
**Modeling and Data
Analysis Working Group**

SOL-SGS-TN-0009

**Metadata Definition for
Solar Orbiter Science**

Author: MADAWG Team

Version: 1.0

Date: 2015.01.13



TABLE OF CONTENTS

1	Introduction	3
1.1	Reference Documents	3
2	Solar Orbiter Data Products Formats	5
2.1	Data Formats and File Names	5
2.1.1	Solar Orbiter Data Format	5
2.1.2	Solar Orbiter File Naming Convention	5
2.2	Definition of Processing Levels	8
3	Solar Orbiter Metadata Standard	13
3.1	FITS Metadata Standard	13
3.1.1	Primary Header Keywords	13
3.1.2	FITS Extensions	54
3.2	CDF Metadata Standard	59
3.2.1	CDF variables	59
3.2.2	Global attributes	63
3.2.3	Variable attributes	71
4	Appendices	80
4.1	UCDs	80
4.1.1	Building UCDs	80
4.1.2	Example of UCDs	81
4.2	Instrument-Specific Requirements	82
4.2.1	Instrument-specific Requirements - TEMPLATE	82
4.2.2	EUI specific issues	83
4.2.3	RPW (IS)	87
4.2.4	SPICE (RS)	92
4.2.5	STIX (RS)	94
4.3	VESPA keywords for Unified Solar Orbiter Catalog (USOC)	95

1 INTRODUCTION

Most Solar Orbiter science goals are based on coordinated observations of its 6 remote-sensing and 4 in-situ instruments [SolO-SRD]. As a direct consequence, the scientific users of Solar Orbiter data will need to access data sets from different instruments and perform scientific analysis on a multi-instrument dataset. Due to the mission's particular orbit and operational constraints, also data products from other space or ground assets might be required to provide the necessary context for the Solar Orbiter data.

However, the scientist/data user in question might not be an expert on all the different types of instruments and data. Connection science in particular requires the combination of very different types of data and the interoperability of analysis software of two quite distinct scientific communities.

It is therefore crucial for the scientific outcome of the mission, and in the benefit of the whole scientific community, to make the full range of Solar Orbiter data easily accessible and easily interpretable by defining data products which are consistently formatted in a transparent and standard manner. All data (both operational and science) will also be stored in the mission archive at the Solar Orbiter SOC (ESAC, Madrid) from which (a subset) will be made available to the science community.

Another important aspect for data provision to the scientific community is the compliance to existing tools and broadly used software. As an answer to the complementarity of recent missions like STEREO, SDO, IRIS, etc. the solar physics community has developed over the last years a wide range of common data archives, retrieval systems and analysis software tools that facilitate the combination of many data types and sources. Therefore it is highly recommended to also make the Solar Orbiter science data compatible to the most widely used software packets, so that data can be digested with tools like (J)Helioviewer, 3Dview, AMDA, (J)Helioviewer, 3Dview, FESTIVAL and ... (TBD)

To facilitate connection of in-situ and remote-sensing observations, compliance with connection models, e.g. magnetic field extrapolations like PFSS, radial constant speed outflow models, and (CME) propagation models like ENLIL, is highly recommended.

In this technical note, we describe a proposed format for Solar Orbiter's data products and metadata.

- [Reference Documents](#)

1.1 Reference Documents

[SolO-SRD]	Solar Orbiter Science Requirements Document, SOL-EST-RS-1858
[LLdata]	Solar Orbiter Low-Latency Data: Concept and Implementation, SOL-SGS-TN-003

[EID-A]	Experiment Interface Document – Part A, Iss. 4.0, SOL.EST.RCD.0050_04_00
[SIRD]	Solar Orbiter Science Implementation Requirements Document SOL-EST-RS-4514
[FITSpaper]	Definition of the Flexible Image Transport System, http://fits.gsfc.nasa.gov/standard30/fits_standard30aa.pdf
[FITS+WCS]	FITS World Coordinate System (WCS), http://fits.gsfc.nasa.gov/fits_wcs.html
[FITSchecksum]	Checksum Keyword Convention, http://fits.gsfc.nasa.gov/registry/checksum/checksum.pdf
[WCSpaper]	Coordinate systems for solar image data, W. Thompson, http://adsabs.harvard.edu/abs/2006A%26A...449..791T
[WCSpaper2]	Precision effects for solar image coordinates within the FITS world coordinate system, W. Thompson, http://adsabs.harvard.edu/abs/2010A%26A...515A..59T (defines some additional keywords)
[WCStime]	WCS paper on time coordinates, A. Rots, http://hea-www.cfa.harvard.edu/~arots/TimeWCS/WCSPaper-IV-v1.1A4.pdf
[ISTP]	ISTP guide, http://spdf.gsfc.nasa.gov/istp_guide/istp_guide.html
[CDF]	CDF User's Guide (V3.5.0), http://cdaweb.gsfc.nasa.gov/pub/software/cdf/doc/cdf350/cdf350ug.pdf

More papers/webpages that can be of interest:

[SECCHI FITS]	http://hesperia.gsfc.nasa.gov/ssw/stereo/secchi/doc/FITS_keywords.pdf
[SWAP FITS]	http://proba2.sidc.be/data/SWAP/level0
[EIS FITS]	http://solarb.uio.no/software/eis_userguide/fits_kw_list.htm
[IRIS FITS]	https://www.lmsal.com/iris_science/doc?cmd=dcur&proj_num=IS0077&file_type=pdf

2 SOLAR ORBITER DATA PRODUCTS FORMATS

As explained in the introduction, it is of crucial importance for the science goals of Solar Orbiter to facilitate scientific analysis of multi-instrument and multi-mission datasets. This requires a minimal set of science data formats that are common to all Solar Orbiter data products and are consistent with commonly used data formats in the science community.

2.1 Data Formats and File Names

2.1.1 Solar Orbiter Data Format

The proposed formats for the Solar Orbiter scientific data products are as follows:

- Image data should be stored in FITS files with standardized FITS keywords (see Sect. Solar Orbiter Metadata Standard). One FITS file can contain one or a series of images.
- Time series should be stored in CDF files per day, or per burst mode period. Metadata should follow the standard described in Sect. CDF Metadata Standard.
- Combined data sets or spectral data (e.g. STIX, SPICE) can be stored in FITS files with binary table extensions or in multi-image FITS files. The primary FITS header should be standardized according to the rules set out in Sect. [FITS Metadata Standard](#).

All science data should include a data header, FITS keywords or CDF global and variable attributes, that follow the Solar Orbiter standard described in [Solar Orbiter Metadata Standard](#). The use of detached headers (metadata in a separate file) is not acceptable.

Higher processing levels of data products might need different formats and are defined in [Definition of Processing Levels](#).

2.1.2 Solar Orbiter File Naming Convention

The naming convention described below is proposed for all files that aren't raw telemetry, i.e. L1 and upwards, including calibration and ancillary data. It is based on both the ISTP file naming convention used in the in-situ community and the needs and common practices in the remote-sensing community.

As a compromise between the needs and conventions of the in situ and remote sensing payloads, Solar Orbiter data files will be named according to the following convention which has been defined with the aim of providing different instrument teams with the flexibility they need within a consistent framework:

source_level_descriptor_dataproduct_datetime_version_freefield.extension

Where the different fields that make up the filename are separated by underscores and are defined as follows:

Field	Description	Mandatory / Optional
source	Identifies the data as from Solar Orbiter. Content will be 'solo' (TBD).	Mandatory
level	Processing level. Content will either be L1, L2, L3, LL0-1, LL0-2, LL0-3, ... according to the Solar Orbiter data processing level standard .	Mandatory
descriptor	The instrument and sensor or detector (if appropriate), separated by a hyphen e.g. METIS or SWA-PAS. This is equivalent to the combination INSTRUMENT-DETECTOR from the FITS metadata. Level 3 products derived from multiple sources can use MULTI as a descriptor.	Mandatory
dataprodu	A means to identify files containing different data types or parameters which have the same processing level e.g. LCURVE and IMAGE could be included for STIX lightcurves and images, respectively.	Optional https://stix.cs.technik.fhnw.ch/confluence/pages/viewpage.action?pageId=18808854#link-talk-103
datetime	The timestamp of the data in the file, i.e. the UTC at which the observation was made by the spacecraft. Formatted according to section 1.2.1 below.	Mandatory
version	The version of the file. Format TBD, but first character must be capital V.	Mandatory
free field	An optional field of which the format is free, but you cannot use a decimal point or underscore.	Optional
extension	File extension: '.fits' for all FITS files, '.cdf' for all CDF, '.jp2' for JPEG2000 files (more extensions might be added later). Simple ASCII text files must also have an extension.	Mandatory

The formats of descriptor and dataprodu are not fixed, and dataprodu need not be included at all, but the following rules must be observed:

- If the dataprodu field is not needed, descriptor and datetime should be separated by two underscores, i.e. the dataprodu field is always there but can be of zero length.
- Sub-instruments/sensors can be included in descriptor, but must be separated by a hyphen rather than an underscore e.g. EUI-FSI. Similarly, dataprodu may contain subfields but these must also be separated by a hyphen. Underscores are reserved as field separators.

2.1.2.1 Datetime Format

1. In common with the rest of the filename, the datetime field will be separated from other fields in the filename by an underscore .
2. As a minimum, the datetime field will be the start of the interval (for time series) or the start of the integration (i.e. content of FITS keyword DATE-BEG) for images.
3. Granularity of datetime can be chosen as appropriate, but date will be separated from time with a capital T. All more significant fields must be included e.g:
 - a. For day granularity files: `yyyymmdd`
 - b. For millisecond granularity files: do not use the decimal point but add the required number of digits to the time field: `yyyymmddThhmmssxyz`
4. An end datetime is optional, but if present shall be separated from the start datetime by a hyphen, '-'.
5. The end datetime shall have the same granularity as start datetime.
6. For time series, in the absence of the end datetime the file duration is assumed to be the granularity of start datetime, e.g. `yyyymmdd` has 1 day duration, `yyyymmddThh` has one hour duration.

In situ normal mode files are encouraged to be daily, without an end datetime. In situ burst mode files are encouraged to be per burst, with an end datetime.

For RS imagery, the recommendation is to only specify the acquisition start time unless the observation is constructed from many acquisitions taken over a substantially long period.

2.1.2.2 Version format (TBD)

How to specify the file version in the filename is still open for discussion, and is linked to the discussion on version keywords in FITS headers and version attributes in CDF.

Data product versions should at least take into account

- software revision numbers
- configuration parameter version, if applicable
- calibration data version
- any other trigger of data reprocessing

The filename could use a combination of the versions above, or a counter that links back to the combination of versions listed above. Additionally we want to limit the number of digits to keep the filename readable.

In any case it is vital that every instrument keeps track consistently.

2.1.2.3 Capitalisation

Current proposal is to keep filenames lowercase, except for a few meaningful letters that should always be capitalised: L for level, V for version, T as separator between date and time, and ANC and CAL as 'levels' for ancillary and calibration data. (TBC)

2.1.2.4 Example Filenames

A Level 2 SoloHI image might have the following filename:

solo_L2_solohi__20181012T0456_V01.fits (note the double underscore between solohi and 2018)

A MAG 16 vectors per second cadence daily file might have the following filename:

solo_L2_mag_16vps_20181012_V02.cdf

SWA-EAS Pitch Angle Distributions from a burst mode might have the following filename:

solo_L2_swa-eas_pad_20181012T045630-20181012T050630_V01.cdf

A highly compressed, yet uncalibrated EUV-FSI 304Å image that got downloaded as Low-Latency data and was produced by the SOC might have the following name:

solo_LL0-2_eui-fsi_304_20201008T121230_V01.fits

and can be distinguished from the same uncalibrated image at science resolution that got downloaded much later from the level in the filename:

solo_L1_eui-fsi_304_20201008T121230_V01.fits

A raw EUV/HRI image from a sequence operating at 10Hz imaging sequence that makes use of the free field would look like

solo_L0_eui-hri_174_20201108T121230899_V01_11234.fits (11234 would be a pointer towards an entry in a database uniquely defining the origin of this image)

A SPICE summary map of the FIP fractionation abundance data may be named:

solo_L3_spice_fip_20201008T121230_V03.fits

2.1.2.5 Calibration and Ancillary Data (FOR DISCUSSION)

Calibration and Ancillary Data will follow the same standard. Calibration files will have CAL included as the level, further subcategories can be added as appropriate in *dataprodukt*, separated by hyphens. E.g.

solo_CAL_mag-ibs__20191001-20201001_V01.cdf

Similarly, ancillary files will have ANC included in *dataprodukt* as the first subfield. SOC-Produced Ancillary files (e.g. the orbit spice kernel) will have 'soc' as the descriptor, e.g.

solo_ANC_soc_orbit_20190101-20190301_V01.spk

2.2 Definition of Processing Levels

Historically, data processing levels have been defined differently in different scientific communities. However, it would be highly beneficial to use a consistent definition throughout the Solar Orbiter data products.

The proposal below is based on the high-level description of [EID-A] and [SIRD] but adds more detail in the description of the data levels. Due to the specific nature of some of the SoLO data products, Instrument Teams (IT) are free to define multiple types of data in each data level, if appropriate. These should be distinguishable through the filename (see [Data Formats and File Names](#)). All data levels available for each instrument should be stored in the mission archive at SOC, although not all data processing levels are necessarily open to the scientific community. Optional data products are formatted *italic* in the table below.

