Astronomical Data Analysis Software and Systems XIII
ASP Conference Series, Vol. 314, 2004
F. Ochsenbein, M. Allen, and D. Egret, eds.

Chandra FITS Dictionary

S. Winkelman, A. Rots, K. Michaud, S. Blecksmith

Chandra X-Ray Center/Smithsonian Astrophysical Observatory

Abstract. Well-defined data and metadata specifications are fundamental for the operation of large observing facilities. We have collected the complete body of FITS keywords, as used in Chandra data products, into a dictionary database that provides easy access to their use and meaning. We present an account of the functionality of the dictionary, as well as a description of the database design and details of the tools which display the dictionary.

1. Introduction

Well-defined data and metadata specifications are fundamental for the operation of large observing facilities. For the Chandra Data Archive almost all data products are in FITS format, following a stringent set of conventions concerning the data formats and header construction. The header keywords form the core of the metadata for our archive, therefore a proper understanding of the rules and conventions that govern their meaning and usage is essential to a correct interpretation of the data.

We have collected the complete body of FITS keywords, as used in Chandra data products, into a dictionary database that provides easy access to their use and meaning. As such, it provides metadata on the metadata. This is of obvious interest to our users, but also, and particularly, for facilitating multi-mission data analysis. In addition, such dictionaries are invaluable as an aide to projects and missions that are in the process of defining their metadata conventions. Finally, the Virtual Observatory also has a clear need for electronic access to metadata definitions.

2. The Dictionary

The Chandra X-ray Center (CXC) FITS dictionary was designed to provide general definitions of the approximately 300 keywords. In addition, the dictionary also provides more detailed definitions of keywords as they are used in the more than 250 distinct types of FITS files produced in the processing of Chandra data. All definitions include a detailed definition, a descriptive string describing the keyword, and a datatype. To complete the definition, the format, units, maximum and minimum values, allowed values, default values, and notes on usage of a keyword are included where appropriate.

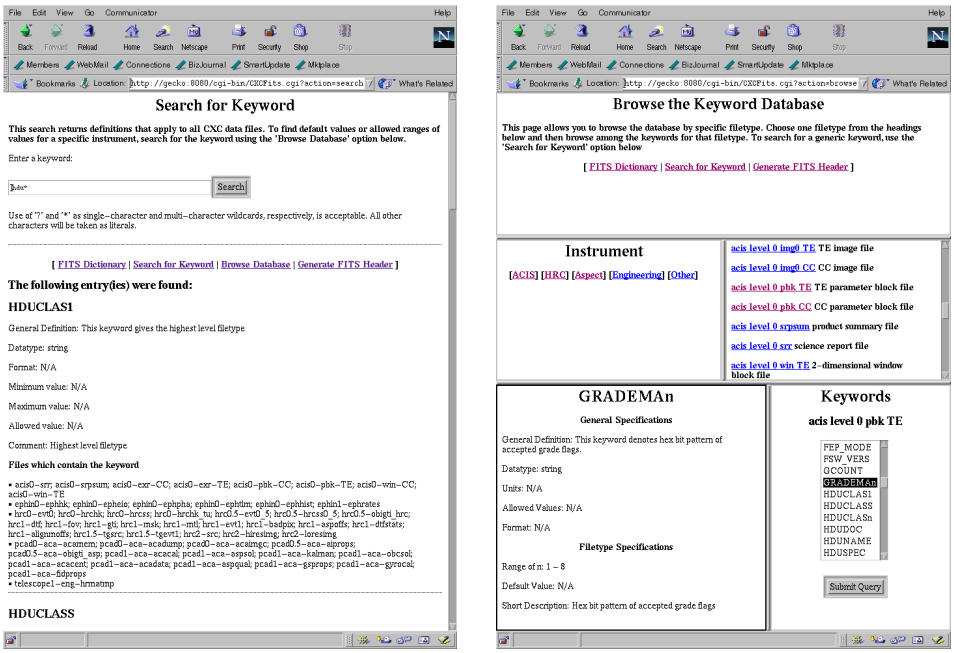

Figure 1. Snapshots of the search and browse screens.

In order to accommodate the various needs of users, the CXC FITS web interface gives users three forms of access to the dictionary.

Search by keyword: This page returns general definitions that apply to all CXC data files. Wildcards can be used in the search. If an exact match cannot be made, close matches will be returned. A list of filetypes using the keyword is also returned. See Figure 1 for a snapshot.

Browse the database: This page allows the user to browse the database by specific filetype. The keyword definition that is returned will be specific to its use in that filetype. For keywords that contain an index (eg., HDUCLASn), each value for the filetype is returned. For keywords that come in groups (eg., TTYPEn, TFORMn, TUNITn, etc.), a link is provided to a table giving the full group definition for the filetype. See Figure 1 for a snapshot.

Generate schematic FITS headers: This page generates a schematic FITS header for a selected filetype. The schematic includes the keywords, default values, FITS comments, datatypes, and formats for each HDU contained in the header. See Figure 2 for a snapshot.

