

The data center for the X-ray Imaging Spectrometer/Telescope STIX

H. Xiao¹, Shane Maloney², Ewan Dickson^{1,4}, Säm Krucker¹, László Etesi¹, Andrea Francesco Battaglia^{1,3}, Daniel Ryan¹, and Other STIX team members

¹ University of Applied Sciences and Arts Northwestern Switzerland (FHNW), 5200 Windisch, Switzerland
e-mail: hualin.xiao@fhnw.ch

² Astrophysics Research Group, School of Physics, Trinity College Dublin, Dublin 2, Ireland

³ ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

⁴ University of Graz, Universitätspl. 3, 8010 Graz, Austria

ABSTRACT

Aims. The Spectrometer/Telescope for Imaging X-rays (STIX) instrument onboard the Solar Orbiter mission launched on February 10, 2020 promises advances in the study of solar flares of various sizes. It is capable of measuring X-ray spectra from 4 to 150 keV with 1 keV resolution binned into 32 energy bins before downlinking. STIX data center is an infrastructure established at FHNW in order to process and archive STIX telemetry data, and to support the operations of the instrument. The automated data processing pipeline turns raw telemetry data into processed information and data products. Processed information and data products are archived at the data center. STIX data center provides various tools to visualize the information and data products for the solar physics community.

Methods. Methods.

Results. Results.

Key words. Solar flares – Data Center – STIX data products – Data processing pipeline

1. Introduction

Solar Orbiter is a Sun-observing mission of the European Space Agency that addresses the interaction between the Sun and the heliosphere. It was launched on Feb. 10, 2020 for a nominal mission duration of seven years and a planned extension of three years. It carries ten sets of instruments for comprehensive remote-sensing and in-situ measurements. Solar Orbiter will perform detailed measurements of the Sun as close as 0.28 AU and for the first time look at its uncharted polar regions (Forveille & Shore (2020)). Its goal is to address the central question of heliophysics "How does the Sun create and control the heliosphere?". It is designed to identify the origins and causes of the solar wind, the heliospheric magnetic field, the solar energetic particles, the transient interplanetary disturbances, and the Sun's magnetic field. This consists of the study of energetic solar phenomena like flares, solar transients, the solar wind accelerating mechanisms, and the solar dynamo principle.

The Spectrometer Telescope for Imaging X-rays (STIX) is one of the ten instruments onboard Solar Orbiter. It measures X-rays from 4 to 150 keV and takes X-ray images with a few arc-sec angular resolution by using an indirect imaging technique, based on the Moiré effect. The angular resolution is 7 arcsec and the spectral resolution 1.1 keV full-width-half-maximum at 30 keV. Its instrument consists of 32 collimators with grids and 32 pixelated Cadmium telluride detector units called Caliste-SO Krucker, Säm et al. (2020). STIX's main purpose is to study the extremely hot solar plasma and the high-energy electrons accelerated during solar flares. It will address the key science goals of the Solar Orbiter mission by providing information on intensity, spectrum, timing, and location of accelerated electrons near the Sun.

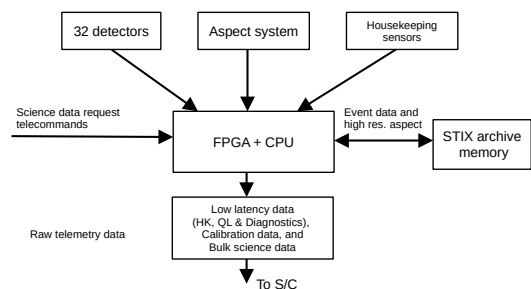


Fig. 1. STIX onboard data flow and data types.

STIX has more than two hundred different telemetry data types. The data have complex data structures and are highly compressed. Being aware of the complexity of the data analysis and of the need to bring the data to the community, a data center is developed at FHNW in order to receive, analyze, archive and distribute the STIX data. The data center turns raw telemetry data into processed information and data products that can be used for scientific analysis. SDC also provides various data visualization tools for the solar physics community. We will describe here STIX raw data types, the flow of data from the users, the main data processing steps, the data products and the tools provided for the community.