Level	Source	Data Type	Format and Metadata content
L0	IT	"Raw" data, unpacked and uncompressed data	FITS or CDF, metadata reflect the information that was available in the TM packets <u>only</u> .
L1	IT	"Engineering" data, uncalibrated	FITS or CDF, metadata follows Solar Orbiter standard for L1 (see Section Solar Orbiter Metadata Standard): <i>Note that this level might not always apply to instruments using a complete processing and calibration pipeline onboard, like PHI.</i>
L2	IT	"Calibrated" data, science quality	FITS or CDF, metadata follows Solar Orbiter standard for L2 (see Section Solar Orbiter Metadata Standard): full attitude information in WCS coordinate frame and time in UTC.
<i>L3</i>	<i>IT</i>	<i>Higher-level data</i>	<i>Data format as appropriate (The format of both Level-3 data and Calibration data can be chosen depending on the type of data product and the objectives. However, as much as possible standard formats should be used (MPEG, FITS, JPEG2000, CDF, PNG, ...)).</i>
CAL	IT	Calibration data	Data format as appropriate. <i>Not all calibration data are necessarily open to the scientific community.</i>
ANC	IT/SOC	Ancillary data	Data format as appropriate. <i>Not all ancillary data are necessarily open to the scientific community.</i>
LL01	SOC	LL engineering data, output of LL pipeline	FITS or CDF, metadata follow Solar Orbiter standard, with some specifics for LL-0.1 data (see LL section to be added later) : time in OBT, attitude in instrument reference frame.
LL02	SOC	Operational LL data, enhanced with S/C HK	FITS or CDF, metadata follow Solar Orbiter standard, with some specifics for LL-0.2 data (see LL section to be added later) : time in UTC, attitude in WCS coordinate frame.

Level	Source	Data Type	Format and Metadata content
LL03	SOC	Operational LL data, in "quicklook" format	'Quicklook' data in PNG or JPEG2000 (details TBC)
LL2	IT	"Calibrated" LL data	<i>FITS/CDF. Note that this level is optional. In the case LL data are of scientific quality, the calibrated product should be called 'L2'.</i>
LL3	IT	"Calibrated" LL data	<i>JPEG2000. Note that this level is optional. In the case LL data are of scientific quality, the calibrated product should be called 'L3'.</i>

The data levels can be described in more detail as follows:

- Level 0 data are basically telemetry packets, that are decommutated and decompressed and formatted in standard formats, i.e CDF or FITS files. The metadata should contain only the information that was available from the packet headers. No metadata are added from elsewhere. L0 data may be structured per sensor or instrument unit, similar to how the data were packed onboard.
- Level 1 data are still uncalibrated data. These data are formatted in standard formats, i.e. FITS files for images (1 FITS file per image or image sequence) and CDF (daily files or per burst period) for time series. They form the basis of the scientific quality data in Level 2. They are still expressed in engineering units (counts) and may contain extra engineering metadata from other sources. Still, this type of data is not yet of science quality. It may also lack details on the instrument's boresight or attitude at the time of the data acquisition. Examples are "raw" images in counts, not corrected for flatfield, dark current,... ; particle spectra in counts ; MAG data in which S/C disturbances are still included; etc.
In case Level 1 data are distributed to the scientific community, they should be accompanied by calibration files and processing software to enhance them to Level 2 for scientific analysis.
- Level 2 data are calibrated data, ready for scientific analysis. Depending on the type of instrument, "full calibration" might have a different meaning. Data products for which this is appropriate, instrument counts/measurements will be calibrated to physical units. For other instruments (e.g. broadband imagers) it might be more appropriate to limit calibration to the correction of data for all known instrumental defects, and calibrate detector counts in exposure-normalized counts rather than in physical units. In any case, level 2 data should be the reference science product for the scientific user, shielding the user from any instrument peculiarities. All the available orbital and attitude information should be applied and coordinates expressed in WCS coordinate systems.
Level 2 data are likely to contain subcategories, depending on the science goal. Time series might for example be distributed both in full temporal resolution as resampled to lower cadence. These data should however be suitable for quantitative analysis, otherwise they should be categorized as Level 3 data.

- Level 3 data are higher-level data products, in the sense of derived data products that either combine several 'single' level-2 data products or are a derived data product that is the result of data analysis. Examples of the former are: movies for imagers, images that combine FOVs of different sub-units, mosaics, graphical plots of dynamic spectra, etc.
Examples of derived, value-added data products are: event catalogues, graphs/plots linking in-situ to remote-sensing measurements, derived plasma beta, Alfvén Mach number, ...
Many of the data products in this category may be only defined after launch. Also Level 3 data are not necessarily limited to one instrument.
- Calibration data are data used typically to derive the Level 2 data from the Level 1 data. The calibration data are made from any available data, incl. pre-launch test data, spacecraft attitude information, some Level 0/1 data (e.g. dark images) and appropriate spectrum synthesis models. The format of the calibration data is chosen depending on the type and objective of the data product but should re-use as much as possible the data formats used for Level 1-3 data.
- Ancillary data are any data that are provided to the scientific community as context for the scientific data. These are data products derived from S/C or payload HK TM that are useful for scientific analysis, and could include: instrument status for all payload, SPICE kernels for orbit and S/C attitude, time conversion files (OBT to UTC), time periods of burst modes, science operation plans, etc.
The data format will depend on the type of ancillary data. DAWG recommends to also include instrument-specific ancillary data, in the form of an 'engineering logbook' that accompanies each dataset and lists the available engineering information such as special observing modes, anomalies, data gaps, etc.
- Low-Latency data are special in the sense that there might be two distinct categories of LL data, both produced from the same TM products (i.e. science TM with LL APID).
 - The first category are the so-called "operational LL data" that are necessary for mission planning and S/C pointing definitions at the SOC. These data products are time critical and are therefore produced at the SOC: LL TM is first reformatted by instrument-provided LL data pipelines into Level LL-0.1 data, and then post-processed by SOC software to apply common operations (like conversion to UTC, application of the spacecraft attitude, and known instrument alignments), resulting in Level LL-0.2 data. SOC will also produce a quicklook format of the latter, LL-0.3 -data, for visualization in the Low-Latency Data Visualisation tool. All 3 SOC-produced 'operational LL data' will be available through the mission archive, at least to the instrument teams.
 - During times of low downlink performance, LL data might be the only recent data available from Solar Orbiter during several months. Therefore, instrument teams might want to produce an enhanced, better calibrated, version of the LL data starting from the original Level-0 LL TM packets. To distinguish both versions, we propose to call the optional instrument-produced data LL-2 and LL-3 data, depending on the format (science data format vs quicklook format), or L-2 and L-3 in case the LL data are of sufficient science quality and do not need distinction from nominal L2 data.

- It is proposed to format Quicklook imagery made from LL data as JPEG2000 images, or at least in a data format compatible to (J)Helioviewer.

3 SOLAR ORBITER METADATA STANDARD

The following sections describe a proposal for a standardized list of metadata keywords. FITS keyword lists are defined in Sect. 3.1 and CDF metadata in Sect. 3.2.

3.1 FITS Metadata Standard

Below we propose a list of keywords that should be provided in SoIO science FITS files of Level 1 (engineering data) and Level 2 (calibrated science data). The proposal follows the FITS standards [\[FITSpaper\]](#) and use of the World Coordinate System (WCS) [\[WCSpaper\]](#) for position and pointing metadata. The proposed list of FITS keywords is based on the metadata definitions used for the STEREO/SECCHI [\[SECCHI FITS\]](#) and PROBA2/SWAP [\[SWAP FITS\]](#) and also inspired by keyword definitions of Hinode/EIS [\[EIS FITS\]](#) and IRIS [\[IRIS FITS\]](#) instruments.

In the keyword lists below, grouped by metadata type, we distinguish 3 types of keywords:

- M: Mandatory keywords to comply to the standard FITS format.
- P: Keywords we propose to be mandatory in the Solar Orbiter FITS files. They should comply with the description and units described below.
- O: Keywords that are optional and can be defined if applicable to the instrument in question. However, again, if they are defined with the proposed name they should be provided in the units and follow the description below.

Note that keywords marked with * are optional in FITS standard but if they are present in a FITS header, they must be used as defined in the standard.

Instrument-specific keywords will be added later as appendices to this document.

3.1.1 Primary Header Keywords

Basic Keywords

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
SIMPLE	T	file conforms to basic FITS standards	T (for true)	Boolean	SIMPLE is a logical keyword specifying whether the file conforms to basic FITS standards; Solar Orbiter data should always comply, so T (for True) is only valid option here.	M	All
BITPIX	16	number of bits per data pixel, rounded up	8, 16, 32, 64 for binary integers -32, -64 for single/double precision floating points	Byte	BITPIX states the number of bits per pixel and can only take values 8, 16, 32, 64 for binary integers and -32 or -64 for single, resp. double precision floating points (see table 8 in [FITSpaper]) so the actual bit depth should be rounded up if necessary.	M	All
NAXIS	2	number of axes in data cube	0, ..., 4+	Byte	The NAXIS keyword indicates the number of dimensions in the data (0 indicates header only).	M	All
NAXIS1	1024	length of data axis 1	positive integer	Integer	There should be a NAXISx keyword for each of the dimensions specified above, that describes the number of pixels along that axis. The example given here is for a 2D image with 1024x1024 pixel resolution. A	M	All


Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					3D data cube would need an extra keyword NAXIS3.		
NAXIS2	1024	length of data axis 2	positive integer	Integer	Length of the second axis (y, number of rows in 2D image).	M	All
EXTEND	F	F = no extensions T = extension(s)	T, F (for true, false)	Boolean	EXTEND specifies whether the FITS file contains extension tables or multiple images or not, and is proposed to be a standard keyword in SoIO data FITS headers. Mandatory keywords in extension table headers are specified in Table 10 of [FITSpaper] and SoIO standard keywords in the section Secondary Header Keywords (Extension Images and Tables) .	P	All

General Description Keywords




Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
FILENAME	'solo_L2_mag_16vps_20181012_V02.cdf'	FITS filename	see filename convention	String	Name of FITS file. FILENAME should follow the conventions laid out in	P	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					Sect. 2.1 (Data Formats and File Names), with the date/time consistent with the information in DATE-BEG below.		
FILE_RAW	<i>TBD</i>	Raw filename	<i>TBD</i>	String	Optional keyword referring to a raw file the FITS file was derived from. (TBC)	O	L1+
FILE_PACK	<i>TBD</i>	Raw telemetry package filename	<i>TBD</i>	String	Optional keyword referring to the raw telemetry package file, if applicable. (TBC)	O	All
APID		APID number of associated TM	positive	Integer	APID number of associated TM.	O	L0, 1
DATE	'2020-01-27T14:23:57.167'	[UTC] FITS file creation date	any date+'T'+time	String	Date and time of FITS file creation, in UTC, in ISO-8601 format 'yyyy-mm-ddThh:mm:ss.sss'	P	L1+

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
DATE-OBS	'2020-01-22T14:00:00.000'	[UTC] Deprecated, same as DATE-BEG.	any date+'T'+time within mission duration	String	Time of observation, in UTC, in ISO-8601 format. Historically, this has usually referred to the start of the observation, though it has sometimes been used for the midpoint. The [WCStime] paper defines two new keywords, DATE-BEG and DATE-AVG, to replace DATE-OBS. To support both older and newer software, Solar Orbiter files will contain both DATE-OBS and the DATE-BEG, which will have the same value.	P	L1+
DATE-BEG	'2020-01-22T14:00:00.000'	[UTC] Start	any date+'T'+time	String	Start time of observation,	P	L1+

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
		time of observation	time within mission duration		in UTC, in ISO-8601 format. DATE-BEG replaces DATE-OBS that got depreciated.		
DATE-AVG	'2020-01-22T14:00:30.000'	[UTC] Average time of observation	any date+'T'+time within mission duration	String	Average time of observation, in UTC, in ISO-8601 format. This will usually be the midpoint between the start and end times. Mandatory for Solar Orbiter as all attitude and WCS metadata are based on this time.	P	L1+
DATE-END	'2020-01-22T14:01:00.000'	[UTC] End time of observation	any date+'T'+time within mission duration	String	End time of observation, in UTC, in ISO-8601 format.	P/O? 	L1+
TIMRDER	0.2	[s] Estimated random error in time values			Estimated random error in time values, expressed in seconds.	O	


Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
TIMSYERR	0.1	[s] Estimated systematic error in time values			Estimated systematic error in time values, expressed in seconds. The total error is the root-summed-square of the random and systematic errors.	O	
OBT_BEGIN	254518633	start acquisition time in OBT	positive float	Long	Mandatory keyword that contains the start time of data acquisition in onboard time units (floating-valued).	P	All
OBT_END	254518854	end acquisition time in OBT	positive float	Long	OBT_END expresses the end time of data acquisition in onboard time units (floating-valued). It is only mandatory in case the end acquisition time was contained in the TM packets.	O	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
LEVEL	'L1'	Data processing level	'L0', 'L1', 'L2', 'L3', 'LL01', 'LL02', 'LL2', 'LL3', 'ANC', 'CAL'	String	Data processing level, should be consistent with definition in Sect. 2.2 (Definition of Processing Levels)	P	All
CREATOR	'EUI-EDG v3.2'	FITS creation software	any	String	Name of the software that created, or last touched, the FITS file, with version (TBC ).	P	All
ORIGIN	'SSOC'	<name of institution>	any	String	Location where the FITS file has been generated, e.g. 'SSOC' for 'Solar Orbiter Science Operations Centre', or the institute name, or ...	P	All
Calibration  version (keyword to be defined)						P	L2+
SW  version						P	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
(keyword to be defined)							

Instrument and Observation Configuration

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
OBSRVTRY	'Solar Orbiter'	Satellite name	'Solar Orbiter'	String	Observatory or satellite name, in this case Solar Orbiter.	P	L1+
TELESCOPE	'SOLO/EUI/F SI'	Telescope/Sensor name	"SOLO/" + \$INSTRUME/ [+ \$DETECTOR]	String	The full description of the telescope or sensor that took the measurement. In case of instruments with several telescope channels (different optical paths), the full chain should be specified and the channel should be added (TBC). For in situ instruments, if a sensor has multiple apertures, and the data come from only one of those, the aperture should be specified as	P	L1+

