The Journal of Open Source Software

dfitspy: a dfits|fitsort implementation in python

Romain Thomas[1]

1 European Southern Observatory, Av. Alonso de Córdoa 3107, 7630355 Vitacura, Santiago, Chile

DOI: 10.21105/joss.01249

Software

- Review
- Repository
- Archive

Submitted: 29 January 2019
Published: 14 March 2019

Summary

The FITS format (Flexible Image Transport System) (Hanisch et al., 2001) is a widely used format to store astronomical data. It is used to store a lot of different types of data such as 1D or 2D spectra, 3D data cubes. It has been developed in the late 1970 to reach its final form almost two decades ago. FITS files are built with two components. The data themselves are stored as tables and contains any types of data. A header is built containing set of keywords-value pairs aiming at describing the data themselves.

Accessing and displaying metadata inside FITS files is important in order to get an overview of their content properties without having to read the data themselves. It is particularly useful when dealing with large amount of files at once. Tools have been already publicly available for years with the dfits and fitsort algorithms (the documentation is available here https://www.eso.org/sci/software/eclipse/eug/eug/node8.html). The main limitation is that they are stand-alone programs useable only in a terminal. They can not be used natively inside another program.

The python module presented in this paper, `dfitspy`, is a project that migrates the main dfits and fitsort capabilities to python. It is a metadata searcher/displayer for FITS files. As dfits and fitsort, `dfitspy` is able to display in the terminal the result of a metadata search and is able to grep certain values of keywords inside large samples of files. Therefore it can be used directly with the command line interface. Nevertheless, `dfitspy` can be, and it is its strength, imported as a python module and the user can use these functionnalities inside another python code or the python interpretor.

dfitspy as a terminal command

A command line interface has been included in `dfitspy` so it can be used as a Terminal command. A typical command is:

```
dfitspy -f Test_data/* -k author,number,type --grep 2dspec
```

This command will search in all the FITS file present in the *Test_data* directory. `dfitspy` will search for three keywords in the header: author, time and type. Finally, the terminal will display only the file where `2dspec` is in the requested keyword values. The terminal output is similar to the dfits|fitsort combination. It displays, in a column fashion, each file with the requested keyword its corresponding values:

```
filename     author           number        type
----------   ------------     ---------     ------
file1.fits   R. Thomas        49098.26      2dspec
file2.fits   R. Thomas        79098.26      2dspec
file3.fits   R. Thomas        69198.26      2dspec
file4.fits   R. Thomas        79498.26      2dspec
file5.fits   R. Thomas        89098.26      2dspec
file6.fits   R. Thomas        79498.26      2dspec
```

dfitspy as a Python module

To be used as a Python module, `dfitspy` must be imported. Then a set of command have to be used in order to produce the final list of filenames/keywords/values. In short, three main commands must be used:

First of all, import the module:
`import dfitspy`

Then, the files must be gathered:
`listfiles = dfitspy.get_files(['all'],'Test_data/')`

And the list of keywords must be prepared, and eventually the grepping values:
```
listkeys = ['author', 'number', 'type']
grepping = ['2dspec']
```

Finally, we can fitsort the files and eventually grep.
`fitsortgrep = dfitspy.dfitsort(listfiles, listkeys, grepping)`

The final output is stored as a dictionnary of files for which each keywords/values is given. It can also be displayed in the same way as for the terminal output (see above).

Availability

`dfitspy` is a GPL licensed software and the source code is available at https://github.com/astrom-tom/dfitspy. The full documentation is available at https://astrom-tom.github.io/dfitspy/build/html/index.html .

Acknowledgements

The author would like to thank the Journal of Open Source Software to give the opportunity to researchers to publish their softwares and to the referee of this paper for helpful comments.

References

Hanisch, R. J., Farris, A., Greisen, E. W., Pence, W. D., Schlesinger, B. M., Teuben, P. J., Thompson, R. W., et al. (2001). Definition of the Flexible Image Transport System (FITS), *376*, 359–380. doi:10.1051/0004-6361:20010923