



PHEBUS Data Processing, Quick-Look Analysis & Archiving Current Status and Plans

Aurélie Reberac



Data Processing & Archiving

Current status and plans

- **Data products/archive : levels, format, content, filenaming & organisation**
- **Data processing and Quick-Look Analysis implementation**
- **Data & quick-look products access, availability & use**

PHEBUS data products & archive Levels, format, content filenaming & organisation

PHEBUS archive content and structure



PHEBUS archive = collection of all operational and science data produced by the PHEBUS instrument onboard Bepi during all the mission phases

- **Identifier (LID) = `urn:esa:psa:bc_mpo_phebus`**
- **Products levels** = raw, partially processed, calibrated, derived
- **data_* collections** = primary data products (i.e Science & HK)
- **Miscellaneous collection** = other “extra” PHEBUS data (TM events...) + Phebus operation catalog
- **browse, geometry** = supplementary products associated to science products
- **calibration_* collections** = calibration products and derived calibration files
- **All collections split by mission phases, and inside a mission phase, by year and month (YYYYMM)**
- **All products are compliant with the PDS4* archiving standard**

```
bc_mpo_phebus/
|--data_raw/
|--data_partially_processed/
|--data_calibrated/
|--data_derived/
|--miscellaneous/
    |--extra_tm/
    |--pheb_misc_operation_log.csv
|--calibration_files/
|--calibration_raw/
|--calibration_partially_processed/
|--browse_partially_processed/
|--browse_calibrated/
|--browse_derived/
|--geometry/
```

```
bc_mpo_phebus/
|--data_raw/
    |--cruise
        |--202203
        |--202204
        ...
    |--mercury
        |--202603
        |--202604
        ...
```

**Planetary Data System, v4 data standards (NASA)*

PHEBUS raw data products

Content & format -- HK

- Identifier (LID) = **urn:esa:psa:bc_mpo_phebus:data_raw**
- Products = Raw (uncalibrated) instrument housekeeping and science data
 - TM sorted by data types, de-packetized, time-ordered, science images decompressed/reconstructed
 - Raw products saved into a FITS/PDS4 compatible format, e.g. binary tables, 2D or 3D arrays
 - Engineering data in DN (digital numbers), measurements in counts/s.
- **HK FITS file content = All HK collected during one switch ON-OFF of the instrument**
 - One 2D binary table extension per HK type (Temperatures, Voltages, DPU, Scanner, Detector...up to 8)
 - Each column = one parameter, each row : one record of all parameters at a given time

Index	Extension	Type	Dimension
0	Primary	Image	0
1	PHEB_Temperature_Housekeeping	Binary	27 cols X 181 rows
2	PHEB_Voltage_Housekeeping	Binary	18 cols X 181 rows
3	PHEB_Detector_Housekeeping	Binary	51 cols X 301 rows
4	PHEB_DPU_Housekeeping	Binary	101 cols X 181 rows
5	PHEB_Scanner_Monitor_Housekeeping	Binary	30 cols X 160 rows
6	PHEB_Windows_Parameter_Housekeeping	Binary	19 cols X 2 rows
7	PHEB_TM_frequencies_Housekeeping	Binary	31 cols X 6 rows

Example of FITS extensions in a HK FITS file

FITS : Flexible Image Transport System

Extract of the Temperature HK table

EUV_Temp_Elec	FUV_Temp_MCF	FUV_Temp_Elec	IF_Temperature	Starsys_Temperature	OBT_HK_Temp
J	J	J	J	J	23A
RAW	RAW	RAW	RAW	RAW	OBT
Modify	Modify	Modify	Modify	Modify	Modify
3358	3272	3270	3300	3354	2022-04-16T07:12:56.450
3357	3272	3270	3300	3354	2022-04-16T07:13:01.450
3358	3272	3270	3300	3354	2022-04-16T07:13:08.449

Note: visualisation of FITS file content with the NASA FITS Viewer (FV) tool <https://heasarc.gsfc.nasa.gov/ftools/fv/>

