Data processing pipelines and the data center for the X-ray spectrometer/imager STIX onboard Solar Orbiter

Hualin Xiao

Shane Maloney

Säm Krucker

Paolo Massa

Lastufka Erica

Andrea Francesco Battaglia

Ewan Dickson

Frederic Schuller

André Csillaghy

Ryan Daniel

László Etesi

Nicky Hochmuth

Collier Hannah

Gordon Hurford

Olivier Limousin

Astrid M. Veronig

Alexander Warmuth

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. It purposes include processing and archiving STIX data, supporting operations of the instrument and providing data and tools to the solar physics community. It consists of databases and automated data processing pipelines, various tools for data analysis. The automated data processing pipelines turn raw telemetry data into processed information and data products. Processed information and data products are archived at the data center.

Results.

# 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 (). Its goal is to address the center 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 arcsec angular resolution by using an indirect imaging technique, based on the Moiré effect. STIX consists of 32 collimators with grids and pixelated Cadmium telluride detector unit Caliste-SO. The main science objective of STIX is to study the extremely hot solar plasma and the high-energy electrons accelerated during solar flares. STIX provides intensity, spectrum, timing, and location of accelerated electron information of solar flaresn. For details on STIX instrumentation and its scientific capabilities, we refer to the instrument paper ().

During nominal science operations, STIX continuously acquires data. They are compressed and formatted into different types of telemetry packets onboard by STIX onboard data processing unit (IDPU). The total number of telemetry packets reaches hundreds. Being aware of the complexity of STIX data analysis and the need of bringing the data to the solar physics community, a data center was developed at FHNW. The data center software consists of a set of software packets, which allow The data center (SDC) receives, analyzes, archives and distributes STIX data. It also supports STIX in-flight operations. It turns raw telemetry data into processed information and produces data products that can be used for scientific analysis. It also provides various data visualization tools to the solar physics community. The purpose of this paper is to describe the processing pipelines, the main algorithms of automated data processing, data products and tools provided by the STIX data center. STIX data center processes STIX telemetry to generate a set of widely usable products, as well as performing a quick-look analysis to access the data quality and discover solar flares. Data are distributed to users and archived at STIX data center. STIX data center also provides various data analysis tools, and giving support to users.

This paper is organized as follows: Section [2](#sec:raw-data) briefly introduces STIX raw telemetry types, data flow and telemetry processing pipelines. At the end it is a summary [7](#sec:summary).

# STIX raw telemetry data

STIX continuously observes high-energy events on the Sun at 4 – 150keV. Photons emitted by solar flares are detected with 32 subcollimators (a 12-pixel-detector behind a front and rear grid). While passing through the front and rear grids, the flares generate a modulation pattern over the 12 pixels of each detector. The pattern can then be used to reconstruct images and to do spectroscopy. Other data products include lightcurves, flare information, spectra, and background information. The nominal telemetry budget of STIX is 50 bps. STIX is far from earth, not all data can be downloaded from the spacecraft. We have low latency data. For bulk science data have to be requested. STIX continuously collects energy deposition information from 32 detector units, the aspect system, and engineering sensors in the nominal observation mode. The collected data are first processed by the FPGA and the onboard flight software After the prompt processing, low latency telemetry data are directed to the storage in the spacecraft whereas high time resolution pixel data are written to STIX onboard archive memory for later processings. STIX transmits data to the spacecraft in the form of binary packets. STIX raw telemetry data can be classified into four categories: housekeeping data, diagnostic data, quick-look data and science data.

In the next sections, we briefly introduce the main raw telemetry data types. For details about STIX raw telemetry data, we refer to STIX instrument paper ().

## Housekeeping data

STIX housekeeping data (HK) contain engineering data that measure temperatures, voltages, currents, the status of switches, averaged signal readouts from the four photodiodes in the aspect system, detector trigger counts, and the flags indicating status bits of the onboard software and the archive memory. The data is used to monitor instrument status, and the instrument pointing. HK data are generated as long as STIX is powered on. During nominal operations, a housekeeping packet is generated every 64 seconds, which results in a daily raw telemetry of 143 kiB. All STIX sensors and the IDPU will produce housekeeping data that will be received on the ground with the highest priority for monitoring instrument status and health.

## Quick-look data

STIX has only one operational mode in which Low Latency data is produced, and that is NOMINAL mode, which is the regular science mode. STIX generates four different types of quick-look (QL) data:

* STIX quick-look light curves. STIX quick-look light curves are time series of detector-sumed counts in five energy bands, with a default integration time of 4 seconds. Counts are not corrected for dead time, transmission, impacts of the presence of the attenuator or rate control regimes. STIX quick-look light curve data are transmitted to the ground as other low latency data. Typically, STIX data centers receive them within tens of minutes to two days after being generated on board.
* Background monitor light curves. Background monitor light curves are similar to the Quick-look light curves but only counts recorded by the background detector pixels are included, and the integration time is 8 seconds.
* Variance data are onboard computed variance of 40 successive detector-summed count rates based on 0.1 sec integration.
* Quick-look spectra, which are energy snapshots of energy spectra (32 science energy bins) for each of the 32 detectors with 32 sec exposure time. STIX takes a snapshot of energy spectra every 1024 seconds in the nominal observation mode.
* Calibration spectra. Calibration spectra are accumulated for events from 128 weak onboard Ba radioactive sources (a total activity of 4500 Bq) when the solar count rate is small compared to the background. Identification of such quiet periods is done autonomously, based on the presence of a TC-specified time gap between two successive photons. ADC readouts of each channel are accumulated in individual spectrum. After an accumulation a long period (typically 24 hours), the spectra are formmated and redirected to in the quick-look low-latency data.

