An Introduction to X-ray Data

… a general introduction to file formats & definitions common to X-ray Astronomy

Mike Corcoran/NICER & HEASARC

Some History

1960's - 1990's: HEA data stored in proprietary formats in private (non-networked distributed) archives

- Data hard to access/transport (example: US ROSAT data in IRAF/PROS format, German ROSAT data in MIDAS/EXSAS format)

- hard to analyze (file converters, big/little endian architecture problems, software installation/compilation issues, library conflicts…)

→ prevented maximum use of science data (long-term variability, multi-wavelength/multi-messenger analyses practically impossible)

The HEASARC

https://heasarc.gsfc.nasa.gov

The High Energy Astrophysics Science Archive Research Center (along with other SARCS) was formed by NASA to:

- **Maintain and disseminate data** from previous and current high energy astrophysics and (later) cosmic microwave background missions;

- Provide **software and data analysis** support for these datasets;

- Maintain and provide the **necessary scientific and technical expertise** for the processing and interpretation of the data holdings;

- Develop and maintain **multi-mission analysis** and support tools

- Provide **catalogs of observations** and ancillary information for the data holdings;

- **Coordinate data, software and media standards** with other astrophysics sites; and

- **Support outreach** in high-energy and cosmic microwave background astrophysics.

Fundamental Standard: FITS

- Flexible Image Transport System: Archive data format mandated by NASA
- Originally developed at NRAO (Wells, Greisen & Harten, 1981) for transport of image date (the “IT” in FITS), Version 4 (latest version) approved 2016
- Simple structure: ASCII header with 80-character “cards”; Reserved keywords (“SIMPLE = T” shows that the file conforms to the FITS standard)
- “Extensions” to primary header/image approved in late 80’s/early 90’s: includes storing data as ASCII or binary tables (image extensions also allowed); variable-length vectors to efficiently store sparse arrays
- added standards for world coordinate system specification

https://fits.gsfc.nasa.gov/fits_documentation.html

Sample FITS Primary Header

```
SIMPLE  =                    T / file does conform to FITS standard
BITPIX  =                   16 / number of bits per data pixel
NAXIS   =                    0 / number of data axes
EXTEND  =                    T / FITS dataset may contain extensions
COMMENT   FITS (Flexible Image Transport System) format is defined in 'Astronomy
COMMENT   and Astrophysics', volume 376, page 359; bibcode: 2001A&A...376..359H
TELESCOP= 'NICER   '           / Telescope (mission) name
INSTRUME= 'XTI     '           / Instrument name
TARG_ID =                 7053 / NICER target catalog ID number
OBSERVER= 'CORCORAN, MICHAEL'  / Observer or Principal Investigator
TITLE   = 'NICER''S CLEAR VIEW OF THE HOT COLLIDING WIND SHOCK IN WR 140' / Scie
OBS_ID  = '6672012301'         / Observation ID
ORIGIN  = 'NASA/GSFC'          / origin of fits file
TLM2FITS= 'NICER2FITSv1.31_FSWv20201222' / Telemetry converter version number
PROCVER = 'l0-master_20230518' / Processing script version number
SOFTVER = 'Hea_22Aug2023_V6.32.1_NICER_2023-08-22_V011a' / NICER Software Versio
CALDBVER= 'xti20221001'        / NICER CALDB Version
SEQPNUM =                    3 / Number of times the dataset processed
CHECKSUM= 'US4naP2kUP2kaP2k'   / HDU checksum updated 2023-08-16T14:50:04
DATASUM = '         0'         / data unit checksum
DATE-OBS= '2023-08-04T07:07:40' / Start date of observations
DATE-END= '2023-08-04T13:30:20' / End date of observations
TSTART  =     302598461.974372
```

https://fits.gsfc.nasa.gov/fits_documentation.html

The HEASARC Conventions

In the 1990's the HEASARC and representation of existing/upcoming missions (the "OGIP FITS Working Group, OFWG) worked to define general data formats for:

- Standard mission/instrument names
- X-ray photon event lists
- spectral file formats (spectra, detector responses/sensitivity)
- timing files
- point spread functions & maps
- weighted map files
- calibration file keywords
- keywords to specify min/max values in table columns

Templates available at
https://heasarc.gsfc.nasa.gov/docs/heasarc/fits_template_example.html

bitly

Implementation

- Standards/conventions first developed for ROSAT and ASCA missions (emphasized imaging, spectral analysis, lightcurves)