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					well.		
INSTRUME	'EUI'	Instrument name	'EUI', 'METIS', 'PHI', 'STIX', 'SoloHI', 'EPD', 'MAG', 'SPICE', 'RPW', 'SWA' <i>to be updated by ITs</i>	String	This keyword should reflect the 'official' instrument acronym, i.e. how you would want it to appear in a publication. Capitalisation does not necessarily have to match the filename. Instrument names in INSTRUME should be used consistently (see valid data range). Subunits, if applicable, are specified through the keyword DETECTOR.	P	All
DETECTOR	'FSI'	Instrument subunit or sensor	<i>to be defined by instr teams</i>	String	To specify the subunit/sensor, e.g. 'FSI', 'HRI_EUV', 'HRI_LYA' for EUI, 'VL', 'UV' for METIS, ... to be defined by instrument teams	O	L1,2
OBJECT	'Full Sun'	Type of object observed	TBD 	String	The use of the keyword OBJECT is TBD. If used, we should	O	L1+

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					define a standard list of options, like 'Full Sun, AR, CH, QS, ...'		
OBS_MODE	'Synoptic_low'	Observation mode	any string (Limited list of observation modes per instrument)	String	Description of the observation mode or 'study' that has been used to acquire this image (series). This keyword/value would be used to find all data that have been acquired with similar settings (varying parameters might still apply). There should be a limited list of observation modes defined per instrument. These are also considered the building blocks ('ingredients') of the SOOP campaigns.	P	L1,2
STUDY_ID	7	ID of the observation mode used	any	Integer	Similar to OBS_MODE but numeric identifier per observation mode or study. Optional keyword.	O	L1,2
FILTER	'<filter	Description of	any	String	To specify the filter in multi-	O	L1,2

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
	name>'	filter used			filter instruments Optional keyword, but format can be mandated based in any commonalities between instruments' proposed usage (comments welcome).		
WAVELNTH	174.0	[Angstrom] characteristic wavelength of observation	<i>TBD by ITs</i>	Float	WAVELNTH contains the characteristic wavelength at which the observation was taken, in Ångstrom. It is optional as it might not be appropriate to all data products. It applies to instruments or sensors using a narrow bandpass. For EUV imagers, this keyword typically identifies the most prominent emission line in the bandpass.	O	L1,2
WAVEMIN	170.1	[Angstrom] Min wavelength where response	<i>TBD by ITs</i>	Float	The shortest wavelength at which the net (approximate)	O	L1,2

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
		> 0.05 of max			response function becomes 0.05 times the maximum response. Included for SVO-like search purposes, to identify useful observations with contributions from unexpected lines. If the response function is not known out to the wavelength(s) where this happens, the relevant min/max value(s) should be set to the limits of the area in which the response function is known/estimated. Only applies to bandpass instruments (both narrow-band and wide-band) and spectrometers.		
WAVEMAX	178.2	[Angstrom] Max wavelength where response	<i>TBD by ITs</i>	Float	The longest wavelength at which the net (approximate)	O	L1,2

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
		> 0.05 of max			response function becomes 0.05 times the maximum response. Included for SVO-like search purposes, to identify useful observations with contributions from unexpected lines. If the response function is not known out to the wavelength(s) where this happens, the relevant min/max value(s) should be set to the limits of the area in which the response function is known/estimated. Only applies to bandpass instruments (both narrow-band and wide-band) and spectrometers.		
WAVEBAND	'ENERGY_L'	Bandpass description	any string	String	Description of the wavelength band. Typically used for wider	O	L1,2

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					bandpasses containing several significant lines.		
XPOSURE	10.0	[s] Total effective exposure time	Positive	Float	Total effective exposure time of the observation, in seconds. For observations made of multiple exposures summed together, this would be the total sum of the individual effective exposure times.	P	L1,2
NSUMEXP	10	Number of detector readouts summed together	1, 2, 3, ...	Integer	Number of images summed together to form the observation. The average effective exposure time of an individual exposure can be calculated as XPOSURE / NSUMEXP.	O	L1,2
TELAPSE	10.0	[s] Elapsed time between beginning and end of observation	Positive	Float	Total elapsed time between the beginning and end of the complete observation in seconds,	O	L1,2

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					including any dead times between exposures. In other words, this is the time from the start of the first exposure, to the end of the last exposure. For rasters, this would be the duration of the entire raster.		
TIMESYS	'UTC'	System used for time keywords	See [FITS+WCS]	String	System used for time keywords. For Solar Orbiter, this should be 'UTC'. See Greisen and Calabretta in [FITS+WCS] for a complete list of possible values.	P	L1+
TRIGGERD	'STIXflag'	Event that triggered observation, default 'none'	Any string	String	Proposal, ITs to comment or propose alternatives: This keyword specifies whether the acquisition(s) was/were triggered by a flag, either raised internally or communicated through Inter-instrument	P	L1,2

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					communication (service 20). Acquisitions that were completely commanded from ground have default value 'none'.		

Campaign Information

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
SOOP_ID	'SOOP1.14'	SOOP campaign name	<i>TBD with SOOP naming convention</i>	String TBD	Identification of the SOOP this observation belongs too. It is TBD ⚠️ how we can link to multiple SOOPs. Options are (1) to allow a list of IDs separated by TBD symbol or (2) add multiple instances of the keyword. The source is the SOOP plan as agreed upon by SOWG and distributed by SOC at time of MTP.	P	L1+
TARGET	'AR'	Type of target from planning	'AR', 'QS', 'limb', 'CH', 'sun center', ⚠️ TBC	String	The target as defined in the SOOP description. The source is the SOOP plan as agreed upon by SOWG and distributed by SOC at time of MTP.	P	L1+

The source of this information should be the science plan, distributed to ITs after long-term and medium-term planning, to make the content consistent throughout the data products. It is TBD how/in what format this information will be passed on to instrument teams.

Description of Data Content

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
BSCALE*	1.0	ratio of physical to array value at 0 offset		Float	To avoid storing calibrated data as large floating-point numbers, a multiplicative factor BSCALE and an additive factor BZERO can be provided, to enable each data value to be represented as an integer ($N_{x,y}$). The true value of each pixel would then be $N_{x,y} * BSCALE + BZERO$. For raw data (counts), the default values of BSCALE and BZERO would be 1.000 and 0.000. More details on the use of these keywords in [FITSpaper] sect. 4.4.2.5.	O	L1+
BZERO*	0.0	physical value for the array value 0		Float	See description BSCALE.	O	L1+
BTYPE	'Intensity'	Type of data		String	Descriptive text for the type of data stored in the array.		
BUNIT*	'DN'	Units of physical value, after application of BSCALE, BZERO	Case sensitive. See [FITS+WCS]	String	Units of physical value, after application of BSCALE, BZERO. See Greisen and Calabretta in [FITS+WCS] for a complete description of how to form unit	P	L1+

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					strings.		
DATAMIN*	10.0	minimum valid physical value	min value in data (except BLANK)	Float	DATAMIN/MAX refer to min and max data values within the file. These keywords are proposed to be mandatory as it helps scaling images/data, however in combined datasets it is probably more appropriate in the table headers than in the primary header ⚠️ (TBC).	P	
DATAMAX*	3588.0	maximum valid physical value	max value in data (except BLANK)	Float		P	
BLANK*	-2147483648	Value marking undefined pixels (before the application of BSCALE, BZERO)	integer outside valid data range [DATAMIN, DATAMAX]	Integer	The value specified for keyword BLANK gives the data value that is used to mark undefined pixels, i.e. before BZERO and BSCALE are applied. It shall only be used in arrays with integer data, i.e. with positive values of BITPIX. More information in [FITSpaper] sect. 4.4.2.5. Each instrument team can define a value that fits their needs and should use it consistently. Blank pixels in floating point data are to be marked with the IEEE Not-a-Number (NaN) mechanism.	P	

Image Relative to Detector and Electronics

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
PXBEGn	1	First read-out pixel in dimension n	Positive, starting at 1	Integer	PXBEGn and PXENDn describe how subfields (in 2D images) and partial readouts in higher dimensional data cubes are handled. The first pixel that has been read out in dimension n is specified in PXBEGn, the last pixel in that dimension in PXENDn. For example, if PXBEG1=1 and PXEND1=NAXIS1, the detector has been fully read out in dimension 1. Note that PXBEG/END refer to pixels in unbinned images, in other words subfields are extracted before binning.	O	
PXENDn	512	Last read-out pixel in dimension n	Positive, smaller or equal to NAXISn	Integer	Value of the last pixel that has been read out in dimension n (see PXBEGn description).	O	
NBINn	2	on-board binning factor in dim <n>	Positive, starting at 1	Integer	NBINn specifies the binning factor in the n'th dimension that has been applied onboard, i.e. the actual number of pixels binned in that direction. Default value (no binning) is 1. If the rebinned pixels are kept as they are (not binning back), their resolution - as compared to the instrument's nominal resolution - can also be inferred from the resolution in arcsecs provided in the positional	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					parameters CUNITi below. However some instrument teams might also choose to always bin back rebinned images to their original (detector) resolution. This would be most appropriate at least for imagers that apply partial rebinning or different bins at different parts of the detector (e.g. METIS). In the latter option, it is necessary to describe in the metadata which binning was applied onboard. ★		
NBIN	4	total binning (product of NBINn)	Positive	Integer	The value of NBIN (or alternatively BINNING) is the product of all NBINn values above. ★	O	

★ Note that it is currently under discussion whether keywords which are associated to binning should be part of the metadata standard. If all instruments have very specific needs to specify binning parameters, it is probably best to leave it out of the standard. Instrument teams also have the option to bin their data back to the original (i.e. detector) resolution and describe the binning that took place onboard, or to produce a FITS file in the reduced resolution (no extra keyword needed in this case).

Some instrument-specific plans on rebinning are listed below:

- METIS: depending on the observation mode METIS applies partial rebinning of pixels, with different factors in the inner circle (2° or 2.5°) than in the outer part. Therefore METIS FITS files will be binned back to the detector resolution. A binning keyword will specify the 'binning map' ID that has been used onboard.
- SPICE: might bin with different factors along each dimension. Binned resolution will be kept as is in FITS files and binning factors will be specified with keywords NBINn and NBIN as in the table above.

Parameters Describing Onboard Processing

like IS_PROC and COMPRESS were discussed. They should only appear in this cross-instrument standard if their meaning can be defined and made sense for several instruments. Below an example open for discussion:

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
COMPRESS	'Lossless'	Data compression quality	choice among 5 possible words, see description	String	A cross-instrument description of the data compression quality: "None" - no compression applied "Lossless" - lossless compression "Lossy-high quality" - lossy compression applied. Data quality is mostly unaffected. "Lossy-strong" - strong lossy compression applied. Analysis of data should take into account frequent compression artefacts. "Lossy-extreme" - extreme lossy compression applied. Data quality is visibly affected		
COMPQUA	25.0	[%]	any	Float	Nominally, the		

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
L		Compressed/uncompressed quality ratio			compressed information content divided by the uncompressed information content, expressed as a percentage. The aim of a numeric value is to have a sortable item, so that users/programs can pick the best observation that satisfies any other (search) criteria.		

World Coordinate System (WCS) Attitude Parameters

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
WCSAXES	4	Number of WCS axes	>= NAXIS	Integer	Number of WCS axes. This number can be larger than NAXIS, with the understanding that the additional axes are degenerate, i.e. have a dimensionality of 1. If present, must	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					precede all other WCS keywords. An example use would be provide 2D pointing information for a single slit-spectrograph exposure with NAXIS=2.		
WCSNAME	'Helioprojective-cartesian'	Name of coordinate system		String	Name of coordinate system. This is typically 'Helioprojective-cartesian' for the typical image coordinate system which has solar north along the +Y axis.	P	
CTYPE1	'HPLN-TAN'	Helioprojective longitude (Solar X)		String	Coordinate name and projection along axis 1. CTYPE1[2] describe the nature of the images - the projection of the solar disk. 'HPLN' is the Helioprojective westward angle, and 'HPLT' is the Helioprojective northward angle. The units (arcsec)	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					are given in CUNIT1[2] . The projection is normally 'HPLT-TAN', though other values are possible for telescopes with very large fields of view.		
CTYPE2	'HPLT-TAN'	Helioprojective latitude (Solar Y)		String	Coordinate name and projection along axis 2.	P	
CUNIT1	'arcsec'	Units along axis 1.	Case sensitive. See [FITS+WCS]	String	Units along axis 1. See Greisen and Calabretta in [FITS+WCS] for a complete description of how to form unit strings.	P	
CUNIT2	'arcsec'	Units along axis 2.	Case sensitive. See [FITS+WCS]	String	Units along axis 2.	P	
PC1_1	0.8660254	WCS coordinate transformation matrix		Float	Coordinate transformation matrix. This could be used to describe the effect of the spacecraft roll angle, or more complex transformations. The example shown is for a rotation of 30	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					degrees.		
PC1_2	-0.5000000	WCS coordinate transformation matrix		Float	Coordinate transformation matrix.	P	
PC2_1	0.5000000	WCS coordinate transformation matrix		Float	Coordinate transformation matrix.	P	
PC2_2	0.8660254	WCS coordinate transformation matrix		Float	Coordinate transformation matrix.	P	
CDELTA1	3.17	[arcsec] Pixel scale along axis 1		Float	Pixel scale along axis 1. For projections that allow the pixel scale to vary with position, this is the pixel scale at the reference pixel location given by the CRPIXi keywords.	P	
CDELTA2	3.17	[arcsec] Pixel scale along axis 2		Float	Pixel scale along axis 2.	P	
CROTA	30.0	[deg] Rotation angle		Float	Data rotation angle expressed in the PCj_i matrix. Suitable mainly for 2D telescope images. Leaving out the axis number in	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					the name allows this to co-exist in FITS headers with the PCj_i matrix. In past missions, the bulk rotation of the image about the direction normal to the image plane have been described by CROTA1[2] . However, we propose to replace them by a more flexible and complete alternative which is to define the transformation matrix $s_i * m_{ij}$ using the keywords PCj_i allowing for a skewed axes and fully general rotation of multi-dimensional arrays.		
CRVAL1	0.0	[arcsec] Value of reference pixel along first axis.		Float	Value along first axis (X in example) at pixel location specified by CRPIXi keywords. Setting	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					CRVAL1=0, CRVAL2=0 means Sun center, and the reference pixel must be calculated accordingly.		
CRVAL2	0.0	[arcsec] Value of reference pixel along second axis.		Float	Value along second axis (Y in example) at pixel location specified by CRPIXi keywords.	P	
CRPIX1	512.5	[pixel] Reference pixel location along axis 1		Float	Reference pixel location along axis 1. Must be consistent with the CRVALi values. For a 1024x1024 image, setting the reference pixel to CRPIX1=512.5, CRPIX2=512.5 selects the center of the image, and the CRVALi values must be calculated accordingly.	P	
CRPIX2	512.5	[pixel] Reference pixel location along axis 2		Float	Reference pixel location along axis 2.	P	
CRDER1	0.2	Estimated random error	>= 0.0	Float	Representative average value	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
		along axis 1			over the data range of the random error in the coordinate values along axis 1, in the units given by CUNIT1 .		
CRDER2	0.2	Estimated random error along axis 2	≥ 0.0	Float	Representative average value over the data range of the random error in the coordinate values along axis 2, in the units given by CUNIT2 .	O	
CSYER1	0.1	Estimated systematic error along axis 1	≥ 0.0	Float	Representative average value over the data range of the systematic error in the coordinate values along axis 1, in the units given by CUNIT1 . The total error is the root-summed-square of the random and systematic errors.	O	
CSYER2	0.1	Estimated systematic error along axis 2	≥ 0.0	Float	Representative average value over the data range of the systematic error in the coordinate	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					values along axis 2, in the units given by CUNIT2 . The total error is the root-summed-square of the random and systematic errors.		
LONPOLE	180.0	[deg] Native longitude of the celestial pole		Float	Native longitude of the celestial pole. For azimuthal projections such as '-TAN', this defaults to 180 degrees, but may differ for other projections.	O	
SPECSYS	'HELIOCEN'	Reference frame for spectral axes	'HELIOCEN'	String	Reference frame used for calculating the effect of the spacecraft velocity on spectral coordinates. Used in conjunction with VELOSYS keyword.	O	
VELOSYS	-165.20536	[m/s] Radial velocity of spacecraft relative to Sun		Float	Radial velocity of the spacecraft relative to the Sun, in meters per second. Positive velocity is outward from	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					the Sun. Used in conjunction with SPECSYS keyword.		