## Diagnostic and event data

When a failure is detected directly by STIX, the instrument’s response includes sending a “dedicated” telemetry event report with appropriate diagnostic information included.

## Bulk science data

Bulk science data are different combinations of summing and compression of the basic pixel data stored in STIX onboard archive memory, which can be used for spectrocopy and imaging. To cope with the limited available telemetry, they are not automatically included in the telemetry. Onboard formation of the science data is invoked by data request telecommands. Each science request can select subsets of energies, pixels and detectors. Six different science data can be generated:

* Raw pixel data is the least processed data and contains uncompressed counts for the selected energies, pixels and detectors;
* Pixel data is essentially the same as Level-0 but the counts are compressed onboard before being sent to the ground;
* Compressed pixel data are further compressed counts from the L1 pixel data, in which are summed down to 4 before compression.
* Visibilities are further reduces the data by combining the four summed pixel counts into complex visibility which is also compressed.
* Spectrograms are detector summed spectrograms;
* High-time resolution aspect data.

It should be noticed that the data levels described above are on-board data levels. They are sent to the ground in the format of raw binary packets and considered as level-0 at STIX data center.

[tb:raw\_types]

Housekeeping data, quick-look data and calibration data are directed to the common low latency data stored in SSMM in the spacecraft. The coverage, data rate and latency of different types of STIX raw telemetry data are summarized in Table [[tb:raw\_types]](#tb:raw_types).

# Data reception

During nominal science operations, Low latency data are down-linked in the very next ground station pass regardless of orbital geometry, whereas science data are only down-linked when the bandwidth permits. The downloaded instrument telemetry data are first processed by ESA’s ground segment software at the Solar Orbiter mission control center. Then they are distributed to instrument teams by the ESA EGOS Data Dissemination System (EDDS) () according instrument teams’ preset conditions.

Telemetry data received by STIX data center from EDDS are in the binary format, and have the same information content as the original telemetry generated by STIX. The low-latency data arrives at STIX data center with delays ranging from a few minutes to a few days, depending on whether there are antenna passes, whereas science data may arrive several weeks after being generated onboard.

In addition to the telemetry data, STIX data center also receives SPICE kernels () from the science operations center. They contain information of spacecraft ephemeris and clocking calibration factors required for time conversions.

# Data processing pipelines

## Data processing pipeline overview

![image](data:application/pdf;base64,)

New telemetry data arriving at STIX data centers is immediately processed by the pipelines as shown in Fig. [[fig:main\_pipelines]](#fig:main_pipelines). The processing is started from raw packet parsing. Parsed level-1 packets are written to a NoSQL database. Then they are selected are processed in 4 different paths, In the first path, housekeeping, quick-look and science packets are selected and used to create L1-A FITS files, which can be used for scientific analysis. The second path selects calibration data from the database and extracts energy calibration factors for instrument monitoring. In the third path, quick-look packets are selected and used to identify solar flares; The fourth path processes flaring data, generating higher data products.

## Raw data parsing and database

New telemetry data arrived at STIX data center are immediately processed by STIX data parser, which is capable of parsing all types of STIX binary telecommands and telemetry packets. Parsing of packets is based on the mission interface database (MIB), which contains information on packet parameters such as names, descriptions, lengths, data types. The parser extracts header and parameters for each raw binary packet using the information in the MIB. Packets after parsing contain raw values of parameters. Spacecraft clock times are further converted to UTC times using the latest version of SPICE kernel data (), compressed integers are decompressed by using a LUT table and housekeeping raw values are converted to human-readable values using the calibration factors or look-up tables stored in the MIB. Packets after the above processing steps are called level-1 packets. They have tree structures, each containing two nodes "header" and "parameters"; the former contains the basic information of the packet such as timestamp, packet type, and the latter contains names, raw values, engineering values (or decompressed integers) and child nodes of a list of parameters. Level-1 packets are written to a NoSQL database. NoSQL database is schemaless, making it ideal to store data with complex structures but small sizes like STIX level-1 packets. NoSQL is also

In addition to packets, the NoSQL also stores other data extracted by the parser during packets parsing, for example,

* Raw file metadata. It contains filename, reception time, observation start time and stop time, MIB and SPICE kernels used by the parser ;
* Bulk science metadata, which contains the observation time range, detector and pixels masks, energy ranges, bulk science data types, and IDs of packets in the database;
* Quick-look light curves. The dataset contains data of counts, time range, energy bins extracted from QL LC packets, allowing for quick access to QL data with web pages or APIs;
* STIX instrument configurations. It contains the instrument configuration parameters for such as energy conversion factors and AISICs, allowing fast tracking of the instrument settings;

The information are necessary for subsequent processing. The database is accessible through web pages at STIX data center website or python APIs.

