What is in this Herschel Science Archive tar file

1. Description of the directory structure
2. File naming convention
3. Structure of the FITS files

1 Description of the directory structure

Herschel observations contain many data products. The data retrieved from the Herschel Science Archive (HSA) as a tar file will unpack into a directory with the request identifier (an arbitrary number) as the name, inside which the data are organised in a tree-like structure of directories:

- The **auxiliary** directory contains all Herschel non-science spacecraft data required directly or indirectly for the processing and analysis of the scientific data.
- The **calibration** directory contains the uplink and downlink calibration products.
- A directory with the observation identifier (<**obsid**>) as the name, which contains the **science data** distributed in sub-directories called level0/0.5/1/2/2.5/3. These levels hold data at different stages of processing: Level 0 are raw, Level 0.5 and 1 are partially processed, Level 2 are of science quality, Level 2.5 are combinations of Level 2 products, and Level 3 are super-combinations of observations (like mosaics or stiched spectra; see the Data Product Overview web page for more details).

A generic unpacked HSA tar file will look like:

```
>request_Id/

      auxiliary/
            h<obsid>auxcontext_<timestamp>.fits.gz
            AcmsTelemetryProduct/     OrbitEphemeris/
            EventsLogProduct/       OrbitEventsProduct /    Pointing/
            Siam /            HorizonsProduct/        SremCalProduct/
            Housekeeping/     SremRawProduct/   MissingTm/
            TeleCommandHistory/           MissionTimeLine/
            TimeCorr/         OOL/        UplinkProduct /

      calibration/
            h<instrument>calibrationcontext_<timestamp>.fits.gz
            Downlink(for HIFI)/     Uplink(for HIFI)/
            photometer(for PACS)/   spectrometer(for PACS)/
            Phot(for SPIRE)/  Spec(forSPIRE)/

      <obs_id>/
            h<instrument><obsid>obs_<timestamp>.fits.gz
            browseImageProduct/     browseProduct/    logObsContext/
            quality/    qualitySummary/
            level0/     level0_5/   level1/     level2/     level2_5/   level3/

      <obs_id>-herschel.ia.obs.ObservationContext-XXXXXX.xml
```

For administrative purposes, an XML file (that you can ignore in most cases) is also included. This is used only when loading the observation into the Herschel Interactive Processing Environment (HIPE).

Many of the FITS files provided do not contain science data but rather references to other products.

These are the so-called *context* files.

The top context file (`h<instrument><obsid>obs_<timestamp>.fits.gz)` inmediately under the `<obs_id>` directory contains the main keywords related to the observation and the instrument setup in the first extension header.

In most cases, you will be interested in the FITS files that contain the science data, which are located in the deepest levels of the `<obs_id>` directory. Hence, you will find the final science products resulting from the pipeline processing in the sub-directories under the level2 and level2_5 directories.

The general structure can be described as:

```
> directory/
      <context_file>.fits.gz
      sub_directory/
            <context_file>.fits.gz
            sub_subdirectory1/
                        <science_data>.fits.gz
            sub_subdirectory2/
                        <science_data>.fits.gz
```

A particular subset of the final science products generated by the pipeline can also be retrieved directly from the HSA User Interface as "`Standalone Browse Products`", either from the retrieval options or by right-clicking in the observation postcard (see the Data Product Overview web page for details on the type of products served).

Further details on products and on the data processing can be found in:

- **Data Analysis Guide** (PDF and HTML)
- **PACS Data Reduction Guide: Spectroscopy** (PDF and HTML)
- **PACS Data Reduction Guide: Photometry** (PDF and HTML)
- **PACS Products Explained** (PDF and HTML)
- **SPIRE Data Reduction Guide** (PDF and HTML)
- **HIFI Data Reduction Guide** (PDF and HTML)
- **Product Definition Document** (PDF and HTML)

2 File Naming Convention

The Herschel product naming convention for exported FITS files takes a format which depends on the type of product. The generic format is as follows, although not all of the items are present in all filenames.

`h<product/instrument><subinst><obsid/od>_<bbid>_<level><type>_<slice>_<timestamp>`

where

- `h` stands for Herschel
- `<product/instrument>`: is the product type such as *aux* for auxiliary products or the instrument name either *hifi*, *pacs or spire* (note that all letters in the filename are lowercase).
- `<subinst>`: This is only relevant for instrument data. Depending on the instrument and on the

type of product, it stands for the subinstrument used, the detector, polarisation mode etc...

- `<obsid/od>`: The observation identifier given in decimal format or the Herschel operational day
- `<bbid>`: Some observational products are split into logical units such as building blocks. When this occurs the bbid is given in hexadecimal format.
- `<level>`: The level of the product is presented here. Level 0 products are represented by 00, Level 0.5 by 05, Level 1 by 10, Level 2 by 20, Level 2.5 by 25 and Level 3 by 30.
- `<type>`: This indicates the type of product as given in the meta keyword TYPE.
- `<slice>`: When data from an observation needs to be split further than by building block or in a way unrelated to building block, the number of the "slice" is given here. If 100 or fewer products result from the split then two digits (yy) represent the slices (in time order). For more than 100 slices, three digits (yyy) are used.

These items are organised in a way intended to produce a logical ordering of the filenames when listed in a directory.

All product names also contain at the end a 13 digit number which is a `<timestamp>` that the system generates when the FITS product is created.

The specific format per product is given in Section 2.10 of the Product Definition Document.

3 Structure of the FITS files

All Herschel FITS files are compatible with the majority of standard FITS readers. They are composed of several extensions:

- The first extension (0) contains always, only a header with information about the observation and the specific product.
- The following extensions (one or more) contain the data (images, spectra, coverage, uncertainties...). For instance, a FITS file from a `SimpleImage` has at least three image extensions called `image`, `error` and `coverage`. A `SpectralSimpleCube` has two three-dimensional datasets, `image` and `coverage`, and one table dataset `ImageIndex`, with two columns relating each cube layer to its wavelength. `SpectralSimpleCube` objects in Level 2.5 HIFI products are made of three extensions, called `image`, `weight` and `flag`.
- The last four (*History* extensions) are common to all Herschel products and contain the history of the data processing (scripts, tasks and parameters used).

A typical Herschel `SpectralSimpleCube` FITS file looks like:

```
XTENSION  EXTNAME         EXTVER EXTLEVEL BITPIX GCOUNT  PCOUNT NAXIS  NAXIS
0                                           32      0      0     0
1 IMAGE    image                           -64      1      0     3  39 x 39 x 29
2 IMAGE    coverage                        -64      1      0     3  39 x 39 x 29
3 BINTABLE ImageIndex                        8      1      0     2  12 x 29
4 IMAGE    History                          32      1      0     0
5 BINTABLE HistoryScript                     8      1      0     2  80 x 7
6 BINTABLE HistoryTasks                      8      1      0     2  35 x 1
7 BINTABLE HistoryParameters                 8      1      0     2  103 x 12
```

More detailed information on the structure of Herschel FITS files is provided in Section 1.15.4 of the Data Analysis Guide.