PHEBUS raw data products

Content & format -- SC



- **SC** = each science data acquisition = a "science header" table + a EUV/FUV data block
 - **Science header** = information about the detector configuration, scanner commanding and real position, ... + **NUV K & Ca counts** (null if not activated)
 - **Data block** = depends on the EUV/FUV detector operating mode, (*no data block if only NUV !*)
 - "pixel" mode ("IMG"): an image, number of counts/s for each pixel of the detector image or ROI (Region of Interest)
 - "histogram" mode ("PHA"): Pulse height distribution curve for each of the 4 corners of the detector resistive anode encoder (RAE) and for the sum of the 4 corners of the RAE
- **SC FITS file content = all acquisitions (header + data block) collected during one observation**

✓ a 2D binary table extension containing all science header (similar to HK tables)

+ if EUV or FUV detector activated in "pixel" (IMG) operating mode:

✓ a 3D array mapping the number of photons counts registered per second onto each of the pixel of the detector ROI (with axis1, axis2 = spectral and spatial dimensions of the ROI) as a function of time (axis3).

or if EUV or FUV detector activated in "histogram" (PHA) operating mode:

✓ a 3D array containing the Pulse Height Distribution spectrum (axis1) for each of the 4 RAE corners (axis2) as a function of time (axis3)

✓ a 2D array containing the Pulse Height Distribution spectrum (axis1) for the sum of the 4 RAE corners as a function of time (axis2)

SC FITS file extensions when only NUV

Index	Extension	Dimension
0	Primary	0
1	PHEB_Science_Header	53 cols X 2410 rows

SC FITS file extensions when XUV in IMG mode

Index	Extension	Dimension
0	Primary	0
1	PHEB_Science_Header	45 cols X 181 rows
2	IMG_3D_PIXEL	256 X 512 X 181

SC FITS file extensions when XUV in PHA mode

Index	Extension	Dimension
0	Primary	0
1	PHEB_Science_Header	45 cols X 6 rows
2	PHA_3D_HIST_corners	4 X 16384 X 6
3	PHA_2D_HIST_sum	65536 X 6

PHEBUS calibrated data products or only partially processed

- Identifier (LID) = **urn:esa:psa:bc_mpo_phebus:data_calibrated**
- Products = calibrated instrument housekeeping and calibrated science data
- Main properties =
 - **HK** : all raw data converted to textual or physical units using PHEBUS on-ground database calibration curves
 - **SC** : Science header data : converted to textual or physical units using PHEBUS database calibration curves
Science image data : dark subtraction, pixel to wavelength and radiometric calibration using calibration products (dark model, pixel to wavelength assignment image, effective area) stored in the calibration_files collection.
- FITS files content/format = identical to HK and SC raw data products (refer to previous slides)

Note : no calibrated science data will be produced during NECP, cruise or flybys

- Identifier (LID) = **urn:esa:psa:bc_mpo_phebus:data_partially_processed**
- Products/main properties = science data products partially processed, i.e only the science header part is converted to textual or physical units (no corrections or calibration applied on the science image data)
- FITS files content/format = identical to SC raw data products (refer to previous slide)

Partially processed science data will be produced for all mission phases

Data products file naming convention

- **HK data products** (= 1 file per switch ON/OFF of the instrument)

phe_<level>_HK_YYYYMMDDTHHMMSS_YYYYMMDDTHHMMSS.fits/xml

with <level> defines the **data level**, i.e “raw” (uncalibrated), “cal” (calibrated)

YYYYMMDDTHHMMSS : **start/end time** in UTC, that should correspond to the switch ON/OFF times of the instrument, rounded to the minute (i.e ‘SS’ = always ‘00’)

- **Science data products** (= 1 file per observation or change in the detector configuration)

phe_<level>_sc_<subunit>_<obs_mode>_YYYYMMDDTHHMMSS_YYYYMMDDTHHMMSS.fits/xml

with <level> defines the **data level**, i.e “raw” (uncalibrated), “par” (partially processed), “cal” (calibrated)

<subunit> defines the **detector(s) used**. It is either <euva> (i.e. EUV alone), <fuva> (i.e. FUV alone), <nuva> (i.e. NUVs alone), <euvn> (i.e. EUV + NUVs), or <fuvn> (i.e. FUV + NUVs).