- Calibration data (in standard format) created for each mission: primary calibration data (pcf; unformatted); basic calibration files (bcf) & calibration product files (cpf) in FITs; include standard header keywords and indexing scheme

- Underlying IO library (FITSIO -> now CFITSIO)

- Now lots of FITS readers/writers (eg: astropy.io.fits)

Software

Standardization allows re-use of software without need for customized data IO (in principle)

HEASARC developed standard tools to interpret/display/filter FITS data conforming to the HEASARC conventions => FTOOLS (now HEASoft)

- custom tools for new missions (NICER, IXPE, XRISM…) incorporated as needed

CIAO (Chandra Interactive Analysis of Observations) developed by CXC

XMM/SAS developed by ESA for customized processing of XMM-Newton data

Fermitools developed by the Fermi Science Support Center for analysIs of Fermi LAT data

etc

- Complete HEASoft source code (all mission & general-use software) (3.2 Gb / 6.5 Gb unpacked); md5 checksum
- Complete HEASoft source code **plus older Xspec model data unneeded by most users** (4.2 Gb / 11 Gb unpacked); md5 checksum

STEP 1 - Select the type of software:

SOURCE CODE DISTRIBUTION (Recommended):

Please note that the source code distribution is recommended - *particularly for Linux users* - due to portability issues that can affect the pre-compiled binaries. Also, a source code distribution is **required** for users who wish to use **local models in XSPEC / PyXspec**.

- **Source Code**

PRE-COMPILED BINARY DISTRIBUTIONS (May experience portability issues):

Please note that the pre-compiled binaries are **not recommended** - *particularly for Linux users* - due to Perl portability issues. Also, note that users who wish to use **local models in XSPEC** or **PyXspec** must get the source code distribution instead. **Pre-compiled binaries for Silicon Macs are currently unavailable but may be added at a later date.**

- **PC - Ubuntu Linux 20.04**
- **PC - Fedora Linux 39**
- **PC - Red Hat Enterprise Linux 8.9**
- **Mac INTEL (macOS 12 Monterey, or newer)**

STEP 2 - Download the desired packages:

Selecting an individual mission package will automatically select a set of recommended general-use tools.

- All
- Mission-Specific Tools
  - **ASCA** **Einstein** **EXOSAT** **CGRO** **HEAO-1** **Hitomi** **INTEGRAL** **IXPE** **MAXI**
  - **NICER** **NuSTAR** **OSO-8** **ROSAT** **Suzaku** **Swift** **Vela** **XTE**
- General-Use FTOOLS
  - **Attitude** **Caltools** **Futils** **Fimage** **HEASARC** **HEASim** **HEASPtools**
  - **HEATools** **HEAGen** **FV** **Time**
- XANADU
  - **Ximage** **Xronos** **Xspec ***
- XSTAR

Common File Extensions

Data products often have the following file extensions:

Type	Description	Extension	Meaning
Photon Event file	Photon arrival time, channel, location	.evt/(_evt.fits)	“Events file”
Detector Effective Area	Detector area vs. energy	.arf	“ancillary response file”
Detector Redistribution Matrix	probability of a photon of energy E appearing in channel C	.rmf	“response matrix file”
Spectrum	Observed distribution of counts in detector channels	.pi/.pha	“pulse invariant”/“pulse height analysis”
Lightcurve	Distribution of counts in some (energy/channel) band vs. time	.lc	“lightcurve file”
Filter file	Engineering/background/detector properties vs. time	.mkf	“makefilter file”

Common FITS tools

A selection of useful tools to use to view/display FITS data:

Tool	Use
SAOImage/ds9	display FITS images with WCS info; define spatial regions
fv	display images, data tables in spreadsheet-like window; filtering
fdump	display contents of file as text to STDOUT or an ASCII file
xselect	extract basic data products
xspec	spectral analysis

fv, fdump, xselect, xspec, nicerdas, and IXPE software distributed as part of HEASoft

Structure of Spectra

```
% fstruct ni6672012301_0mpu7_cl.pha
  No. Type     EXTNAME      BITPIX Dimensions(columns)      PCOUNT  GCOUNT

   0  PRIMARY                 -32     8 7                          0     1
   1  BINTABLE SPECTRUM         8     8(2) 1501                    0     1

      Column Name                Format     Dims       Units     TLMIN  TLMAX
      1 CHANNEL                    J                                 0     1500
      2 COUNTS                     J                   count

   2  BINTABLE GTI              8     16(2) 10                     0     1

      Column Name                Format     Dims       Units     TLMIN  TLMAX
      1 START                      1D                  s
      2 STOP                       1D                  s
```

