File Types, Formats, and Content

Accepted File and Content Types

The types and formats of products that are accepted for HLSP collections are given in the following tables. Note that science products in FITS format should conform to the latest version of the Standard, and may have multiple extensions (MEF).

Science Products

Product	Formats	Data Organization	Notes
Image	FITS	Simple, or MEF with one or more IMAGE extensions	One MEF file should include science and concomitant pixel data, in separate extensions (variance, data quality, exposure map, etc.), where appropriate.
Spectra: 2-D or higher	FITS	Simple, or MEF with one or more IMAGE extensions	One MEF file should include science and concomitant pixel data, in separate extensions (variance, data quality, etc.), where appropriate. One-dimensional specta contained in images must express the WCS using standard header keywords.
Spectra: 1-D		Simple, or MEF with one or more IMAGE or BINTABLE extensions	
Catalog	FITS, csv, DB dump	BINTABLE (FITS), ASCII or DB dump (csv or SQLite)	CSV files should be parsable with common software. Database dumps from e.g. MySQL or PostgresSQL may be acceptable; SQLite files are certainly acceptable.
Light curve	FITS	BINTABLE	Time coordinate information may be represented in keywords or as an explicit table column.
Model or Simulation			Consult with MAST staff. Simulations of observational data would ordinarily be represented in the format appropriate for the product being simulated.

Ancillary Products

Product	Formats	Data Organization	Notes
README	ASCII, HTML, markdown, sphinx	Single flat file, or file with markup; see Example README file, or a simpler ASCII README file	Document that describes the collection contents. It must provide a description of the semantic content and organization of data in the files, and bibliographic information for the published paper(s) that describes the creation and use of the collection. Put details of the science goals, data processing methodology, etc. in the Project Summary file.
Manifest	json, yaml, csv	Structured; see example file in yaml format: MANIFEST.yml	Document that describes the product delivery manifest. It must provide the organization of data among the files. Details of the science goals, data processing methodology, etc. must be included in the Project Summary file.
Graphics	PDF, PostScript, gif, jpeg, png		Plots and illustrations to be used as previews for image, light curve, spectra, atlases, (possibly) catalogs. Also appropriate for elements of the project description.
Animation	MP4, WebM		Appropriate for some data product previews, particularly for models or simulations.
Project Summary	Outline-oriented structure: ASCII, HTML, sphinx		One or two paragraph description of the science aims of the data collection, for incorporation into a web site for your collection. May also include a description of the observing program (including sky coverage). Must include literature reference(s) to the methodology used to create the HLSPs. For small HLSP collections this content may instead be folded into the README file.

For **all** data files be sure to:

- follow the HLSP file naming convention.

For **science data** files be sure to:

- include the required HLSP metadata (i.e., keywords) in the headers

Non-Accepted File and Content Types

Proprietary Formats and Certain Content Types

Content that cannot be archived includes: publications, project tar files, or any files in a proprietary format (e.g., Microsoft Office).

Software

It can be useful for contributors to associate software with their data products. Reasons include:

- Documenting the processing software
- Visualizing the data products
- Providing an analysis tool or Jupyter notebook that is tuned to the HLSP collection
- Providing back-end services that operate on the collection data

MAST embraces the idea of associating software with data, but does not accept or support source code.

> The staff at MAST encourage contributors to use third-party software repositories such as GitHub, and to register your software with the Astrophysics Source Code Library (ASCL). A link to your software repository can be included on the MAST website, and the link will be added to our master list of astronomical software that is tied to HLSP collections.
>
> **Please note:** links to a repository must point to a version of the resource (software, notebook, or documentation) that corresponds to the delivered version of the HLSP data products.

Only very limited support is presently available for back-end (analytical) services connected to HLSP collections, but more support is expected in the future.

Certain FITS File Organization

FITS is a fairly general format, and allows for multiple ways to organize data within a single file. However, some potential choices of data organization within FITS files have not seen wide use and are presently not supported for HLSP collections, including:

- Binary tables containing images (any dimensionality) within table cells
  - Table cells may contain simple 1-D arrays, however
- ASCII table extensions (however, BINTABLE extensions **are** supported)

For Further Reading...

- File naming convention
- Required metadata
- FITS Standard