< obs_mode > defines the **observation operating mode** (4 letters, refer to the table in next slides)

YYYYMMDDTHHMMSS : **start/end time** in UTC, that should correspond to the start/end times of an observation, rounded to the minute (i.e ‘SS’ = always ‘00’)

Note : one may have 1 HK product and several science products during one PHEBUS switch ON/OFF

Calibration operating modes

Identifier in TM	Description	Textual Value in PHEBUS database	obs_mode	Collections	Detector mode
1	EUV PHA Calibration	C_EUV_GAIN	EPHA	calibration_raw calibration_partially_ processed	PHA
2	FUV PHA Calibration	C_FUV_GAIN	FPHA		PHA
3	NUV K Calibration PM Curve Plate	C_NUV_K_PLAT	KCPL		N/A (EUV/FUV OFF)
4	NUV Ca Calibration PM Curve Plate	C_NUV_CA_PLAT	CCPL		N/A (EUV/FUV OFF)
6	Dark Current Parking Bracket	C_DARK	DRKH if PHA DRKI if IMG		PHA or IMG
7	EUV Analog Noise 1 point	C_EUV_1PULSE	E01H if PHA E01I if IMG		PHA or IMG
8	EUV Analog Noise 16 point	C_EUV_16PULSE	E16H if PHA E16I if IMG		PHA or IMG
9	FUV Analog Noise 1 point	C_FUV_1PULSE	F01H if PHA F01I if IMG		PHA or IMG
10	FUV Analog Noise 16 point	C_FUV_16PULSE	F16H if PHA F16I if IMG		PHA or IMG

EUV/FUV calibration modes : 1, 2 : to check the detector HV setting and ageing; **6**: to measure the dark current;
7 to 10 : to check the spatial distribution of the counts on the detector

Science operating modes

Identifier in TM	Description	Textual Value in PHEBUS database	obs_mode	Collections	Detector mode
11	Fixed LOS Day Side	EXODAY_DRIFT	XODD	data_raw data_partially_processed data_calibrated	IMG
12	Fixed LOS Night Side	EXONIGHT_DRIFT	XOND		
13	Fixed LOS Nadir	NADIR_DRIFT	SNDD		
14	Feature Tracking on Surface	SPOT_TRACK	STRK		
15	Feature Tracking in Exosphere	EXO_TRACK	XTRK		
16	Variable LOS Day Side	EXODAY_SCAN	XDSC		
17	Variable LOS Night Side	EXONIGHT_SCAN	XNSC		
18	Stellar / Inertial Pointing	STAR_INERT	STIN		
19	Stellar / Normal Pointing	STAR_DRIFT	STDF		
20	Stellar Normal Pointing with tracking	STAR_TRACK	STTK		
21	IP Background / Inertial Pointing	SKY_INERT	IPIN		
22	IP Background / Normal Pointing	SKY_DRIFT	IPDF		
23	IP Background scan	SKY_SCAN	IPSC		
24	Solar Corona	CORONA_TRACK	CRNA		
25	Comet / Inertial Pointing	COMET_INERT	CMTI		
26	Comet / Normal Pointing with scan	COMET_SCAN	CMTS		
27	Roll Mode	SKY_ROLL_DRIFT	FLIP		

List of operating modes dedicated to science, defined according to the pointing and/or the target

PHEBUS Data processing & Quick-Look Analysis tools Overview & Status

Data processing & Quick-Look Analysis Purpose and Responsibilities



- **Data processing software ("pipeline")**

To convert the instrument telemetry into scientific data products (raw through to calibrated level + meta-data, geometry)

- **Quick-look analysis software ("QLA")**

To generate **overview plots** of the instrument HK and science data (to assess the **instrument health and performance** and to give a **first view of the scientific measurements**)