In addition, alternative coordinate systems can be specified by adding sets of the above WCS keywords with a single letter of the alphabet appended, e.g. WCSNAMEA, CRPIX1A, etc. Each letter specifies a separate coordinate system, and all keywords for that coordinate system, even if they don't vary from the default coordinate system, should be included. (Note that SPECSYS and VELOSYS actually relate to the orbital position of the spacecraft, rather than the spacecraft attitude, but are included here because they also can have an alternate coordinate letter appended to them, e.g. **SPECSYSA**, **VELOSYSA**.)

Attitude Parameters (Example for Spectral Data)

The following example shows how the definitions above *could* be extended to a three-dimensional data cube containing spectral data.

```

NAXIS = 3
NAXIS1 = 1024 / Number of pixels in spectral dimension
NAXIS2 = 10 / Number of mirror positions in raster
NAXIS3 = 800 / Number of pixels along slit
WCSNAME = 'Helioprojective-cartesian'
CTYPE1 = 'WAVE' / Wavelength
CTYPE2 = 'HPLN-TAN' / Helioprojective longitude (Solar X)
CTYPE3 = 'HPLT-TAN' / Helioprojective longitude (Solar Y)
CUNIT1 = 'Angstrom ' / Units along spectral dimension
CUNIT2 = 'arcsec ' / Units along second axis (Solar X)
CUNIT3 = 'arcsec ' / Units along second axis (Solar Y)
CDELTA1 = 0.1 / [Angstrom] Wavelength scale
CDELTA2 = 4.0 / [arcsec] Scan step size
CDELTA3 = 1.1 / [arcsec] Slit pixel size, resolution along slit
CRVAL1 = 300. / [angstrom] Reference wavelength
CRVAL2 = 0.0 / [arcsec] Solar X coordinate of reference pixel
CRVAL3 = 30.0 / [arcsec] Solar Y coordinate of reference pixel
CRPIX1 = 512.5 / [pixel] Pixel coord. of reference wavelength
CRPIX2 = 1.0 / [pixel] Reference pixel is first exposure in raster
CRPIX3 = 400.5 / [pixel] Reference pixel is slit center
PC1_1 = 1.0 / No wavelength transformation
PC2_2 = 1.0 / No rotation of SPICE scan direction wrt solar north
PC3_3 = 1.0 / No rotation of SPICE scan direction wrt solar north

```

This group of keywords describes the relationship between the detector and target. All SolO metadata on position and attitude should be expressed in WCS coordinate systems. Helioprojective coordinates are celestial spherical coordinates that use

the sun to define their pole and origin [WCSpaper]. As this is a spherical coordinate system a map projection is used. The most common one in solar physics is the TAN projection, which is used when a solar image is focused onto a flat focal plane, such as an APS detector. Helioprojective-Cartesian coordinates, as chosen in the example above, are an angular coordinate system that measure latitude from disk centre towards solar North in theta_y, and longitude towards West limb in theta_x. Calibrated Level-2 images should be corrected for distortion and rotation effects and always be aligned with Solar North.

FITS headers can specify coordinates in multiple coordinate systems, and other useful coordinates like heliographic and heliocentric coordinate systems can be added. These can be combined in the way described by W. Thompson in Fig 4 of [WCSpaper].

Solar Ephemeris Data keywords (from SWAP and SECCHI)

All ephemeris values are to be calculated at the time DATE-AVG. (Note that, because two data sources have been combined together here, the values may not be consistent with each other.)

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
RSUN_ARC	952.565823288377	[arcsec] Apparent photospheric solar radius	Positive	Float	Apparent photospheric solar radius in arc seconds.	?	
RSUN_REFF	6.955508E+08	[m] Assumed physical solar radius	Positive	Float	Assumed photospheric solar radius, in meters. The intended purpose of this keyword is to document what assumption was made in determining the instrument plate scale, so that it can be backed out for comparison with measurements assuming a different solar radius.	O	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
SOLAR_B0	-4.18794601814401	[deg] s/c tilt of solar North pole			The tilt angle of the solar North pole toward (positive) or away from (negative) the spacecraft, in degrees. Equals HGLT_OBS , CRLT_OBS .		
SOLAR_P0	16.242648	[deg] s/c celestial North to solar North angle		Float	The apparent angle in the sky, as seen from the spacecraft location, between celestial north and solar north.		
SOLAR_EP	5.92376011723103	[deg] s/c ecliptic North to solar North angle		Float	The apparent angle in the sky, as seen from the spacecraft location, between ecliptic north and solar north.	?	
CAR_ROT	2158	Carrington rotation number		Integer	Carrington rotation number associated with pixel given by CRPIXi.		
HGLT_OBS	-4.18794601814401	[deg] s/c heliographic latitude (B0 angle)		Float	Stonyhurst heliographic latitude of the spacecraft, in	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					degrees.		
HGLN_OBS	0.00131206248030214	[deg] s/c heliographic longitude		Float	Stonyhurst heliographic longitude of the spacecraft, in degrees.	P	
CRLT_OBS	-4.18794601814401	[deg] s/c Carrington latitude (B0 angle)		Float	Carrington heliographic latitude of the spacecraft, in degrees. Because of the way the two coordinate systems are defined, HGLT_OBS and CRLT_OBS are always the same. However, both should appear in the FITS header.	P	
CRLN_OBS	25.6633839152	[deg] s/c Carrington longitude (L0 angle)		Float	Carrington heliographic longitude of the spacecraft, in degrees.	P	
DSUN_OBS	150707994180.746	[m] s/c distance from Sun	Positive	Float	Distance of the spacecraft from the center of the Sun, in meters.	P	
HEEX_OBS	150707994015.506	[m] s/c Heliocentric Earth Ecliptic X		Float	X axis position of the spacecraft in	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					Heliocentric Earth Ecliptic coordinates, in meters.		
HEEY_OBS	4059466.73665196	[m] s/c Heliocentric Earth Ecliptic Y		Float	Y axis position of the spacecraft in Heliocentric Earth Ecliptic coordinates, in meters.	P	
HEEZ_OBS	5772938.06008995	[m] s/c Heliocentric Earth Ecliptic Z		Float	Z axis position of the spacecraft in Heliocentric Earth Ecliptic coordinates, in meters.	P	
HCIX_OBS	109043378382.	[m] s/c Heliocentric Inertial X		Float	X axis position of the spacecraft in Heliocentric Inertial coordinates, in meters.	P	
HCiy_OBS	94332806329.1	[m] s/c Heliocentric Inertial Y		Float	Y axis position of the spacecraft in Heliocentric Inertial coordinates, in meters.	P	
HCIZ_OBS	-12321790987.1	[m] s/c Heliocentric Inertial Z		Float	Z axis position of the spacecraft in Heliocentric Inertial coordinates, in meters.	P	
HCIX_VO	-24982.158	[m/s] s/c		Float	Spacecraft	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
B		Heliocentric Inertial X Velocity			velocity along the X axis in Heliocentric Inertial coordinates, in meters per second. Keywords for velocities in other coordinate systems can also be formed by substituting _VOB for _OBS, but will not be inertial velocities, and are therefore optional.		
HCIZ_VO B	-17026.140	[m/s] s/c Heliocentric Inertial Y Velocity		Float	Spacecraft velocity along the Y axis in Heliocentric Inertial coordinates, in meters per second.	P	
HCIZ_VO B	2231.6757	[m/s] s/c Heliocentric Inertial Z Velocity		Float	Spacecraft velocity along the Z axis in Heliocentric Inertial coordinates, in meters per second.	P	
HAEX_OBS	-65400700049.4	[m] s/c Heliocentric Aries Ecliptic X		Float	X axis position of the spacecraft in	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					Heliocentric Aries Ecliptic coordinates, in meters.		
HAEY_OBS	129087613712.	[m] s/c Heliocentric Aries Ecliptic Y		Float	Y axis position of the spacecraft in Heliocentric Aries Ecliptic coordinates, in meters.	P	
HAEZ_OBS	-315707554.534	[m] s/c Heliocentric Aries Ecliptic Z		Float	Z axis position of the spacecraft in Heliocentric Aries Ecliptic coordinates, in meters.	P	
HEQX_OBS	-139113055426.	[m] s/c Heliocentric Earth Equatorial X		Float	X axis position of the spacecraft in Heliocentric Earth Equatorial coordinates, in meters.	P	
HEQY_OBS	37903753496.0	[m] s/c Heliocentric Earth Equatorial Y		Float	Y axis position of the spacecraft in Heliocentric Earth Equatorial coordinates, in meters.	P	
HEQZ_OBS	-12321790987.1	[m] s/c Heliocentric Earth Equatorial Z		Float	Z axis position of the spacecraft in Heliocentric Earth Equatorial	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					coordinates, in meters.		
GSEX_OBS	632918.814814495	[m] s/c Geocentric Solar Ecliptic X		Float	X axis position of the spacecraft in Geocentric Solar Ecliptic coordinates, in meters. <i>Note: There's some discussion in the WCS group on whether geocentric coordinates are useful for Solar Orbiter.</i>	P	
GSEY_OBS	-4059466.73664814	[m] s/c Geocentric Solar Ecliptic Y		Float	Y axis position of the spacecraft in Geocentric Solar Ecliptic coordinates, in meters.	P	
GSEZ_OBS	5772938.06009963	[m] s/c Geocentric Solar Ecliptic Z		Float	Z axis position of the spacecraft in Geocentric Solar Ecliptic coordinates, in meters.	P	
EAR_TIME	349.06169	[s] Time(Sun to Earth) - Time(Sun to S/C)		Float	Difference in light travel time between the Sun to Earth (heliocenter to geocenter) and the light	?	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					travel time from the Sun to the spacecraft (heliocenter to s/c), in seconds. Positive if the spacecraft-Sun distance is smaller than the Sun-Earth distance.		
SUN_TIME	149.59788	[s] Time(Sun to S/C)	Positive	Float	Light travel time from the Sun (heliocenter) to the spacecraft.	?	
DATE_EAR	'2020-01-22T14:05:49.062'	[UTC] Start time of observation, corrected to Earth	any date+'T'+time within mission duration	String	Start time of observation corrected for the difference in light travel between the Sun to Earth (heliocenter to geocenter) and the light travel time from the Sun to the spacecraft (heliocenter to s/c), in UTC, in ISO-8601 format. In other words, this is DATE-BEGIN <i>plus</i> EAR_TIME.		
DATE_START	'2020-01-22T13:57:30.402'	[UTC] Start time of	any date+'T'+ti	String	Start time of observation		

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
		observation, corrected to Sun	me within mission duration		corrected for the difference in light travel from the Sun to the spacecraft (heliocenter to s/c), in UTC, in ISO-8601 format. In other words, this is DATE-BEG <i>minus</i> SUN_TIME.		

Parameters Closing off Metadata Description

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
COMMENT*	'\$<comment text>\$'				This keyword may be used to supply any comments regarding the FITS file.	O	
CHECKSUM	'VVdCWTZ9VTbCVTZ9'	HDU checksum updated at <date/time>	character string	String	The value field of the CHECKSUM keyword shall consist of an ASCII character string whose value forces the 32-bit 1's complement checksum accumulated over the entire FITS HDU to equal negative 0. Documentation on how to generate the	P	

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					checksum in [FITS Checksum].		
DATASUM	'2189405276'	data unit checksum updated at <date/time >	character string	String	The value field of the DATASUM keyword shall consist of a character string containing the unsigned integer value of the 32-bit 1's complement checksum of the data records in the HDU (i.e., excluding the header records). Documentation on how to generate the data checksum in [FITS Checksum].	?	
HISTORY*	'<software name>'			String	This keyword should be used to describe the history of steps and procedures associated with the processing of the associated data.	P	
END					This keyword has no associated	M	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					value. Bytes 9 through 80 shall be filled with ASCII spaces (decimal 32 or hexadecimal 20). The END keyword marks the logical end of the header and must occur in the last 2880-byte FITS block of the header.		