3. The Database

The CXC follows a stringent set of conventions regarding header construction. A filetype is composed of one or more Header-Data Units (HDU). Each HDU's header contains a set of components. These components consist of fixed sets of

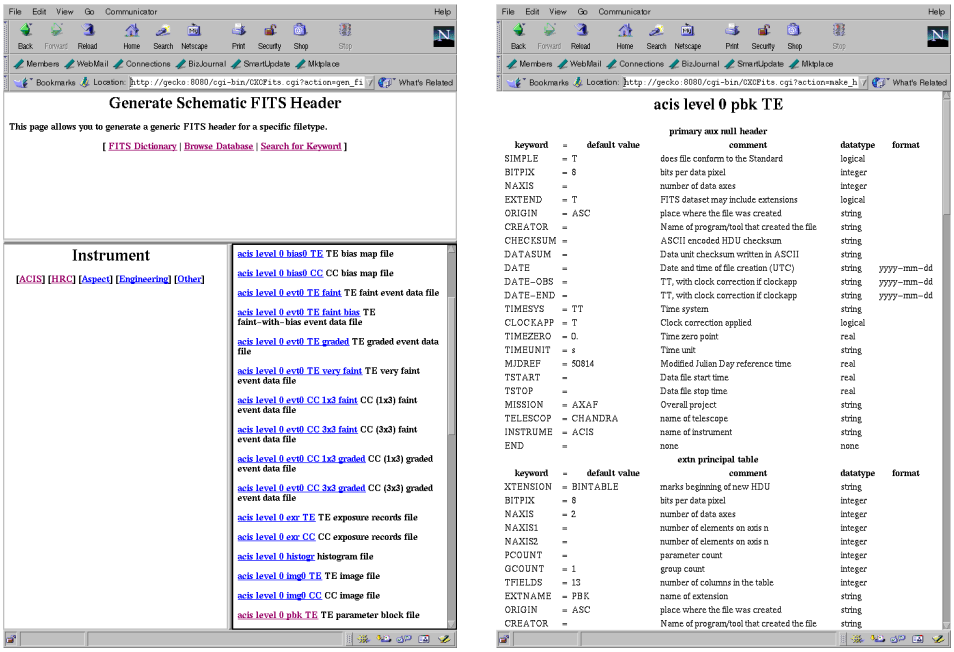

Figure 2. Snapshots of the generate header screen and a sample header.

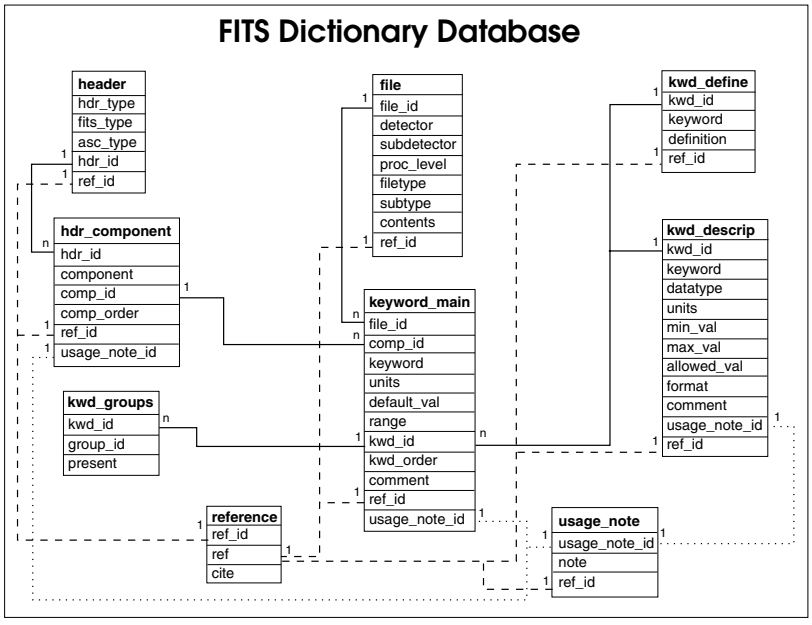

Figure 3. Diagram of the relationship (one-to-one or one-to-many) between the columns used to join tables in the database. The dotted lines are links between the usage_note table and the rest of the database. The dashed lines are links between the reference table and the rest of the database.

keywords. In addition, each filetype may contain additional keywords. There are also ordering requirements for header components and keywords. As a result, the dictionary database became quite involved. The database layout is loosely based on the keyword database design used by the Space Telescope Science Institute. See Figure 3 for a diagram of the database structure. The tables fall into three categories: those which define the keywords to varying degrees of specificity; those which provide locations of keywords within headers; and those which provide references for the information within the database.

4. Conclusion

The database and dictionary interface are mostly mission-independent, possibly requiring minor revisions to one table in the database and several queries in the CGI script to reflect a mission's categorization of its FITS files. As such, we offer our package to any project that has a need for this facility.

Acknowledgments. This work is supported by HEASARC grant 16613384 and NASA contract NAS8-39073 (CXC).

Links

— http://cxc.harvard.edu/contrib/arots/fits/ascfits.ps
— http://cxc.harvard.edu/contrib/arots/fits/content.txt
— http://heasarc.gsfc.nasa.gov/docs/heasarc/fits_overview.html
— http://archive.stsci.edu/fits/users_guide/
— http://www.dpt.stsci.edu/keyword/keyword_design.html