Lightcurves

```
% fstruct ni6672012301_0mpu7_cl.lc
  No. Type     EXTNAME      BITPIX Dimensions(columns)      PCOUNT  GCOUNT

   0  PRIMARY                 32     0                           0     1
   1  BINTABLE RATE            8     20(4)  53                   0     1

      Column Name                Format     Dims       Units     TLMIN  TLMAX
      1 TIME                       D                   s
      2 RATE                       E                   count/s
      3 ERROR                      E                   count/s
      4 FRACEXP                    E

   2  BINTABLE GTI             8     16(2)  10                   0     1

      Column Name                Format     Dims       Units     TLMIN  TLMAX
      1 START                      1D                  s
      2 STOP                       1D                  s
```

IXPE Events File

```
In [1]: from astropy.io import fits
In [2]: hdu = fits.open('ixpe01250701_det1_evt2_v02.fits.gz')
In [3]: hdu.info()
Filename: ixpe01250701_det1_evt2_v02.fits.gz
No.    Name      Ver    Type      Cards   Dimensions   Format
  0  PRIMARY       1 PrimaryHDU      62   ()
  1  EVENTS        1 BinTableHDU    217   1502002R x 10C   [J, D, 16X, 16X, J, E, E, E, D, D]
  2  GTI           1 BinTableHDU     53   72R x 2C   [D, D]

In [4]: hdu['EVENTS'].data.columns
ColDefs(
    name = 'TRG_ID'; format = 'J'
    name = 'TIME'; format = 'D'; unit = 's'
    name = 'STATUS'; format = '16X'
    name = 'STATUS2'; format = '16X'
    name = 'PI'; format = 'J'; unit = 'chan'
    name = 'W_MOM'; format = 'E'
    name = 'X'; format = 'E'; unit = 'pixel'; coord_type = 'RA---TAN'; coord_unit = 'deg'; coord_ref_point
= 299; coord_ref_value = 255.2436; coord_inc = -0.00072222
    name = 'Y'; format = 'E'; unit = 'pixel'; coord_type = 'DEC--TAN'; coord_unit = 'deg'; coord_ref_point
= 299; coord_ref_value = -46.1857; coord_inc = 0.00072222
    name = 'Q'; format = 'D'
    name = 'U'; format = 'D'
)
```

Finding Data

Observations of interest can be found by using the HEASARC Xamin (or Browse) interfaces supported by the HEASARC

Chandra data can be found using CHASER (https://cda.harvard.edu/chaser/) as well

Catalogs include:

- Observation catalogs: performed or planned observation information (pointing direction, exposure time, PI, etc)
- Source catalogs: derived source characteristics (automatic or manual processing)

heasarc.gsfc.nasa.gov/xamin/xamin.jsp

HEASARCDEV NASA Box pywebspecdev NuSTAR CCF IXPE caldb fornaxdev nicervisdev FAU sixte EmployeeInfo.aspx xamin Beta Jack_pha ASD Computing LEM

Xamin Web Interface

XAMIN SEARCH

Clear/reset: Tables Target/Constraints Options All | Session Options | Help

Quick Search: Tables, positions, times, ... Query examples

Product Cart: 0 row(s)