3.1.2 FITS Extensions

Besides the primary header and data unit (HDU), a FITS file may contain one or more HDU extensions. These extensions are formatted in the same manner as the primary HDU, with a ASCII text header followed by a data unit. The following keywords apply to FITS extensions in general.

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
XTENSION	'IMAGE'	Extension type	'IMAGE ', 'TABLE ', 'BINTABLE'	String	Name of the type of extension being used. Must be one of the three currently defined conforming extensions, 'IMAGE ', 'TABLE ', or 'BINTABLE', must be the first keyword in the header, and must be padded out with spaces to be at least 8 characters long. <i>Cannot</i> appear in the primary HDU.	M	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
BITPIX	8	Number of bits per data unit.		Byte	Number of bits per data unit. Value may depend on the type of extension.	M	All
NAXIS	2	Number of axes in data cube		Byte	Number of data axes. Value may depend on the type of extension.	M	All
NAXISn	128	Number of data units along axis n.		Integer	Number of data units along axis n. This has different meanings in different extension types.	M	All
PCOUNT	0	Parameter count			Value depends on the type of extension.	M	All
GCOUNT	1	Group count			Value depends on the type of extension.	M	All
EXTNAME		Extension name		String	Name of the extension. May also appear in the primary HDU.	P	

Depending on the type of extension, some additional keywords may be mandatory, as discussed below.

3.1.2.1 Types of extensions

IMAGE extensions

By far the simplest type of extension is the IMAGE extension. This is formatted exactly the same as the primary HDU, with the exception that the **XTENSION** keyword replaces the **SIMPLE** keyword, and the header will contain the keyword/value combinations PCOUNT=0, GCOUNT=1. Otherwise, all the keywords defined for the primary header are equally applicable to IMAGE extensions, with the same requirements on which keywords are mandatory or optional. An example header for an IMAGE extension would be

```
XTENSION= 'IMAGE' / Extension type
BITPIX  = 16      / Number of bits per data pixel
NAXIS   = 2      / Number of data axes
```


NAXIS1 = 512 / Number of pixels along axis 1
 NAXIS2 = 512 / Number of pixels along axis 2
 PCOUNT = 0 / Parameter count
 GCOUNT = 1 / Group count
 EXTNAME = 'Helium II 304 Angstroms'
 ...
 END

Binary table (BINTABLE) extensions

A much more complicated, but very powerful, type of extension is the binary table extension, expressed by setting XTENSION='BINTABLE'. A binary table is organized as rows and columns. Each column has a data type, which may be multidimensional. The following primary keywords are used to set up the binary table.

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
XTENSION	'BINTABLE'	Extension type	'BINTABLE'	String	Specifies that what follows is a binary table.	M	All
BITPIX	8	Number of bits per data unit.		Byte	This is always 8 for binary tables. The actual data types within the table are specified separately.	M	All
NAXIS	2	Number of axes in data cube		Byte	This is always 2 for binary tables, to signify that the table is organized into columns and rows.	M	All
NAXIS1	128	Number of bytes in each data row.	Positive	Integer	Number of bytes in each data row, summed from all the columns in the table.	M	All
NAXIS2	10	Number of rows in the table	Positive	Integer	Number of rows in the table.	M	All
PCOUNT	0	Number of			Number of bytes	M	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
		bytes following table			that follow the table in the supplemental data area called the heap. This is to support the variable length arrays option.		
GCOUNT	1	Group count			This is always 1 for binary tables.	M	All
TFIELDS	5	Number of table columns	Positive	String	Number of columns in the table.	M	All
TFORMn	'20I'	Data format		String	String describing the format of the data in column n, where n runs from 1 to TFIELDS . The example '20I' represents a vector of 20 16-bit integer values. See [FITSpaper] for a complete description of how TFORMn keywords are formed.	M	All
TTYPEn	'Intensity'	Data type		String	Name of the data in column n.	P	All
TUNITn	'DN/s'	Data units		String	Units of the data in column n. See description for BUNIT keyword.	P	All
TDIMn	'(5,4)'	Data dimensions		String	String specifying the dimensions of multidimensional data. Not used if the data only have one dimension.	P	All

A number of the keywords defined for primary header have equivalents defined for use in binary table columns. These equivalents are given in the following table, where n represents the binary table column number, i and j represent axis numbers (from 1-9), and a represents an optional letter from A to Z signifying an alternate coordinate system.

Keyword Description	Primary Array	BINTABLE vector	
		Primary	Secondary
Coordinate dimensionality	WCSEXESa	WCAXna	
Axis type	CTYPEia	iCTYPn	iCTYna
Axis units	CUNITia	iCUNIn	iCUNna
Reference value	CRVALia	iCRVLn	iCRVna
Coordinate increment	CDELTia	iCDLTn	iCDEna
Reference point	CRPIXja	jCRPXn	jCRPna
Transformation matrix	PCi_ja	ijPCna	
Transformation matrix	CDi_ja	ijCDna	
Coordinate name	WCSNAMEa	WCSNna	
Random error	CRDERia	iCRDna	
Systematic error	CSYERia	iCSYna	
Coordinate native longitude	LONPOLEa	LONPna	
Date/time of observation	DATE-OBS	DOBSn	
Average date/time of observation	DATE-AVG	DAVGn	
Scaling parameter	BSCALE	TSCALn	
Zero offset	BZERO	TZEROn	
Null value for integers	BLANK	TNULLn	

Keywords which apply to the HDU as a whole, and which do not vary between columns, are put in the header in the normal fashion. For example, if all the data has the same average observation time, then **DATE-AVG** is placed in the header instead of the individual **DAVGn** for every column.

The FITS Green Bank convention allows keyword values to be stored in a column of the binary table. Such values then pertain all the columns in the table, but can vary by row. This is done by setting the value of **TTYPEn** to the name of the keyword. For example, if column 10 in the table had **TTYPE10** = 'DATE-BEG', then the contents of this column would have the values of the beginning date/time of the observation for each row in the table. The datatype of the column would

be appropriate for the keyword that it is storing. Thus, TTYPE10 = 'DATE-BEG' would be stored as an ASCII string column, while TTYPE11 = 'CRVAL1' would be stored as a floating pointing scalar column. More information about the FITS Green Bank convention is available at the Registry of FITS Conventions at http://fits.gsfc.nasa.gov/fits_registry.html.

3.2 CDF Metadata Standard

This section presents the data and metadata conventions for the Solar Orbiter CDF Level 1 (engineering data) and Level 2 (calibrated science data) data sets. The CDF format shall be used for time series data files, produced by the in situ instruments.

Metadata conventions for the CDF data sets shall be as much as possible compliant with the International Solar Terrestrial Physics (ISTP) guidelines [ISTPstandard]: http://spdf.gsfc.nasa.gov/istp_guide/istp_guide.html

Section 3.3.1 describes the CDF variable recommendations. Section 3.3.2 and 3.3.3 gives respectively the lists of global and variable attributes to be used.

In the global/variable attribute lists, grouped by metadata type, we distinguish 3 types of keywords:

- M: Mandatory keywords to comply to the ISTP guidelines.
- P: Keywords we propose to be mandatory in the Solar Orbiter CDF files.
- O: Keywords in *italic* are optional and can be defined if applicable to the instrument in question.

Any additional variables and attributes specific to instrument CDF data sets can be implemented as long as they are referenced in the dedicated TN document appendices.

3.2.1 CDF variables

In a CDF format file, each data parameter (e.g., time, flux, etc.) is represented by a given CDF variable. This CDF variable is identified by its name and a list of given variable attribute.

There are 3 types of CDF variables:

- "Data", which corresponds to the variables of primary importance (e.g, density, magnetic fields, flux).
- "Support_data", which corresponds to the variables of secondary importance (e.g., time, frequency, energy_bands).
- "Metadata", which provides labels for "data" and "support_data" types (e.g., "Bx, By, Bz" labels for a magnetic field components "data" variable).

3.2.1.1 General conventions

The general conventions for the CDF variables for Solar Orbiter are provided in the following list:

- CDF variables description and naming conventions shall be compliant with the ISTP guidelines. In addition, CDF variable names shall contain capital letters only and shall not exceed 63 characters.
- CDF variables shall be described using the appropriated variable attributes (see "Variable attributes" section below).
- CDF data files shall contain at least the expected variables listed in the section 1.1.5.

3.2.1.2 Time conventions

The time conventions for the CDF variables are provided in the following list:

- According to the ISTP guidelines, all of the CDF files shall contain at least one primary CDF variable called "EPOCH", which shall be the time reference variable. This variable shall be a CDF_TIME_TT2000 data type, and it shall be the first to be defined in the skeleton files.
- Data sets including multiple temporal resolutions shall use "EPOCH_i" (i=1,N) variables to define the multiple time samples.
- Level 1 and Low Latency data files shall provide the onboard S/C time values in a CDF variable "SCET".
- In addition to "EPOCH" and "SCET", level 1 and low latency data files should also contain a CDF variable named "ACQUISITION_TIME". This variable can be used to store the CCSDS format time values as written in the TM packet data.
- Any other time variables derived from "EPOCH" (e.g., milliseconds of day, julian days, etc.) can be provided using the CDF variable conventions.

It should be noted that the CDF_TIME_TT2000 data type is defined as an 8-byte signed integer with a fixed Time_Base=J2000 (Julian date 2451545.0 TT or 2000 January 1, 12h TT), Resolution=nanoseconds, Time_Scale=Terrestrial Time (TT), Units=nanoseconds, Reference_Position=[Spacecraft barycenter](#). Given a current list of leap seconds, conversion between TT and UTC is straightforward ($TT = TAI + 32.184s$; $TT = UTC + \text{deltaAT} + 32.184s$, where deltaAT is the sum of the leap seconds since 1960; for example, for 2009, deltaAT = 34s).

3.2.1.3 Coordinate system conventions

The coordinate system conventions for the CDF variables are as follows:

- Level 2 vector and tensor time series should always be presented in a scientific coordinate system. Spacecraft-centric RTN coordinates in a Cartesian representation are the preferred choice for Solar Orbiter. Other allowed coordinate systems are defined below. Further variables containing the same data in spacecraft or instrument coordinates are optional. In the event that instrument rather than spacecraft coordinates are included, the rotation matrix from instrument to spacecraft coordinates should also be included in the file.
- Level 1 data should be presented in instrument coordinates, together with the rotation matrix to spacecraft coordinates.
- Coordinate systems will be described by four variable attributes: COORDINATE_SYSTEM which gives the name of the coordinate system; FRAME_VELOCITY which can take either the value 'Observatory' where no corrections have been applied to the data or 'Inertial' where quantities (e.g. electric field or plasma flow velocity) have been corrected for spacecraft motion

relative to an inertial frame (HCI); FRAME_ORIGIN which gives the origin of the reference frame where this is not implicit in the value of COORDINATE_SYSTEM; and REPRESENTATION_i which gives the representation (['x','y','z'] for Cartesian; ['r','p','t'] for spherical polar; ['r','p','z'] for cylindrical polar) of the ith dimension of the variable.

3.2.1.4 Spacecraft position

It is suggested that at least the low cadence, normal mode, Level 2 files should also include the spacecraft position as a variable. This should be expressed in HCI Cartesian Coordinates and be treated in the same way as any other variable.

3.2.1.5 Data quality

CDF files shall include the QUALITY_FLAG and BITMASK variables to inform on data quality. They are both time dependent parameters and can apply to one or several science parameters.

QUALITY_FLAG shall be a CDF_UINT1 flag providing a human readable high level information about the quality.

Quality	Meaning
0	Bad data
1	Known problems, use at your own risk
2	Survey data, possibly not publication-quality
3	Good for publication, subject to PI approval
4	Excellent data which has received special treatment

BITMASK shall be a CDF_UINT2 flag providing a computer readable quality information. The values and meanings of BITMASK can differ from instrument to another and shall be reported in the instrument specific requirements appendices.

3.2.1.6 Expected variables

The following table provides a summary list of the variables expected to be found in the Solar Orbiter CDF data files. The "LEVEL" column indicates the processing level(s) for which the variable is mandatory or optional.

NAME	DESCRIPTION	TYPE	LEVEL	DATA TYPE	COMMENT
EPOCH	Primary time variable as defined in the section	M	All	CDF_TIME_TT2000	See http://spdf.gsfc.nasa.gov/istp_guide/variables.html#Epoch .

NAME	DESCRIPTION	TYPE	LEVEL	DATA TYPE	COMMENT
	1.1.2				
EPOCH_i	i-th time variable as defined in the section 1.1.2	M	All	CDF_TIME_TT2000	Only mandatory for multiple temporal resolutions.
SCET	Onboard S/C time	P	L1, LL	TBD	
ACQUISITION_TIME	Time of acquisition as returned in the instrument packet data	O	L1, LL	TBD	Not necessarily the same than SCET
JULIAN_DAY	Time in julian days	O	L2, L3	CDF_REAL8	
SYNCHRO_FLAG	Flag to check instrument time synchronisation	O	L1, LL	CDF_UINT1	e.g., Check time synchronisation between the RPW DPU time and sub-systems times.
QUALITY_FLAG	Human readable high level parameter	P	All	CDF_UINT1	Can apply to one or several science parameters
BITMASK	Computer readable quality parameter	P	All	CDF_UINT2	Can apply to one or several science parameters
POST_GAP	Flag that	O	All	CDF_UINT1	See

NAME	DESCRIPTION	TYPE	LEVEL	DATA TYPE	COMMENT
_FLAG	indicates data gap in the records.				http://spdf.gsfc.nasa.gov/istp_guide/variables.html#Post%20Gap%20Flag
INTERPOL_FLAG	Flag that indicates if the current record is real or interpolated	O	All	CDF_UINT1	
SC_HCI_POS	S/C Cartesian coordinates in the HCI system.	P	L2, L3	TBD	Only mandatory for low cadence, level 2/3.
SC_HEE_POS	S/C Cartesian coordinates in the HEE system	O	L2, L3	TBD	Only mandatory for low cadence, level 2/3.
SBM_FLAG	Flag that indicated if a SBM mode is currently on/off.	O	All	CDF_UINT1	

3.2.2 Global attributes

Global attributes are used to provide information about the data set as an entity. Global attributes for Solar Orbiter CDF data sets are divided into 4 categories:

- Attributes defined by the ISTP (http://spdf.gsfc.nasa.gov/sp_use_of_cdf.html)
- Attributes associated to the Space Physics Archive Search and Extract dictionary (SPASE, <http://www.spase-group.org/>)

- Attributes derived from the Virtual European Solar and Planetary Access (VESPA, <http://voparis-europlanet-new.obspm.fr/planetary/data/eqn/query/all/>)
- Additional attributes specific to Solar Orbiter CDF data sets.