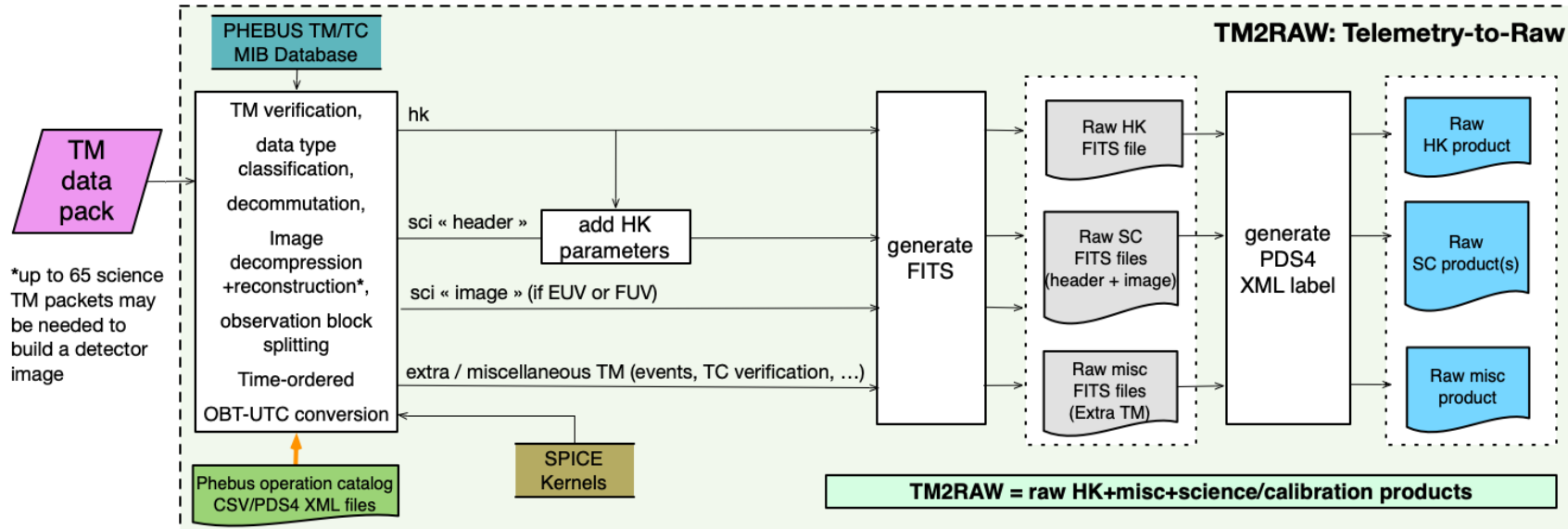
Software	Responsibilities			
	Specification	Development	Operations	Maintenance
pipeline	PHEBUS team	PHEBUS team*	SGS team	PHEBUS team*
QLA	PHEBUS team	SGS team (web-interface , dynamic plots) ** PHEBUS team (operation HTML reports)	SGS team PHEBUS team	SGS team PHEBUS team

* pipeline fully developed and maintained by the PHEBUS team and delivered to SGS for integration into the SGS system. The SGS is responsible for integrating and testing the pipeline, and for systematic generation and archiving of PHEBUS data products.