Current selections: target=eta carinae x

Send query

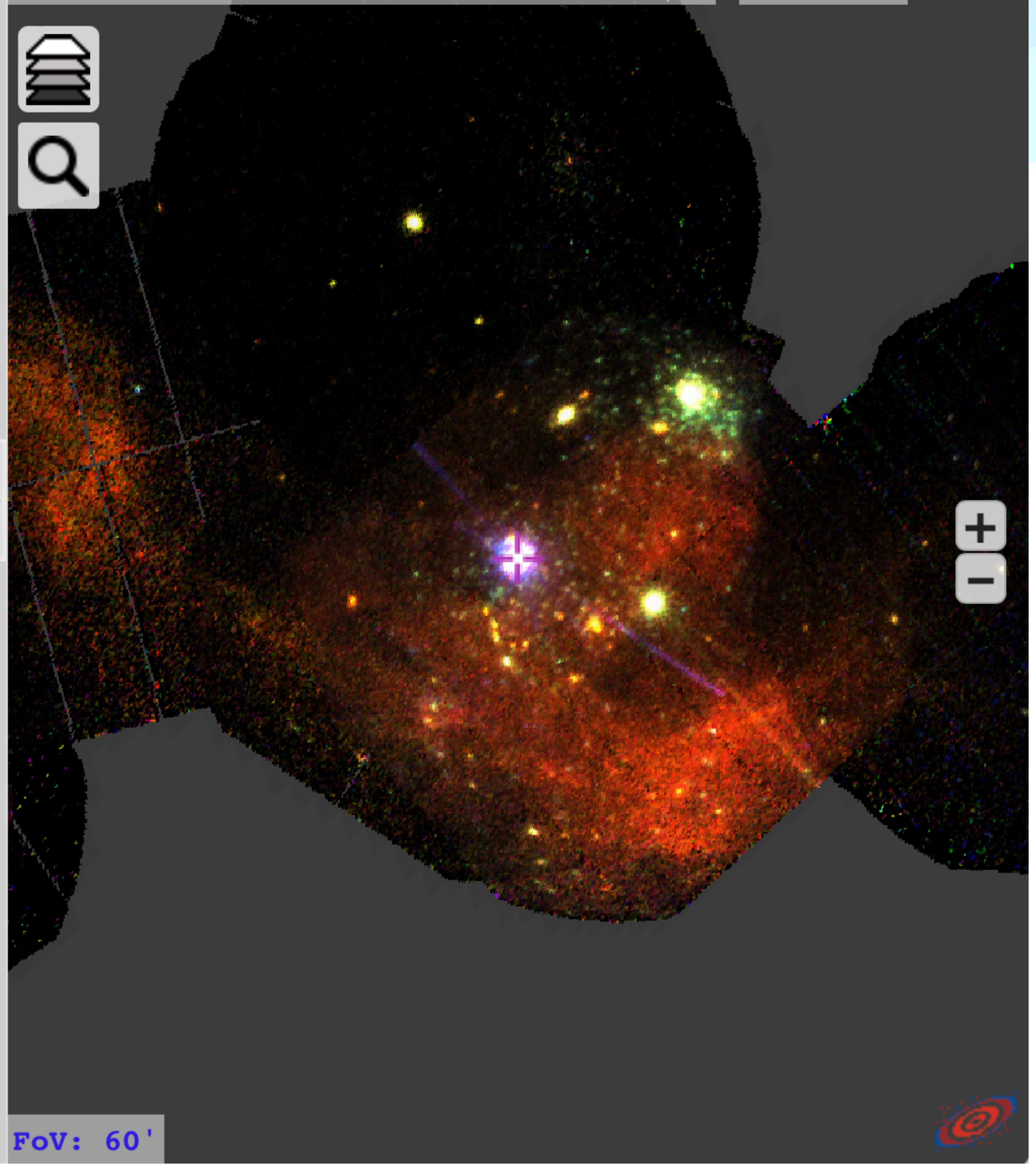
Enter table name here or select from dropdown

Or search for table by keyw Search

Alternatively, go to Available Tables pane below to browse

View: Pan, zoom and shift/click to set query region

J2000 10 45 03.550 -59 41 04.00 SIN



FoV: 60'

Matches in HEASARC Catalogs

l1:Local tables @ eta carinae

	Table	Description	Matches	Regime	Mission	Type
1	a1	HEAO 1 A1 X-Ray Source Catalog	1	X-ray	heao1	
2	a1point	HEAO 1 A1 Lightcurves	1	X-ray	heao1	
3	a2lcpoint	HEAO 1 A2 Pointed Lightcurves	1	X-ray	heao1	
4	a2lcscan	HEAO 1 A2 Scanned Lightcurves	4	X-ray	heao1	
5	a2led	HEAO 1 A2 LED Catalog	1	X-ray	heao1	
6	a2point	HEAO 1 A2 Pointing Catalog	1	X-ray	heao1	
7	a3	HEAO 1 A3 MC LASS Catalog	1	X-ray	heao1	
8	aavsovsx	AAVSO International Variable Star I…	7	Optical		star
9	acrs	Astrographic Catalog of Reference …	1	Optical		star
10	agilecat	First AGILE Catalog of High-Confid…	1	Gamm…	agile	

Conclusions

Standardization offers these benefits:

- minimizes (reduces) learning curve
- allows for re-use of data pipelines, analysis tools
- facilitates multi-(wavelength, messenger) analyses
- faciltates open science

But there are challenges

- difficult to define/implement/enforce
- need for missions to develop custom software
- lack of resources for smaller missions to (re-)format data
- one-size does not necessarily fit all
- advances in technology, specific mission can difficult to anticipate/standardize: need for flexible/extensible standards