All of the global attributes shall be CDF_CHAR data types.

3.2.2.1 ISTP attributes

The following table provides the list of ISTP global attributes that should be used for Solar Orbiter CDF data sets.

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
Project	The name of the project.	M	"ESAC>European Space Astronomy Center"	It can allow more than one entry
Source_name	The mission or investigation that contains the sensors. It shall use the ISTP format "PREFIX>Suffix" (e.g., "SOLO>Solar Orbiter"), where the value of the prefix shall correspond to the "source" field in the file naming convention.	M	"SOLO>Solar Orbiter"	
Discipline	The science discipline and subdiscipline.	M	"Space Physics>Interplanetary Studies"	

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
Data_type	ISTP defined exchangeable data product type.	M		See http://spdf.gsfc.nasa.gov/istp_guide/gattributes.html#Data_type
Descriptor	The name of the instrument and sensor or detector, separated by a hyphen, that collected the data. (e.g., "RPW-TDS>Radio and Plasma Waves, Time Domain Sampler", "SWA>Solar Wind Analyzer", ...) It shall use the ISTP format "PREFIX>Suffix" (e.g., "MAG>Magnetometer"), where the value of the prefix shall correspond to the "descriptor" field in the file naming convention.	M		
Data_version	This attribute identifies the	M		It shall be an integer increment starting at 01.

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
	version of a particular CDF data file for a given date. The value of the version shall be consistent with the "version" field in the filename.			
Software_version	The version of the software that generated the CDF.	P		
Skeleton_version	The skeleton file version number.	O		It shall be an integer increment starting at 01.
PI_name	First initial and last name of the PI.	M		e.g., "M.Maksimovic"
PI_affiliation	A recognizable abbreviation of the PI affiliation.	M		e.g., "LESIA, Observatoire de Paris-CNRS".
TEXT	Description of the experiment.	M		It shall be used to describe the instrument and to provide references. The first entry shall be reserved for the description, followed by as many entries as references. Reference entries shall be sorted by decreasing importance.
Instrument_type	The ISTP defined	M		See http://spdf.gsfc.nasa.gov/istp_guide/gattribution

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
	instrument type. Multi-valued.			s.html#Instrument_type
Mission_group	The assigned name of the mission or project.	M	"Solar Orbiter"	
Logical_source	Source_name, data_type, and descriptor information.	M		See http://spdf.gsfc.nasa.gov/istp_guide/gattribution.html#Logical_source
Logical_file_id	The name of the CDF file using the ISTP naming convention.	M		See http://spdf.gsfc.nasa.gov/istp_guide/gattribution.html#Logical_file_id
Logical_source_description	Full words associated with the Logical_source.	M		See http://spdf.gsfc.nasa.gov/istp_guide/gattribution.html#Logical_source_description
File_naming_convention	Description of the file naming convention.	O		It shall be consistent with the file naming convention.
Rules_of_use	Citability and PI access restrictions. This may point to a World Wide Web page	M		

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
	specifying the rules of use.			
Generated_by	The generating data center/group.	M		The name of the Institute that produce the data file. (TBC)
Generation_date	Date stamp for the creation of the file.	M		It shall be the data file creation date in ISO8601 format ("YYYY-MM-DDTHH:MN:SS") (TBC)
Acknowledgement	Text string at PI disposal allowing for information on expected acknowledgment if data is citable.	M		
MODS	History of modifications made to the CDF data set.	P		A new entry shall be added each time a modification is made to the CDF data set. Each entry has to specify the date of the modifications, the person responsible and a summary of the changes. (TBC)
LINK_TEXT	Text describing on-line data available at PI or Co-I web sites.	O		TBD
LINK_TITLE	The title of the web site holding on-line data available at PI or Co-I web sites.	O		TBD
HTTP_LINK	The URL for the PI or Co-I	P		TBD

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
	web site holding on-line data.			
TEXT_supplement_1	An attribute that can be used for providing additional information about dataset.	O		
Parents	This attribute lists the parent CDF(S) for files of derived and merged data sets. Subsequent entry values are used for multiple parents. The syntax for a CDF parent would be e.g. "CDF>logical_file_id".	O		

3.2.2.2 SPASE attributes

The following table provides the list of SPASE (<http://www.spase-group.org/>) global attributes for Solar Orbiter CDF data sets.

NAME	DESCRIPTION	TYPE	COMMENT
CAVEATS	Information which may be important in the avoidance of the misuse of the resource, for instance the assumptions or limitations on data processing modelings or inversions.	O	
spase_DatasetResourceID	The SPASE ResourceID assigned of the	O	

NAME	DESCRIPTION	TYPE	COMMENT
	NumericalData resource the data file is part of.		
spase_DatasetResource	The SPASE XML description of the dataset that corresponds to the SPASE ResourceID.	O	
spase_GranuleResourceID	The Granule ResourceID assigned to the data file.	O	
spase_GranuleResource	The SPASE XML description of the dataset that corresponds to the SPASE Granule ResourceID	O	

3.2.2.3 VESPA attributes

The following table provides the list of the Solar Orbiter CDF global attributes derived from the VESPA keywords.

NAME	DESCRIPTION	TYPE	COMMENT
TARGET_NAME	Name of the target	P	e.g., 'Sun'
TARGET_CLASS	Class of the target	P	e.g., 'Star'
TARGET_REGION	Region where the target in the observed	P	e.g., 'Corona'
ACCESS_FORMAT	Format of the file	O	e.g., 'CDF' or 'CDF V3.5'
DATAPRODUCT_TYPE	High level scientific organization of the data	O	
TIME_MIN	The date and time of the beginning of the first acquisition for the data contained in the file	P	
TIME_MAX	The date and time of the end of the last acquisition for the data contained in the file	P	
REFERENCE	Bibcode, DOI or URL	O	

More details about the VESPA keywords can be found in the appendix [VESPA keywords for Unified Solar Orbiter Catalog \(USOC\)](#).

3.2.2.4 Additional attributes

The following table provides the list of additional global attributes for Solar Orbiter CDF data sets.

NAME	DESCRIPTION	TYPE	COMMENT
------	-------------	------	---------

NAME	DESCRIPTION	TYPE	COMMENT
Data_product	Type of data product. It shall use the ISTP format "PREFIX>Suffix" (e.g., "HIST1D>1D histogram"), where the prefix shall correspond with the "dataprodu" field in the file naming convention.	P	
Free_field	Description of the free field in the filename. It shall use the ISTP format "PREFIX>Suffix" (e.g., "NORM>Normal mode"), where the value of the prefix shall correspond to the "free" field in the file naming convention.	P	Only mandatory if the "free" field is provided in the file name.
SOOP_ID	SOOP campaign identifier	P	It shall be one entry per SOOP. It shall let empty if there is no SOOP in the file.
Dataset_ID	Data set identifier	O	It can refer to a unique instrument data set identifier.
Process_ID	Process identifier	O	It can refer to a unique process identifier in a instrument data processing pipeline.
Job_ID	Processed job identifier	O	It can refer to a unique processed job identifier in a instrument data processing pipeline.
LEVEL	Data processing level as defined in the Solar Orbiter conventions. It shall use the ISTP format "PREFIX>Suffix" (e.g., "L1>Level 1 data processing"), where the value of the prefix shall correspond to the "level" field in the file naming convention.	P	

3.2.3 Variable attributes

Variable attributes are linked with each individual variable, and provide additional information about each variable.

Solar Orbiter CDF data files shall use the variable attributes listed in the following tables.

3.2.3.1 ISTP attributes

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
FIELDNAM	holds a character string (up to 30 characters) which describes the variable.	M	CDF_CHAR	
CATDESC	is an approximately 80-character string which is a textual description of the variable and includes a description of what the variable depends on.	M	CDF_CHAR	
DEPEND_0	explicitly ties a data variable to the time variable on which it depends. All variables which change with time must have a DEPEND_0 attribute defined. The value of DEPEND_0 is 'Epoch', the time ordering parameter for ISTP/IACG.	M	CDF_CHAR	
DEPEND_1	ties a dimensional data variable to a support_data variable on	M	CDF_CHAR	Only mandatory for 2d variables

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
	which the 2-th dimension of the data variable depends.			
DEPEND_2	ties a dimensional data variable to a support_data variable on which the 3-th dimension of the data variable depends.	M	CDF_CHAR	Only mandatory for 3d variables
DEPEND_3	ties a dimensional data variable to a support_data variable on which the 4-th dimension of the data variable depends.	M	CDF_CHAR	Only mandatory for 4d variables
DISPLAY_TYPE	tells automated software what type of plot to make and what associated variables in the CDF are required in order to do so.	M	CDF_CHAR	Only mandatory for data variable
FILLVAL	is the number inserted in the CDF in place of data values that are known to be bad or missing. Fill data are always non-valid data.	M	Variable data type	Only mandatory for time varying variables. See http://spdf.gsfc.nasa.gov/istp_guide/vattributes.html#FILLVAL for allowed values.

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
FORMAT	is the output format used when extracting data values out to a file or screen.	M	CDF_CHAR	Only mandatory if not using FORM_PTR
LABLAXIS	should be a short character string (approximately 10 characters, but preferably 6 characters - more only if absolutely required for clarity) which can be used to label a y-axis for a plot or to provide a heading for a data listing.	M	CDF_CHAR	Only mandatory if not using LABL_PTR
LABL_PTR_1	is used to label a dimensional variable when one value of LABLAXIS is not sufficient to describe the variable or to label all the axes.	M	CDF_CHAR	Only mandatory if not using LABLAXIS
LABL_PTR_2	is used to label a dimensional variable when one value of LABLAXIS is not sufficient to describe the variable or to label all the axes.	M	CDF_CHAR	Only mandatory if not using LABLAXIS

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
LABL_PTR_3	is used to label a dimensional variable when one value of LABLAXIS is not sufficient to describe the variable or to label all the axes.	M	CDF_CHAR	Only mandatory if not using LABLAXIS
UNITS	is a character string (no more than 20 characters, but preferably 6 characters) representing the units of the variable.	M	CDF_CHAR	Only mandatory if not using UNIT_PTR
UNIT_PTR	has as its value a variable which stores the character strings (up to 20 characters per character string) representing the units of the original variable, which can be added to a data listing heading or plot label.	M	CDF_CHAR	Only mandatory if not using UNITS
VALIDMIN	holds value which is the minimum value for a particular variable that are expected over the lifetime of the	M	Variable data type	Only mandatory for time varying data and support_data

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
	mission.			
VALIDMAX	hold values which is the maximum value for a particular variable that are expected over the lifetime of the mission.	M	Variable data type	Only mandatory for time varying data and support_data
SCALEMIN	is value which can be based on the actual value of data found in the CDF data set or on the probable uses of the data, {lem e.g.}, plotting multiple files at the same scale.	M	Variable data type	Only mandatory for time varying data and support_data
SCALEMAX	is value which can be based on the actual values of data found in the CDF data set or on the probable uses of the data, {lem e.g.}, plotting multiple files at the same scale.	M	Variable data type	Only mandatory for time varying data and support_data
VAR_TYPE	identifies a variable.	M	CDF_CHAR	
SCALETYP	indicates whether the variable should have a linear	M	CDF_CHAR	Only mandatory for data and support_data if not using SCAL_PTR

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
	or a log scale as a default.			
SCAL_PTR	is used for dimensional variables when one value of SCALTYP is not sufficient.	M	CDF_CHAR	Only mandatory for data and support_data if not using SCALETYP
VAR_NOTES	holds ancillary information about the variable and can be any length.	M	CDF_CHAR	
MONOTON	indicates whether the variable is monotonically increasing or monotonically decreasing.	O	CDF_CHAR	
TIME_BASE	Reference frame where the time is measured (e.g., J2000)	P	CDF_CHAR	Only mandatory for time variables
TIME_SCALE	Scale of the time (e.g., Terrestrial Time, UTC)	P	CDF_CHAR	Only mandatory for time variables
REFERENCE_POSITION	Position in the reference frame where the time is measured (e.g., Rotating Earth Geoid, Spacecraft barycenter, Sun barycenter)	P	CDF_CHAR	

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
ABSOLUTE_ERROR	Value of the absolute error	O	Variable data type	
RELATIVE_ERROR	Value of the relative error	O	Variable data type	
FORM_PTR	has as its value a variable which stores the character strings (up to 20 characters per character string) representing the desired output format for the original variable.	M	CDF_CHAR	Only mandatory if not using FORMAT

3.2.3.2 Additional attributes

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
SI_CONVERSION	A string that defines the conversion needed to base SI units, e.g. "1.0E-9>T" for DC Magnetic field data in nT.	P	CDF_CHAR	
COORDINATE_SYSTEM	For nonscalar data, could contain coordinate system and whether cartesian/polar e.g. 'GSE_XYZ', 'HCI_RTP'.	P	CDF_CHAR	Only mandatory for coordinate system variables
FRAME_ORIGIN	gives the origin of the reference frame where this is not	P	CDF_CHAR	Only mandatory for coordinate system variables

NAME	DESCRIPTION	TYPE	CDF DATA TYPE	COMMENT
	implicit in the value of COORDINATE_SYSTEM			
FRAME_VELOCITY	can take either the value 'Observatory' where no corrections have been applied to the data or 'Inertial' where quantities (e.g. electric field or plasma flow velocity) have been corrected for spacecraft motion relative to an inertial frame (HCI)	O	CDF_CHARACTER	Only mandatory for coordinate system variables
REPRESENTATION_i	gives the representation (['x','y','z'] for Cartesian; ['r','p','t'] for spherical polar; ['r','p','z'] for cylindrical polar) of the ith dimension of the variable.	P	CDF_CHARACTER	Only mandatory for coordinate system variables
UCD	UCD keywords as defined by the IVOA	O	CDF_CHARACTER	http://www.ivoa.net/documents/latest/UCD.html

4 APPENDICES

4.1 UCDS

4.1.1 Discussion on Unified Content Descriptors (UCD)

UCD stands for *Unified Content Descriptor*. It is an IVOA (International Virtual Observatory Alliance) standard. Generic documentation can be found [here](#). The introduction of the UCD standard description is recalled here:

The Unified Content Descriptor (UCD) is a formal vocabulary for astronomical data that is controlled by the International Virtual Observatory Alliance (IVOA). The vocabulary is *restricted* in order to avoid proliferation of terms and synonyms, and *controlled* in order to avoid ambiguities as far as possible. It is intended to be flexible, so that it is understandable to both humans and computers. UCDS describe astronomical quantities, and they are built by combining words from the controlled vocabulary.