** co-development approach – 2 tools

SGS : *Bepi Science Ground Segment @ESAC*

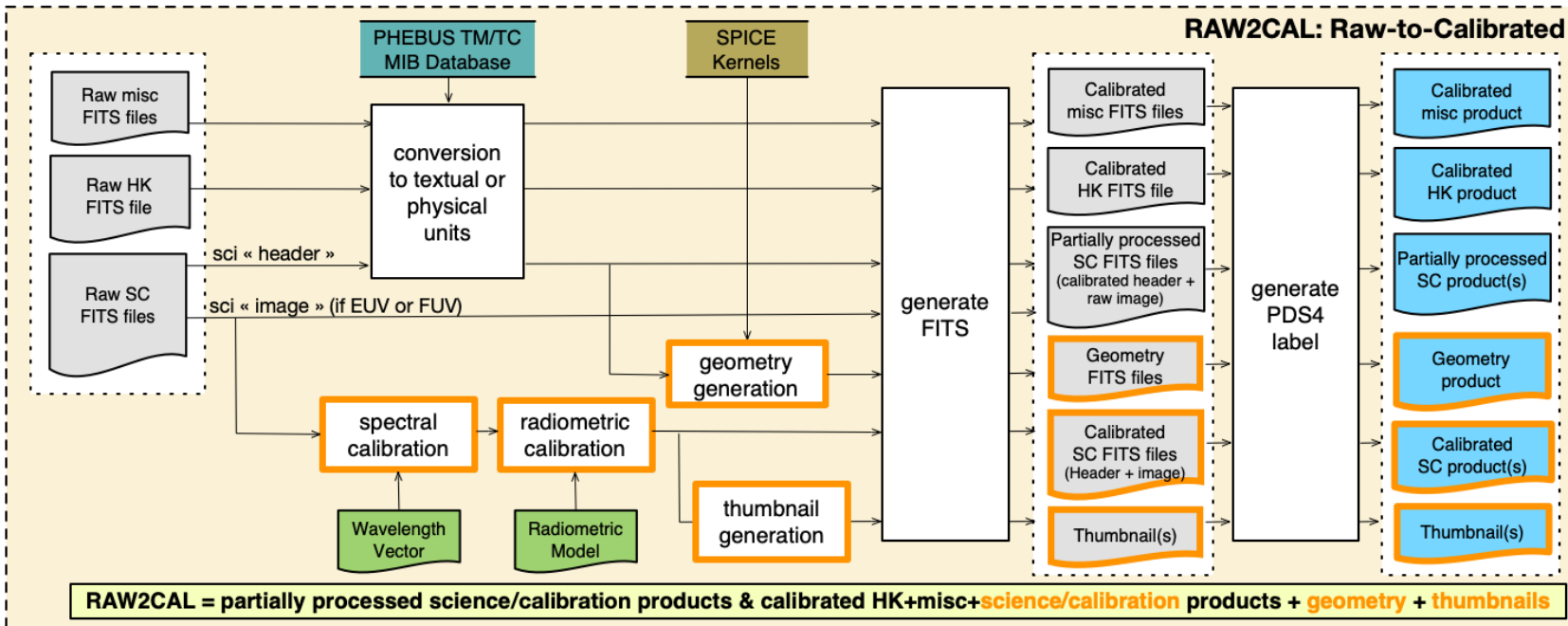
Data processing software ("pipeline")



