HST Keyword Dictionary

Daryl A. Swade [1], Lisa Gardner, Ed Hopkins, Tim Kimball, Karen Lezon, Jim Rose [1] and Bernie Shiao

Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, Email: swade@stsci.edu

Abstract. STScI has undertaken a project to place all HST keyword information in one source, the keyword database, and to provide a mechanism for making this keyword information accessible to all HST users, the keyword dictionary, which is a WWW interface to the keyword database.

1. Overview

The HST OPUS[2] data processing pipeline receives the telemetry stream from the spacecraft and formats the data as either GEIS or standard FITS files. In both cases header keywords are used to characterize the data. Along with forming the headers for the science data, most keyword values are ingested into HST archive catalog fields where they are accessible for archive searches and data characterization. One goal of this project is to strive for unity in keyword definitions across all HST science instrument data. In order to establish standards for HST keywords and to provide a central location for keyword information, STScI has undertaken a project to place all keyword information in one source - the keyword database.

In addition, a Web-based interface to the keyword database, the keyword dictionary[3], has been developed as a mechanism for making this keyword information accessible to all HST users.

2. Keyword Flow

The keyword database provides the capability to trace the flow of a keyword value from source, through headers, to the archive catalog.

Keyword values for the HST science data are derived from three main sources: the Proposal Management Database (PMDB), spacecraft telemetry as defined in the Project Database (PDB), and calculations performed during data processing. The PMDB contains information derived from the submitted proposals and their scheduling process. The PDB contains mnemonic definitions

---

[1]Computer Sciences Corporation

[2]http://www.stsci.edu/software/OPUS/

[3]http://archive.stsci.edu/keyword/

for downlinked telemetry values and aperture positions. A table in the keyword database links a keyword to the source of its value.

Headers are constructed and the keyword values are populated in the OPUS (science) data processing pipeline and the OMS (engineering and FGS) data processing pipeline. All science data keywords are inserted into the headers during a processing step called Generic Conversion. Some of the keyword values may be left blank at this stage and populated in subsequent pipeline steps. For example, the OPUS pipeline includes the instrument specific STSDAS calibration step that will update some keyword values.

Along with forming the headers for the data, most keyword values are cataloged into HST archive database fields where they are accessible for archive searches and data characterization. Tables in the keyword database associate the header keywords with the archive catalog field into which they are ingested.

3. Keyword Database

The keyword database is a relational database that contains all relevant information about a keyword. It is the definitive source of keyword information for HST data processing and archiving software systems. This information includes the keyword source, datatype, default value, possible values, units, header comment, and textual long description with detailed information. The keyword database also contains all the necessary information on how to build headers from the individual keywords, which archive catalog fields are populated by each keyword, and information necessary to build DDL for StarView screens. A detailed description of the keyword database schema is available through the keyword dictionary Web site.

4. Information Contained in the Keyword Database

The following keyword related information is contained within the keyword database fields.

- Keyword name - name of the keyword following FITS standards

- HST instrument - the instrument onboard HST to which the keyword applies (the same keyword can be used for multiple instruments)

- Datatype - datatype of the keyword value: character, I2, I4, R4, R8, or L1

- Units - the units of the keyword value

- Default value - a default value for each keyword may be specified

- Possible values - a list of all possible values for a given keyword. This may be an enumerated list in which case all values are stored in the database, or a range in which case the minimum and maximum values are stored.

- Source - the source of the keyword value (PMDB, PDB, or calculated; see above)

- Long description - an arbitrarily long text field which describes any relevant information about a keyword or how its value is derived

- Header structure information - all the information necessary to construct each HST data header including which keywords to include and their order

- Optional flag - a flag to identify a keyword as optional. Optional keywords will not appear in the header if their value is blank.

- Short comment - a short (<48 character) description of the keyword for inclusion in the headers

- Archive catalog table - the archive catalog table to which the keyword is ingested

- Archive catalog field - the field in the archive catalog table to which the keyword is ingested

- StarView DDL - information necessary to construct StarView screens

5. Keyword Database Products

The following products are currently generated from the keyword database.

- ICD-19: The HST project document that defines the headers for HST science data files. This document now resides on-line within the keyword dictionary.

- OPUS load files: The OPUS data processing system uses database load files generated by the keyword database to populate tables in the operational database which define the keywords contained in the data headers, the order those keywords appear in the headers, and the source of the keyword values.

- OMS keyword file: In a manner similar to OPUS, the OMS engineering data processing system uses the keyword database to define the contents and order of keywords in observation logs.

- Archive catalog verification file: The archive catalog is verified to insure all keywords ingested into the catalog correspond to the keyword value in the header. This file contains archive catalog keyword/fieldname mapping.

- StarView DDL files: StarView is the interface to the HST archive catalog. DDL files provide the database attributes defined on StarView screens.

6. Keyword Database Configuration Process

A formal configuration management procedure has been established to control modifications to the database. There are three versions of the keyword database active at any one time. Development of new keywords and headers is performed in keyworddev. After development the new keywords enter a test phase and

the contents of keyworddev are copied to keywordtst. New development can continue in keyworddev as the previously developed keywords are tested. After successfully completing the test phase keywordtst is copied into the operational keyword database. In addition to the three versions of the keyword database, the latest version of the OPUS load files derived from the keyword database are configured in a CMS software archive.

7.   Keyword Update Process

STScI has set up a Keyword Review Team (KRT) to consider all proposed changes to HST headers. Proposals to change a keyword or any keyword attribute are submitted to the KRT coordinator. Proposals have originated from a number of sources such as science instrument teams, software developers, system engineers, user support personnel, etc. Once the proposed changes are clearly documented, the proposal is forwarded to the rest of the KRT for consideration and evaluation of any impacts. When approved the change is implemented in the development version of the keyword database, and the revised headers enter into the formal software test cycle.

During development of headers for new science instruments placed aboard HST during servicing missions, the process is streamlined in that entire headers are considered by the KRT instead of individual keywords.

8.   Keyword Dictionary

The HST Keyword Dictionary Web Interface is written in Perl using the Sybperl module. For Javascript-enabled browsers, a small window can be popped up that allows quick searches "on the side". Frames were used where possible in the initial design, but they encumber the interface (and the CGI script) somewhat; the next generation will design them out.

An administration tool was written in Tcl/Tk. (At the time, Web security was judged to be insufficient for an administrative interface.) The keyword database administrator can use this tool to load new or edit existing database records.

9.   Keyword Dictionary Capabilities

The HST keyword dictionary allows for access to keyword information through a number of different approaches. It is possible to search for a specific keyword, browse all the keywords in alphabetical order for a specific HST instrument, or generate schematic headers for each science file generated from HST data. All information about each keyword contained within the keyword database is accessible through any of these routes. It is also possible to look in the development, test, or operational versions of the keyword database.

Since the keyword dictionary accesses the same database that provides the information and instructions on how to construct the headers in the data processing pipeline, this documentation can never be out of date with respect to current data processing!