2. STIX raw telemetry data

During the normal operation, STIX continuously collects data from the 32 x-ray detector units, the aspect system and house-keeping sensors. The event data from the x-ray detectors are processed in three parallel paths: 1) a primary data path transfer

time-binned event data to the instrument archive memory; 2) a quick-look path generates light curves, snapshots of energy spectra and flare triggers; 3) a calibration path collects events during quiet periods and uses them to form calibration spectrum for each pixel;

Event data and high time resolution aspect readouts are stored in STIX archive flash memory. The processed data is stored on the on-board archive memory for further processing as requested by ground commanding, or used to form telemetry data packets. As shown in STIX raw telemetry data can be classified in four categories: Housekeeping (HK), Quick-look (QL), Diagnostic, and user requested bulk science data. HK data and QL data are directed to the low latency data store in the Solid State Mass Memory (SSMM) of the spacecraft. They are sent down to Earth when during ground antenna passes.

3. Overview of the data flow

STIX telemetry data are first be stored in the spacecraft solid state disk. They are transmitted to ground stations during antenna passes. Telemetry data received by ground stations are processed at the mission control center of ESA and stored in . ESA provides a web interface. The typical latency of the low latency data before they arrive at STIX data center is about x day when there is a pass a. Raw data received at STIX data center has the same structure.

The payload data must be created through the EEDS web page to create a data download request. After receiving the user's data request, EDDS displays the original payload data in the form of hex code and sends it to the stix data center server through the RSYNC protocol. Users can define the time interval of data request. Generally speaking, STIX server requests data from EDDS once a day. The data STIX obtains from EDDS is a HEX code with a time stamp, which is the same as the data sent by STIX to the spacecraft platform after being converted into binary data.

In addition to telemetry data, STIX data center also receive spacecraft ancillary data from ESA, that contains information of spacecraft ephemeris and factors for time conversion.

4. Data processing pipelines at STIX data center

4.1. Data link and data reception

4.2. STIX raw data processing pipeline

4.2.1. Raw data parsing

include a decompression error map here

4.2.2. Background monitoring

Light curves measured during quiet periods of the sun are used for background estimation. Median values and of counts are computed and considered as the background in the selected time frame. They are stored in a database and used for flare identifications.

4.2.3. Flare identification

Quick-look light curves in the energy range 4 to 10 keV are used for solar flare identification. The flare identification procedure consists of two steps:

- Light curve smoothing. In order to filter spikes from electronics and to reduce the amount of variation due to statistics and the onboard integer compression, light curves are smoothed by using the average filtering. The
- Flare identification. Local maximums are selected from smoothed light curves. A local maximum is considered as a flare peak if the counts are exceeds 2 standard deviations above the background and the duration above the background is longer than 1 minute.

For each of identified solar flares, the information such as start time, end time, counts, background subtracted counts is stored in the STIX flare database. It is used for automatic creation of data requests for on-board archived data.

4.2.4. flare ID naming convention

4.2.5. Flare locations

4.3. Calibration data processing

Ba134 radioactive sources with a total activity of about 4000 Bq are placed at the front of each detector. The total activity of the radioactive source is approximately 4000 Ba. When the radioactive source decays, gamma rays are generated. These gamma rays can form peaks in the energy spectrum of the detector. The corresponding energy is known, and the corresponding relationship between ADC and real energy can be calibrated through the position of the peak in the energy spectrum, that is, the calibration coefficient. The figure below is a typical Ba133 gamma-ray energy spectrum measured by STIX CdTe detector. There are three obvious peaks in the energy spectrum, and they correspond to three energies of 30 keV, 35 keV and 81 keV. There are many ways to determine the position of each peak, you can use the ECC method, or use the Gaussian function to fit the left part of the peak.

4.3.1. Flare location using coarse flare data

4.3.2. Flare classification using GOES x-ray flux

4.3.3. FITS file creation

5. STIX data products

- Raw data.
- L1 data.
- L2 data

The latest level 1 FITS IO from Shane has been integrated into the data processing pipeline on pub023 server.

