labels, either detached or attached (embedded at the beginning of the file).

A PDS label, whether attached or detached, consists of lines of ASCII text in the form of keyword = value statements that provide descriptive information about the data file. The label is intended to be readable both by humans and by software. Details of the syntax and semantics of PDS labels can be found in the PDS Standards Reference (Applicable Document 5), and definitions of the keywords used in the label can be found by using the PDS Data Dictionary Lookup web service at http://pds.jpl.nasa.gov/tools/ddlookup/data\_dictionary\_lookup.cfm.

Lines of text in attached and detached labels end with an ASCII carriage return and line feed pair. This allows the files to be read under various operating systems.

#### 3.4. Catalog File Format

Catalog files (suffix .CAT) exist in the Root and Catalog directories. Like PDS labels, they are text files formatted as keyword = value statements which are both human-readable and software-readable. They contain descriptions of the data set, instrument, spacecraft, and mission, as well as personnel contact information and references to published literature. They are called Catalog Files because they are loaded into the PDS online catalog to make the information available to users searching for data.

#### 3.5. Science Data File Formats

One SAM RDR product consists of one PDS label and one or more data files. A label that describes more than one data file is called a combined-detached label (Applicable Document 5). In cases where multiple data files comprise a single RDR product, a list of the data files will appear near the beginning of the label using the PDS file pointer syntax, as in this example:

	/* POINTER: ^FILE ^FILE ^FILE	=	"SM241282 "SM241282	A FILES D A0141RDR1 A0141RDR1 A0141RDR1	B_SPYR_( B_SPYR_(	QMS_PAR. QMS_MAS	AMSX_1.T SDTC 1.T	AB" AB"		
The	name	••		label		in	this	example	would	be
SM2412	28A0141RDR1B	_SPY	R_QMS_XX	XXXXX_1.L	BL.					

Most SAM RDR data products are ASCII text files in the TXT, TAB, and CSV formats described above. These products are viewable in any text editor. TAB and CSV files are also suitable for use in spreadsheet and data base management software.

The SAM Level 2 RDR data set also includes TIFF image files (extension .TIF) as illustrations of the science results reported in the TXT and CSV files. TIFF images may be displayed by many

image processing programs, including free image processing software. They are not viewable in web browsers.

For more information about the format and content of the data products, see the Data Product SIS, Applicable Document 3.

### 4. Archive Generation

#### 4.1. Data Transfer and Validation Methods

SAM RDR products are generated by the SAM science team at NASA Goddard Space Flight Center. The products are delivered to the PDS Geosciences Node at Washington University electronically using FTP according to the schedule in the Mars Science Laboratory Archive Generation, Validation, and Transfer Plan (Applicable Document 2).

#### 4.2. Data Product Sizes and Delivery Rates

Table 4-1 summarizes expected sizes and production rates for the SAM RDR Standard Products.

Product	Product Size	Production Rate	Expected Number of Products for Primary Mission (669 sols)	Expected Total Data Volume for Primary Mission
SAM L0	80 MB	One per experiment	80	6.4 GB
SAM L1A	95 MB	One per experiment	80	7.6 GB
SAM L1B	20 MB	One per experiment	80	1.6 GB
SAM L2	1 MB	One per experiment	80	80 MB

 Table 4-1. Standard Product Sizes and Delivery Rates

#### 4.3. Interface Media Characteristics

All volumes in the SAM RDR Standard Product Archive conform to ISO 9660 standards [ISO 9660, 1988].

#### 4.4. Backup and Duplicates

The SAM Team will retain backup copies of all data delivered to the Geosciences Node until the data have been released to the public.

The Geosciences Node will maintain three backup copies of the SAM RDR archives that are kept online for public use, according to standard PDS policy for backup copies. When the SAM RDR archive is complete at the end of the MSL mission, the Geosciences Node will deliver a copy of the archive to the National Space Science Data Center (NSSDC) for long-term storage, but the Geosciences Node will maintain its own copies of the archive indefinitely.

#### 4.5. Labeling and Identification

All SAM RDR data sets are stored on one PDS archive volume (that is, under one directory structure). Under the top level DATA directory there is a separate subdirectory for each data set,

as described in section 2.2. During the course of the mission, as new data products are delivered to PDS they are added to the existing archive volume. The unique identifier for the SAM RDR archive volume is MSLSAM\_1XXX. If it ever becomes necessary for PDS to subdivide the archive volume into smaller volumes for data storage or transfer, they will be identified as MSLSAM\_1001, MSLSAM\_1002, etc.

# 5. Support Staff and Cognizant Persons

Susan Slavney, PDS Archive Point of Contact for SAM PDS Geosciences Node Washington University St. Louis, Missouri Susan.Slavney@wustl.edu

Paul R. Mahaffy, SAM Principal Investigator NASA Goddard Space Flight Center, Code 699 Greenbelt, Maryland paul.r.mahaffy@nasa.gov

Heather Franz, SAM Point of Contact for PDS NASA Goddard Space Flight Center, Code 699 Greenbelt, Maryland heather.b.franz@nasa.gov

Matthew Lefavor, SAM software support NASA Goddard Space Flight Center, Code 699 Greenbelt, Maryland matthew.lefavor@nasa.gov