A UCD does not define the units nor the name of a quantity, but rather "*what sort of quantity is this?*"; for example **phys.temperature** represents a temperature, without implying a particular unit.

It would be possible to describe astronomical data quantities in a natural language such as English or Hungarian or Uzbek; however, it would be very difficult to expect a machine to 'understand' it in any sense. At the opposite extreme, there is an attempt within the IVOA to describe astronomical data in terms of a hierarchical data model, so that there is a place for everything, and everything is in its place. The UCD vocabulary falls between these extremes, and is intended to be understandable to both humans and computers.

Other explanations can be found at [CDS](#).

There are other standards available for describing the quantity related to a parameter in a data product. For instance, there is the [SPASE \(Space Physics Archive Search and Extract\) "Measurement Type" keyword](#), which is well adapted for space physics measurement, but still very limited. There is also a series of keywords in the NASA/PDS data dictionary that could be used here, but it is not a stable standard.

4.1.2 Building UCDS

A UCD is a string of character. The UCD syntax is described in the [UCD IVOA documentation](#).

There are several online resources to help the user or data provider selecting the correct UCD value.

- [Current list of UCD \(v1.23\)](#)
- [CDS UCD builder](#) (based on syntax analysis)
- [GAVO UCD resolver](#) (based on existing UCD/description occurrences)

NB When trying to build UCDs, it is very common to try to set up a very precise and accurate description of the parameter. This is not the goal of UCDs. The description should stay at a very simple level. Trying to go into too much details usually results in trying to use the UCD concept as a data model for your physical parameters. As UCDs were not meant for this, it doesn't work.

4.1.3 Example of UCDs

Instrument	Parameter Name	Parameter Description	UCD	Notes
EUI	EUI-HRI-17nm	EUI-HRI-17nm measures the photon count per pixel in a EUV band around 17nm.	phot.count;em.UV.10-50nm	check on CDS UCD builder with query "photon count at 17 nm"
		At higher data levels we will normalize with respect to the exposure time	phot.flux;em.UV.10-50nm	
RPW	RPW-HFR-L2	Spectral flux density in the radio range (<20 MHz) using electrical sensors	phys.flux.density;phys.electField;em.radio	

4.2 Instrument-Specific Requirements

4.2.1 Instrument-specific Requirements - TEMPLATE

4.2.1.1 Note on the Template

Copy the complete template document tree and rename this page to "<instrument> (IS or RS)", e.g. "STIX (RS)". See [STIX \(RS\)](#) for an actual example.

4.2.1.2 Issues and Conflicts with Proposal

This section is used to point out possible conflicts between instrument-specific requirements and the current metadata proposal or to add information to the proposal that may affect other instruments (e.g. file naming, versioning, etc.). For "non-controversial" requirements or FITS/CDF metadata keywords, use other sections.

Data Formats and File Names

FITS/CDF Metadata

Other

4.2.1.3 FITS Metadata Standard

Section (e.g. Basic Keywords)

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	D-Type	Level
SIMPLE	T	file conforms to basic FITS standards	T (for true)	Boolean	SIMPLE is a logical keyword specifying whether the file conforms to basic FITS standards; Solar Orbiter data should always comply, so T (for True) is only valid	<instrument>, e.g. STIX	<data type>	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	D-Type	Level
					option here.			

Extra (e.g. engineering, team, etc.)

4.2.1.4 CDF Metadata Standard

Section

NAM E	DESCRIP TION	TYPE	LEV EL	DATA TYPE	COMMENT
EPO CH	Primary time variable as defined in the section 1.1.2	<instrum ent>, e.g. RPW	All	CDF_TIME_T T2000	See http://spdf.gsfc.nasa.gov/istp_guide/variables.html#Epoch .

4.2.1.5 Data Products

Science

Low Latency

4.2.2 EUI specific issues

Disclaimer: none of the present content has been discussed & approved by the EUI consortium. This is merely an early draft of pending issues

4.2.2.1 Filename convention

Following the guidelines in [Updated Proposal for Filename Convention \(09/12/2014\)](#):

- we store 1 image per FITS file (to avoid the complexity that FITS keywords become vectors)
- the third field (descriptor) should be INSTRUME-DETECTOR, which can have the values EUI-FSI, EUI-HR1174 or EUI-HRILya.
- the forth field should be the optional field "data product". For EUI-FSI one could put here the filter (304/174) but there is no equivalent for the EUI-HRILya. In order to keep the filenames homogenous we can therefore add 304 or 174 in the descriptor field: solo_L1_eui-fsi304_date...

, solo_L1_eui-fsi174_date..., solo_L1_eui-hri174_... and omit the optional field 'data product' everywhere.

- in the CAL data level, we could have data products LED, DARK but also FLAT, DEADPIX, ... We also have to decide what to do with the CONTINUOUS READ-OUT mode. In the higher level data products (>=L3) one could have data products JHELIO or ANNOTATED,...
- A raw (L0) EUI/HRI image from a sequence operating at 10Hz imaging sequence that make use of the free field (see next bullet) would look like

SOLO_L0_EUI-HRI174_20201108T121230899_V01_F11234.fits

- we will make use of the free format field F to identify uniquely source images: All filenames originating from the same image onboard can be traced back by doing "ls */solo*F11234.fits"

Solar Orbiter Optional FITS Keywords

TBW

4.2.2.2 EUI specific FITS keywords

Onboard Science Tables

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
SECTABID	'DiscoveryD'	Secondary Table Identifier			Name of program in the secondary table that created this image		
TABEXPID	'DD03'	Table Exposure Identifier		string	Identifier that allows to trace back which line in the secondary table created this image from image metadata		

Mechanisms

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
DOORINT					internal EUI door		
DOOREXT					external heat		

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					shield door		
FILTER	'304_1'		FSI: "304_1", "304_2", "174_1", "174_2" HRI174: ?	string	Position of FILTER wheel (not for HRILya).		
						M	All

Front end electronics and detector settings

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
LEDA	'on', 'off'						
LEDB	'on', 'off'						
HVPU1							
HVPU2							
DETCLEAR	1	Detector Clears	0,1	bit	Always 1 except for Continous Readout (=0)		

Compression settings

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
REBIN							
RECODING							
BIAS							
NUM							

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
...							
						M	All

Single Image Processing

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
COSMICR			on, off		Cosmic Ray Removal		
BADPIXFIX			on, off		Bad Pixel Correction		
TBW offsets & gains							
						M	All

Multi Image Processing

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
						M	All

Data Routing

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
PRIORITY			0-255		Priority Number		
APID					Onboard	M	All

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
					APID		

4.2.3 RPW (IS)

4.2.3.1 RPW

This appendix addresses specific requirements for the RPW instrument.

4.2.3.2 RPW Data Products Formats

This section presents the data products formats for the RPW instrument.

Data Formats and File Names

Following the guidelines of the Solar Orbiter file naming convention:

- The RPW science data for survey modes shall be stored in day granularity CDF files (i.e., `datetime="yyymmdd"`).
- The RPW science data for detected events in Selected Burst Modes (SBM) shall be stored in millisecond granularity CDF files (i.e., `datetime="yyyymmddThhmmssxyz-yyyymmddThhmmssxyz"`), where the start and end datetimes correspond respectively to the beginning and the end of the high cadence time window around the events.

4.2.3.3 CDF Metadata Standard

CDF Variables

BITMASK

The table below lists the BITMASK values and their meanings for the RPW instrument.

Bit	Meaning
0	Reset
1	Bad bias
2	Bias sweeping mode
3	Calibration mode
4	Sensor failure

Bit	Meaning
5	Non-standard operation

Global attributes

NAME	DESCRIPTION	TYPE	DEFAULT VALUE	COMMENT
Spectral_range_min	Minimum value of the spectral range within the file in Hz.	O		It corresponds to the "spectral_range_min" VESPA keyword.
Spectral_range_max	Maximum value of the spectral range within the file in Hz.	O		It corresponds to the "spectral_range_max" VESPA keyword.

4.2.3.4 Data Products

Science

The table below gives the list of science data products for RPW. It includes the values of the corresponding "Descriptor", "Data_product", "Free_field" global attributes.

Product name	Description	Descriptor	Data_product	Free_field
RPW_TNR_NORMAL_PS	TNR Power spectrum	"RPW-TNR>RPW Thermal Noise Receiver"	"PS>Power Spectrum"	"NORM>Normal mode"
RPW_HFR_NORMAL_PS	TNR Power spectrum	"RPW-HFR>RPW High Frequency Receiver"	"PS>Power Spectrum"	"NORM>Normal mode"
RPW_TNR_BURST_PS	TNR Power spectrum	"RPW-TNR>RPW Thermal Noise Receiver"	"PS>Power Spectrum"	"BURST>Burst mode"

Product name	Description	Descriptor	Data_product	Free_field
RPW_HFR_BURST_PS	TNR Power spectrum	"RPW-HFR>RPW High Frequency Receiver"	"PS>Power Spectrum"	"BURST>Burst mode"
RPW_TDS_NORMAL_RSFW	TDS Regular Snapshot Waveform data in normal mode	"RPW-TDS>RPW Time Domain Sampler"	"RSWF>Regular Snapshot Waveform"	"NORM>Normal mode"
RPW_TDS_SBM1_RSFW	TDS Regular Snapshot Waveform data in SBM1 mode	"RPW-TDS>RPW Time Domain Sampler"	"RSWF>Regular Snapshot Waveform"	"SBM1>Selected Burst Mode 1"
RPW_TDS_NORMAL_TSWF	TDS Triggered Snapshot Waveform data in normal mode	"RPW-TDS>RPW Time Domain Sampler"	"TSWF>Triggered Snapshot Waveform"	"NORM>Normal mode"
RPW_TDS_SBM2_TSWF	TDS Triggered Snapshot Waveform data in SBM2 mode	"RPW-TDS>RPW Time Domain Sampler"	"TSWF>Triggered Snapshot Waveform"	"SBM2>Selected Burst Mode 2"
RPW_TDS_NORMAL_STAT	TDS Statistics information	"RPW-TDS>RPW Time Domain Sampler"	"STAT>Statistics"	"NORM>Normal mode"
RPW_TDS_NORMAL_HIST1D	TDS Histogram 1D	"RPW-TDS>RPW Time Domain Sampler"	"HIST1D>Histogram 1D"	"NORM>Normal mode"

Product name	Description	Descriptor	Data_product	Free_field
RPW_TDS_NORMAL_HIST2D	TDS Histogram 2D	"RPW-TDS>RPW Time Domain Sampler"	"HIST2D>Histogram 2D"	"NORM>Normal mode"
RPW_TDS_LFM_RSWF	TDS Regular Snapshot Waveform in LFM mode	"RPW-TDS>RPW Time Domain Sampler"	"RSWF>Regular Snapshot Waveform"	"LFM>Low Frequency mode"
RPW_TDS_LFM_CWF	TDS Continuous Waveform in LFM mode	"RPW-TDS>RPW Time Domain Sampler"	"CWF>Continuous WaveForm"	"LFM>Low Frequency mode"
RPW_TDS_LFM_PSD	TDS Single Power Spectrum in LFM mode	"RPW-TDS>RPW Time Domain Sampler"	"PSD>Single Power Spectrum"	"LFM>Low Frequency mode"
RPW_TDS_LFM_SM	TDS Spectral Matrix in LFM mode	"RPW-TDS>RPW Time Domain Sampler"	"SM>Spectral Matrix"	"LFM>Low Frequency mode"
RPW_TDS_NORMAL_MAMP	TDS continuous HF signal maximum in normal mode	"RPW-TDS>RPW Time Domain Sampler"	"MAMP>Maximum Power"	"NORM>Normal mode"
RPW_LFR_NORMAL_SWF	LFR Snapshot Waveform data in normal mode	"RPW-LFR>RPW Low Frequency Receiver"	"SWF>Snapshot WaveForm"	"NORM>Normal mode"
RPW_LFR_NORMAL_CWF	LFR Continuous	"RPW-LFR>RPW"	"CWF>Continuous WaveForm"	"NORM>Normal mode"

Product name	Description	Descriptor	Data_product	Free_field
	Waveform data in normal mode	Low Frequency Receiver"		
RPW_LFR_BURST_CWF	LFR Continuous Waveform data in burst mode	"RPW-LFR>RPW Low Frequency Receiver"	"CWF>Continuous WaveForm"	"BURST>Burst mode"
RPW_LFR_SBM1_CWF	LFR Continuous Waveform data in SBM1 mode	"RPW-LFR>RPW Low Frequency Receiver"	"CWF>Continuous WaveForm"	"SBM1>Selected Burst Mode 1"
RPW_LFR_SBM2_CWF	LFR Continuous Waveform data in SBM2 mode	"RPW-LFR>RPW Low Frequency Receiver"	"CWF>Continuous WaveForm"	"SBM2>Selected Burst Mode 2"
RPW_LFR_NORMAL_ASM	LFR Averaged spectral matrix in normal mode	"RPW-LFR>RPW Low Frequency Receiver"	"ASM>Averaged Spectral Matrix"	"NORM>Normal mode"
RPW_LFR_NORMAL_BP1	LFR Basic Parameters 1 in normal mode	"RPW-LFR>RPW Low Frequency Receiver"	"BP1>Basic Parameters 1"	"NORM>Normal mode"
RPW_LFR_NORMAL_BP2	LFR Basic Parameters 2 in normal mode	"RPW-LFR>RPW Low Frequency Receiver"	"BP2>Basic Parameters 2"	"NORM>Normal mode"
RPW_LFR_BURST_BP1	LFR Basic Parameters 1 in burst mode	"RPW-LFR>RPW Low Frequency Receiver"	"BP1>Basic Parameters 1"	"BURST>Burst mode"

Product name	Description	Descriptor	Data_product	Free_field
RPW_LFR_BURST_BP2	LFR Basic Parameters 2 in burst mode	"RPW-LFR>RPW Low Frequency Receiver"	"BP2>Basic Parameters 2"	"BURST>Burst mode"

Low Latency

There is only one data product for the RPW low latency data.