Product = FITS + PDS4 XML label

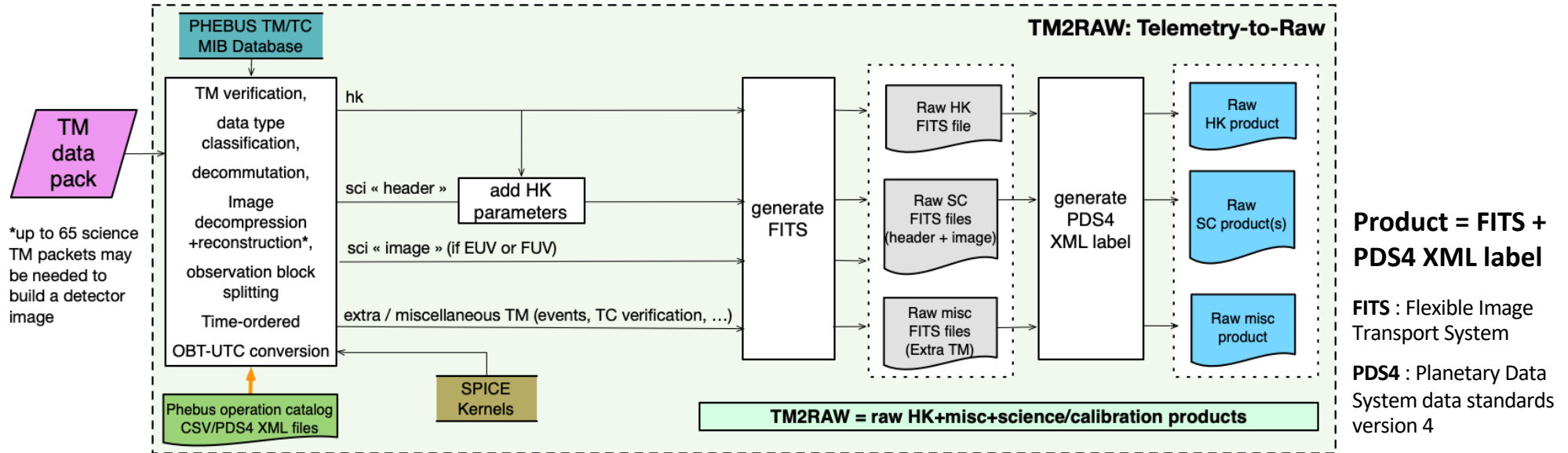
FITS : Flexible Image Transport System

PDS4 : Planetary Data System data standards version 4

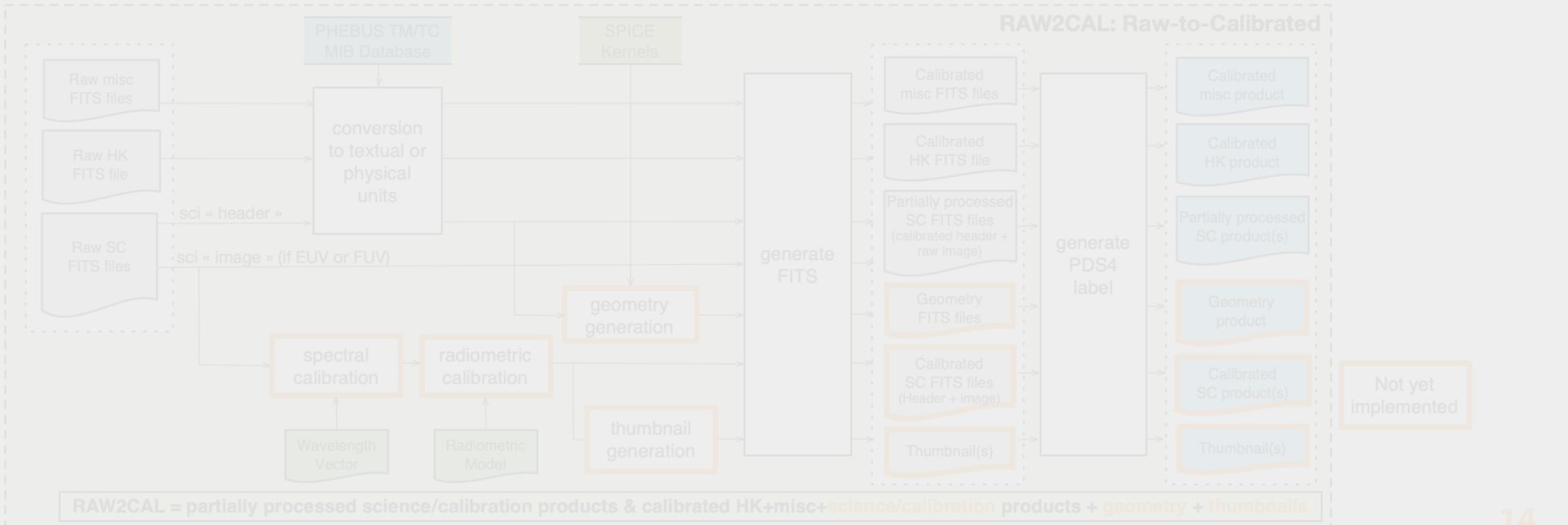


Not yet implemented

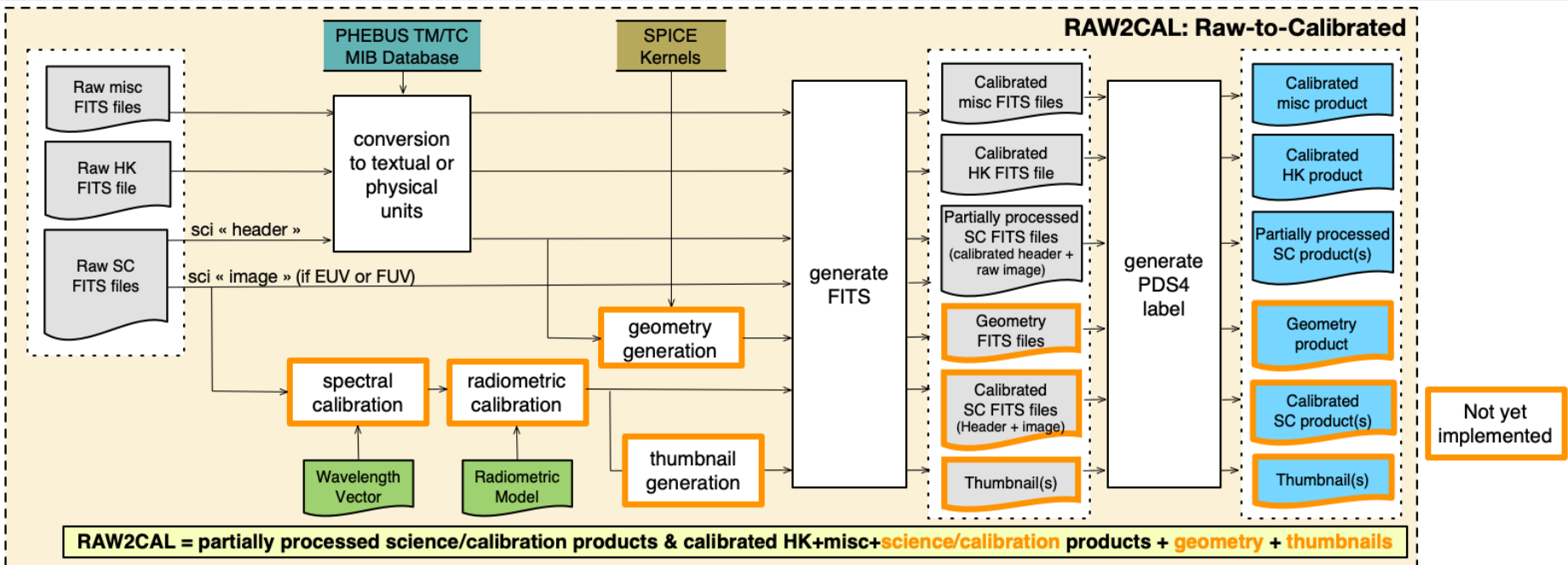
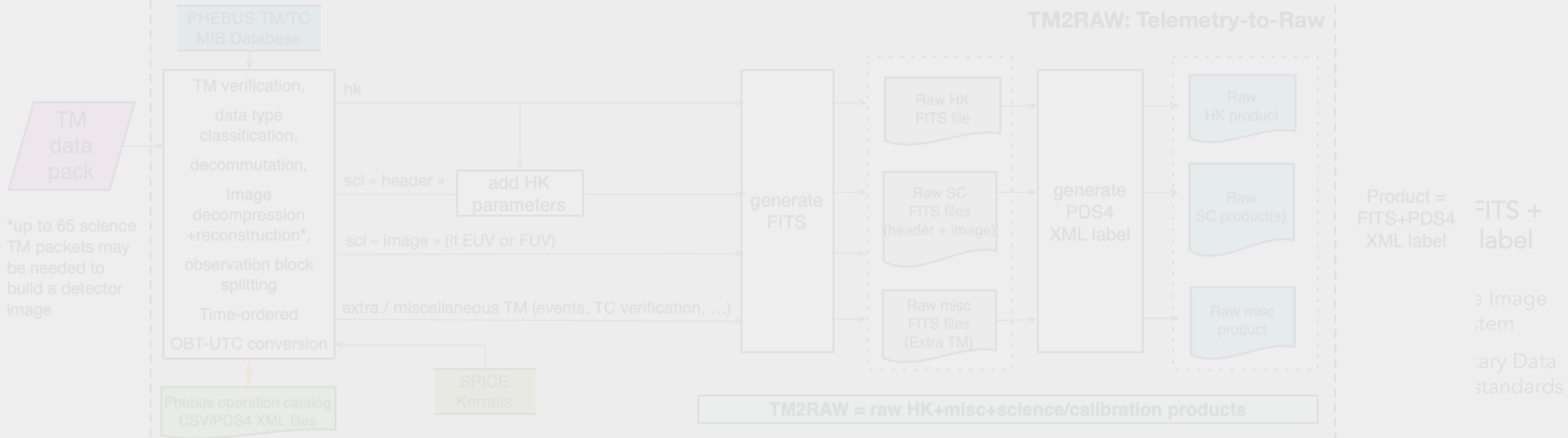
Data processing software ("pipeline")



*up to 65 science TM packets may be needed to build a detector image



Data processing software ("pipeline")



- **Last release 0.3.4** delivered to SGS on **11.6.2022**

- **Produces**

- raw and calibrated HK and misc products (FITS+PDS4 XML label)
- raw and partially processed science products (FITS+PDS4 XML label)

Last release will be deployed by SGS in time for the Mercury flyby#2

- **Updates required before a first official archive review and public data release**

- Time shift found in science products between header parameters and images
- Targets and pointing information to be filled in (requires the reading of the Phebus Operation Catalog)
- Description of fields should be improved
- Generic Array_2D, Array_3D objects could be replaced with specific array (e.g. Array_2D_Image, Array_2D_Spectrum, Array_3D_Image, ...) such that software immediately knows how to treat the data.
- ...