I have recreated fits products for all old telemetry data with the upgraded SW.

The L1 fits files created by this pipeline have a different data level: L1A ('A' here means prerelease/alpha version).

The idea behind L1A data sets is to allow for quicker access to STIX data in the fits format instead of grabbing data from plots or using JSON requests, for operations, debugging etc.

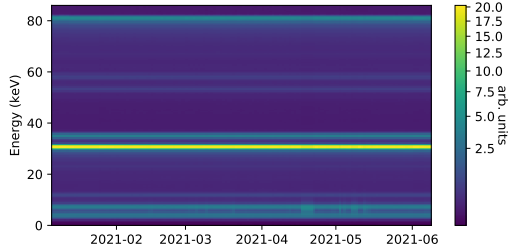
The L1A data sets can be generated within a few minutes after the arrival of a new raw telemetry file.

The differences between L1A and L1 available in Shane's ftp include:

1. Two different L1A data files may have duplicated data
2. L1A data sets are still created for incomplete packets (L1 checks for data completeness)

Table 1. Level 1 data products

Category	Type	Naming convention	Remark
Housekeeping	hk_mini		Housekeeping
	hk_maxi		Housekeeping
Quick-look	light curves		Quick-look
	variance		Quick-look
	spectra		Quick-look
	background monitor		Quick-look
Calibration data	flare location		Quick-look

**Fig. 2.** Caption

3. SPICE kernel data for telemetry files always arrives after one or two days later. So there may be a sub-second difference between the UTC time in fits files (same to times on web pages) and the real time.

Shane's formal L1 release can avoid this issue if they are produced on a later date.

4. L1A contains housekeeping data

The Level0 archive contains TM which has been parsed or decommuted into readable structures but no additional external information is include:

times are not converted to UTC no calibration or conversions applied for STIX we need to decide if we decompress / combine X-ray L0 the count/trigger data at this stage or in the next level L1

copy manual, tree like, json formats name, raw value, eng value, children look-up table, to know description

estimate mongodb benchmark MongoDB benchmark, key value, index, performance

6. Data request procedure

created of data requests manuall checked **data request naming convention yyddmm00 to yyddmm 11**

7. Flare processing

8. Database

8.1. Raw data packet database

8.2. Configuration database

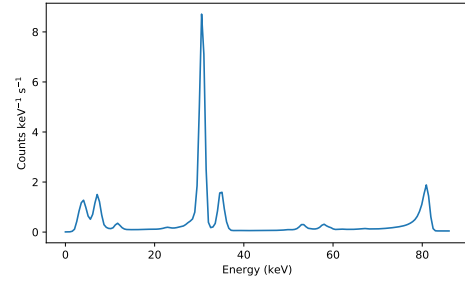
9. Online data visualization tools

9.1. Quick-look light curve

9.2. Science data quick analysis

9.2.1. Calibration data

calibration data products <https://fermi.gsfc.nasa.gov/ssc/data/access/gbm/>

**Fig. 3.** Spectrogram of STIX calibration spectra

9.2.2. Solar Orbiter orbit viewer

When a new data file from the platform is received at the PPDC, it triggers an autonomous start of the dedicated program that decodes and interprets its contents. The binary data contain the spacecraft location, attitude, speed, and GPS timestamps with increments every half second. The GPS timestamps are converted into Unix-timestamps, where the leap seconds are also considered. After processing, the platform data are written to the ROOT format files. The data start and stop time, data processing time, input filename and ID of the output file of each processing are recorded in a dedicated database table. SPICE kernel

Updated once per day.

At the center of the Sun. It is worth mentioning that has to be corrected for. This can be done by using the web tool provided at the auxiliary data center at

10. Data access and APIs

11. Future work

12. Conclusions

References

- Forveille, T. & Shore, S. 2020, A&A, 642, E1
 Krucker, Säm, Hurford, G. J., Grimm, O., et al. 2020, A&A, 642, A15