Product name	Description	Data_product	Free_field	Comment
RPW_LL	RPW low latency data including radio flux around 1MHz, SBM event information and BIAS sweeping calibration data (only once a week).	"LL>Low Latency"		

4.2.4 SPICE (RS)

4.2.4.1 Note on the Template

Copy the complete template document tree and rename this page to "<instrument> (IS or RS)", e.g. "STIX (RS)". See [STIX \(RS\)](#) for an actual example.

4.2.4.2 Issues and Conflicts with Proposal

This section is used to point out possible conflicts between instrument-specific requirements and the current metadata proposal or to add information to the proposal that may affect other instruments (e.g. file naming, versioning, etc.). For "non-controversial" requirements or FITS/CDF metadata keywords, use other sections.

Data Formats and File Names

FITS/CDF Metadata

Other

4.2.4.3 FITS Metadata Standard

Section (e.g. Basic Keywords)

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	D-Type	Level
SIMPLE	T	file conforms to basic FITS standards	T (for true)	Boolean	SIMPLE is a logical keyword specifying whether the file conforms to basic FITS standards; Solar Orbiter data should always comply, so T (for True) is only valid option here.	<instrument>, e.g. STIX	<data type>	All

Extra (e.g. engineering, team, etc.)

4.2.4.4 CDF Metadata Standard

Section

NAM E	DESCRIP TION	TYPE	LEV EL	DATA TYPE	COMMENT
EPO CH	Primary time variable as defined in the section 1.1.2	<instrum ent>, e.g. RPW	All	CDF_TIME_T T2000	See http://spdf.gsfc.nasa.gov/istp_guide/variables.html#Epoch .

4.2.4.5 Data Products

Science

Low Latency

4.2.5 STIX (RS)

4.2.5.1 Issues and Conflicts with Proposal

Data Formats and File Names

- The FITS file names for STIX will need to include energy information (free text field). It may be beneficial to change the order of the file name keywords described in the current proposal [Data Formats and File Names](#) to have the free text switched with the version.

FITS/CDF Metadata

- Will need to think about lightcurves in FITS data (and the impact on the keywords)

Other

4.2.5.2 FITS Metadata Standard

Common

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
						STIX	All

Imaging

Keyword	Example Value	Suggested FITS Comment	Valid Data Range	Value Type	Description	M-Type	Level
ENERGY_L	6	Lower energy band in keV		Byte		STIX	All
ENERGY_H	12	Upper energy band in keV		Byte		STIX	All

Extra (e.g. engineering, team, etc.)

4.2.5.3 Data Products

Science

- Visibilities
- Images
- Light curves
- Spectra
- Spectrogram

Low Latency

- Light curves
- Spectra
- Aspect

4.3 VESPA keywords for Unified Solar Orbiter Catalog (USOC)

We present here the [VESPA \(Virtual European Solar and Planetary Access\)](#) keywords to be studied as keyword names for the Unified Solar Orbiter Catalog (USOC). We also propose to link them to existing FITS and CDF metadata, when they already exist.

The units and allowed values proposed in this series of keywords have been selected for a client-server (computer-to-computer) interaction. The actual user interface shall transform the units and values into human readable and usable units and values (e.g.: the spectral range unit for the query is Hz, but the user interface shall allow the user to chose its preferred unit and transform the user query into a protocol query automatically).

The VESPA keyword list proposed here is the released version available on Nov. 2014.

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	Sol O-FITS keyword	Notes
dataprodukt_type	• im	high level	DATAPRO		See Note [1]

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoI O-FIT S keyword	Notes
	<ul style="list-style-type: none"> • sp • ds • sc • pr • vo • mo • cu • ts • ca • sv 	scientific organization of the data	DUCT_TYPE		
target_name	Sun		TARGET_NAME		
target_class	<ul style="list-style-type: none"> • star • interplanetary_medium 		TARGET_CLASS		Remote data could have "star" and in-situ "interplanetary_medium", but this may not be adequate.
target_region	<ul style="list-style-type: none"> • Chromosphere • Corona • Interior • Photosphere • Solar Wind • ... 		TARGET_REGION	OBJECT	<p><i>The current definition of the FITS "OBJECT" keyword does not fit exactly with this definition. To be mixed with the "Element_name" attribute.</i></p> <p>This parameter optionally identifies the region of interest for the resource, in complement to target_name. This parameter only</p>

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoI O-FITS keyword	Notes
					<p>introduces generic regions, not specific local names, which must be handled using the <code>element_name</code> parameter (see examples above). The best practice is to take the values from standard sources:</p> <ul style="list-style-type: none"> • IVOA thesaurus: http://www.ivoa.net/rdf/Vocabularies/vocabularies-20091007/IVOAT/ • IAU thesaurus http://www.mso.anu.edu.au/library/thesaurus/ + another version: http://www.vocabularyserver.com/trex/en/ The latter seems more recent and more complete (although the interface is not practical) • Spase dictionary http://www.spase-group.org/
time_min	Julian day	The date and time of the beginning of the first acquisition for the data contained in the file	TIME_MIN	DATE-BEGIN	CDF "START_TIME" and FITS "DATE-BEGIN" are usually in ISO-8601 format, and thus are to be transformed into Julian Day at the time of the USOC table generation

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoIO-FITS keyword	Notes
time_max	Julian day	The date and time of the end of the last acquisition for the data contained in the file	TIME_MAX	DATE-END	
time_origin	<ul style="list-style-type: none"> target spacecraft 				This a service level, not data product level metadata
time_scale	UTC				This a service level, not data product level metadata
time_sampling_step_min	seconds	Minimum value of the sampling temporal step within the file.			
time_sampling_step_max	seconds	Maximum value of the sampling temporal step within the file.			
time_exp_min	seconds	Minimum value of the exposure time (=integration time) within the file			
time_exp_max	seconds	Maximum value of the			

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoIO-FITS keyword	Notes
		exposure time (=integration time) within the file			
spectral_range_min	Hz	Minimum value of the spectral range within the file			Standard transformation rules from energy, wavelength or wavenumber to frequency are provided in the documentation
spectral_range_max	Hz	Maximum value of the spectral range within the file			
spectral_sampling_step_min	Hz	Minimum value of the spectral sampling within the file			This is the spectral separation between the centers of two adjacent filters, channels or pixels.
spectral_sampling_step_max	Hz	Maximum value of the spectral sampling within the file			
spectral_resolution_min	Hz	Minimum value of the spectral resolution within the file			This is the spectral bandwidth used for the measurement (Full Width at Half Maximum). In case of a filter camera this is the filter bandwidth; in case of a spectrometer this is the spectral resolution per se.
spectral_resolution_max	Hz	Maximum value of the spectral resolution within the file			

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoIO-FITS keyword	Notes
spatial_frame_type	<ul style="list-style-type: none"> celestial body cartesian spherical healpix 	The spatial coordinate flavour used for defining C1, C2 and C3 axes.			See Note [2]
spatial_coordinate_description	String	Name and/or description of the coordinate system (e.g.: HEEC, Carrington, RTN...)			A list of abbreviations for coordinate systems is being worked on within VESPA project, suggestions very welcome.
spatial_origin	String	Spacecraft or target			
c1min	depends on spatial_frame_type:	Minimum value on axis C1 within the file			
c2min	<ul style="list-style-type: none"> m for distances deg for angles 	Maximum value on axis C1 within the file			
c3min		Minimum value on axis C2 within the file			
c1max		Maximum value on axis C2 within the file			

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoI O-FIT S keyword	Notes
c2max		Minimum value on axis C3 within the file			
c3max		Maximum value on axis C3 within the file			
c1_resol_min		Minimum value of the spatial resolution on axis C1 within the file			
c2_resol_min		Minimum value of the spatial resolution on axis C2 within the file			
c3_resol_min		Minimum value of the spatial resolution on axis C3 within the file			
c1_resol_max		Maximum value of the spatial resolution on axis C1 within the file			
c2_resol_max		Maximum value of the spatial			

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoIO-FITS keyword	Notes
		resolution on axis C2 within the file			
c3_resol_max		Maximum value of the spatial resolution on axis C3 within the file			
instrument_host_name	Solar Orbiter	Name of the facility/Observatory/Mission hosting the instrument			Standard systems outside the Solar Orbiter community will look for a internal metadata to find such information.
instrument_name	<ul style="list-style-type: none"> • MAG • RPW • EPD • SWA • PHI • EUI • SPICE • STIX • METIS • SoloHI 	Name of the instrument			This information is in the file name, but standard systems outside the Solar Orbiter community won't decode the filename, but rather look for a internal metadata to find such information.
measurement_type	UCDs	Description of the observed parameter, expressed using UCDs			See Appendix section on UCDs .

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoIO-FITS keyword	Notes
access_url	URL	Link to the file (at ESAC)			
access_format	CDF or FITS	Format of the file			
access_estsize	kBytes	Size of the file			
preview_url	URL	Link to a preview of the file (format=png)			
element_name	String			OB JEC T	Name of the element observed on the Target (e.g., active region number)
reference	bibcode or DOI or URL		REFERENC E		
particle_spectral_type	<ul style="list-style-type: none"> energy mass mass-per-charge 	Type of particle spectral measurements			This defines the units of the particle_spectral_* attributes
particle_spectral_range_min	<ul style="list-style-type: none"> eV for energy amu for mass amu/q_e for mass-per-charge 	Minimum value of particle spectral range in the file			
particle_spectral_range_max		Maximum value of particle spectral			

VESPA keyword	VESPA keyword unit or list of allowed values	VESPA keyword description	SoIO-CDF keyword	SoI O-FIT S keyword	Notes
		range in the file			
particle_spectral_sampling_step_min		Minimum value of particle spectral sampling step in the file			
particle_spectral_sampling_step_max		Maximum value of particle spectral sampling step in the file			
particle_spectral_resolution_min		Minimum value of particle spectral resolution in the file			
particle_spectral_resolution_max		Maximum value of particle spectral resolution in the file			
processing_level	0 to 5	Data processing level of the data in the file			Using CODMAC processing levels.

Notes:

[1] Detailed description of "data product_type" values:

- **im** = image: scalar field with two spatial axes, or association of several such fields, e.g., images with multiple color planes, from multichannel or filter cameras. Maps of planetary surfaces are considered as images.
- **sp** = spectrum: measurements organized primarily along a spectral axis, e.g., a series of radiance spectra.
- **ds** = dynamic_spectrum: consecutive spectral measurements through time, organized as a time series.
- **sc** = spectral_cube: sets of spectral measurements with 1 or 2D spatial coverage, e.g., imaging spectroscopy. The choice between image and spectral_cube is dictated by the characteristics of the instrument (which dimension is most resolved)
- **pr** = profile: scalar or vectorial measurements along 1 spatial dimension, e.g., atmospheric profiles, atmospheric paths, sub-surface profiles...
- **vo** = volume: other measurements with 3 spatial dimensions, e.g., internal or atmospheric structures.
- **mo** = movie: sets of chronological 2D spatial measurements
- **cu** = cube: multidimensional data with 3 or more axes, e.g., all that is not described by other 3D data types such as spectral cube or volume. This is mostly intended to accommodate unusual data with multiple dimensions.
- **ts** = time_series: measurements organized primarily as a function of time (with exception of dynamical spectra and movies, i.e. usually a scalar quantity). Space-borne dust detector measurements are a typical example of a time series.
- **ca** = catalog: can be a list of events, a catalog of object parameters, a list of features.... e.g., a list of asteroid properties. It can be limited to scalar quantities, and possibly limited to a single element (i.e., one catalog entry). Time_series, Profile, and Catalog are essentially tables of scalar values. In Time_series the primary key is time; in Profile it is altitude or distance; in Catalog, it may be a qualitative parameter (name, ID...).
- **sv** = spatial_vector: list of summit coordinates defining a vector, e.g., vector information from a GIS, spatial footprints...

[2] Detailed description of "spatial_frame_type" value:

- **celestial**: 2D angles on the sky, e. g., right ascension c1 and declination c2 + possibly distance from origin c3 – although this is a special case of spherical frame, the order is different.
- **body**: 2D angles on a rotating body: longitude c1 and latitude c2 + possibly a z c3 coordinate. The best practice is to follow the IAU 2009 planetocentric convention [RD12], in particular eastward longitudes and a north pole located on the north side of the invariant plane of the Solar System for planets and satellites (see [RD12] for small bodies, and Annex A for details). The Z coordinate is by default the distance counted from the center of mass. The spatial_coordinate_description and spatial_origin attributes allow the data provider to indicate different conventions, e. g., to indicate a planetographic frame, or to use altitude above a reference surface as the third coordinate. It is stressed however that using other frames will make comparisons between datasets more difficult.

- **cartesian:** (x,y,z) as (c1,c2,c3). This includes spatial coordinates given in pixels. cylindrical: (r, theta, z) as (c1,c2,c3); angles are defined in degrees.
- **spherical:** (r, theta, phi) as (c1,c2,c3); angles are defined in degrees as in usual spherical systems (E longitude, zenithal angle/colatitude). If the data are related to the sky, “celestial” coordinates with RA/Dec must be used.
- **healpix:** (H, K) as (c1,c2).