- **Does not (yet) produce :**

- browse, geometry and calibrated science products (other home-made tools used)
- index/catalog of science products

- **Next release 0.3.5** including data products updates → **end of sept. 2022**

- **Next release 0.4.0** including geometry and index products → **end of dec. 2022**

- **PHEBUS QLA** (=HTML reports including tables and plots for HK, Science, Misc data)

[HK data report](#)

[Science image data report](#) (FUV dark + stellar observation on March 28, 2022)

[Science header data report](#)

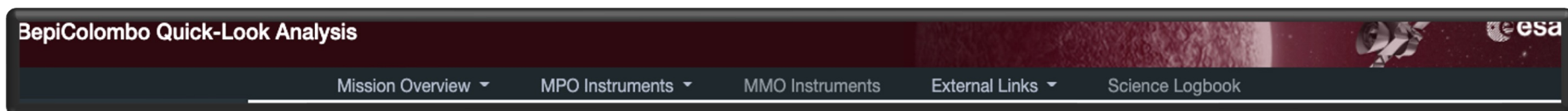
[Miscellaneous data report](#)

Partially processed PHEBUS SCIENCE images Overview



- **BEPI QLA** (=web-interface displaying operational data (ground stations pass information, power consumption, data volumes, telemetry events, ...), and instrument-specific housekeeping and science data views. Data plots are interactive widgets, grouped into thematic dashboards.

<http://bepicolombo.esac.esa.int/bscs-qla/home>



Data & quick-look products

Access, availability & use

Data & quick-look products access, Availability & use

- **Public Releases of (raw & partially processed) data collected during Cruise:**
 - Based on the achievement of the next pipeline releases (i.e september & december 2022), and if products are considered to be « scientifically good enough », the **first archive review should take place in 2023**
 - A specific date for the **first public release** of data collected up to mid-2022 will be agreed between the Bepi Project Scientist and the PI based on the results of the first archive review, and after closure of the review actions (objective : **end of 2023**).
 - Next public releases of data during Cruise will follow the ESA deadlines
- **Public data access :**

ESA PSA (**Planetary Science Archive**) for Bepi Colombo mission :

<https://archives.esac.esa.int/psa/#!Table%20View/BepiColombo=mission>
- **Bepi-Colombo QLA access : for instruments teams only**
 - Request for access/credentials via the PI
- **Other PHEBUS team tools (plans for the end of next year)**
 - Tutorial "PHEBUS data analysis" in the form of python scripts/jupyter notebooks (set of examples to read, manipulate and visualize the calibrated HK and partially processed science products)

Data & quick-look products access, Availability & use

- **Plan for the release of the calibration products**
 - e.g dark current model , calibration factors (flat field, effective area), wavelength assignment
 - **First official version in 2024**
- **Plan for generation and release of calibrated science data products**
 - **Implementation in the pipeline foreseen in 2025**, to be ready for Bepi in orbit phase
 - Will be generated once reconstructed spacecraft attitude is available (i.e ~2 weeks after raw products)
 - **Public releases of these data collected during nominal mission will follow ESA deadlines**
- **Plan for generation of derived products**
 - **Preliminary list of PHEBUS products that will be stored in the PSA:**
 - Calibrated intensities as a function of wavelength with geometry
 - Column densities of relevant species as a function of geometry (including altitude profiles)
 - Scale heights of various species derived from intensity profiles when applicable
 - Interplanetary Background Brightness for H (121.6 nm) and He (58.4 nm)
 - Reference spectra for various bright stars in the UV range
 - Surface albedo at 121.6 nm. Map resolution TBD most likely limited to the polar regions
 - **Public releases of these data collected during nominal mission will follow ESA deadlines**
 - **Release of CRUISE data : TBD**