## Data products

### Level-1 FITS files

After each new raw telemetry file is being parsed, housekeeping, quick-look, and science packets are selected from the NoSQL database successively and checked for data integrity and consistency. Then packets of the same type are merged for creations of pre-release of level-1 data products (Level-1A) in the FITS format (), which is a portable file standard widely used in the astronomy community to store images and tables. It should be noted that the data levels mentioning here are data processing levels. In order not to confuse two conventions, we use different notations to indicate onboard level in this paper. Metadata such as observation time range, creation time, filenames, checksum are written to the primary header data units (HDU) of fits files. They are also written to a dataset in the database.

Level-1A FITS files are generated automatically and available at STIX data center within minutes after the reception of a raw file. It should be noticed that predicted SPICE kernels may be used when the reconstructed spice kernels are not available, and the L1A data are subjected to some known issues due to bugs in the early version of the flight software. After all resources are validated, Level-1 FITS files are created again. Level-1 FITS products have almost the same data structures as Level-1A.

### Level-2 FITS files

Level-1 FITS files are further processed. Dead time correction and energy this make up the second

### Level-3 data products

Level-3 data products are visibilities and images

## Energy calibration

Energies deposited by x-rays are converted from ADC units to keV onboard by using an energy look-up table (ELUT), which is regularly updated with telecommands. ELUTs can be constructed using energy conversion factors obtained by studying positions of photo peaks in the calibration spectra measured for the onboard Ba sources. As energy conversion factors may change with temperature or polarization effects, calibration spectra need to be analyzed promptly on the ground.

![An example of STIX in-flight calibration count spectrum. The three strongest lines, from left to right, are from 31 keV, 35 keV, and 81 keV photons. The first two peaks are fitted by the double-Gaussian function and the high energy peak by the crystal-ball function. ](data:application/pdf;base64,)

An example of STIX in-flight calibration count spectrum. The three strongest lines, from left to right, are from 31 keV, 35 keV, and 81 keV photons. The first two peaks are fitted by the double-Gaussian function and the high energy peak by the crystal-ball function.

The right panel of Fig. [1](#fig:cal-fit) shows an example of STIX count spectrum from in-flight calibration run. The three strongest lines are from 31 keV, 35 keV and 81 keV photons. The first two peaks are fitted with the double-Gaussian function and the third peak by the crystall-ball function (), which consists of a Gaussian core portion and a power-law low-end tail, below a certain threshold. Then a linear line is fitted to the relation between the three fitted peak positions (in ADC unit) and the photon energies in units of keV. The interception and slope indicate the baseline and gain of the channel. The results were compared to those determined using the ECC method (see )), which were found to be consistent within errors. The above steps are performed for each calibration run automatically once the data is available at STIX data center. The results are written to a dataset in NoSQL database. It is accessible through a web page at STIX data center or python APIs.

Calibration factors are monitored continuously. Once significant changes are observed from those used for the construction of the onboard ELUT, a new ELUT is created and then uploaded to STIX after being validated by the operations team.

## Solar flare identification

The in-flight software identifies solar flares by comparing the count rate of the background detector with that of other detectors and packs the results into the QL flare flag and location reports. However, the reports only provide limited information on flares due to the constraints of telemetry bandwidth and the limitation of onboard computing resources, and micro-flares are not reported due to the relative high trigger threshold. Using QL light curves, solar flares can be identified on ground in great detail.

QL light curves of the energy range from 4 to 10 keV are used for solar flare identification on the ground. Identification is based on the fact that the background event rate in that energy range is almost constant over a timescale of days during quiet sun periods. The procedure includes the following steps:

* Light curve smoothing. The selected light curve is filtered using an unweighted moving average filter with a time window of 1 minute to smooth out statistical fluctuations and electric surge spikes;
* "Background" subtraction. A flare may last hours and may have short-duration pulses lying on the main pulse in the light curve. The main pulse can be considered as the "background" of those short-duration pulses. In order to identify those short-duration pulses and estimate their durations, we estimate the "background" using the Statistic-sensitive Non-linear Iterative Peak-clipping algorithm (SNIP), which is widely used in X-ray spectroscopy. Then it is subtracted from the smoothed light curve.
* Identification of flare peaks. Peaks with counts exceeding a threshold of above the quiet Sun background count rates, are selected. The duration of a peak is given by the time difference between the first crossing of the threshold on the left and right sides of the peak;
* Merging of flare peaks. If the time difference between the two consecutive peaks is less than 5 minutes, they are considered to be from the same flare.

![STIX 4 – 10 keV QL light curve recorded from 2022-08-10T21:00:00Z to 2022-08-10T18:00:00Z and identified flares. The light curve was smoothed using a moving average filter with a time window of 1 minute. The identified peaks are marked with plus signs, and flare time ranges are colored in cyan. ](data:application/pdf;base64,)

STIX 4 – 10 keV QL light curve recorded from 2022-08-10T21:00:00Z to 2022-08-10T18:00:00Z and identified flares. The light curve was smoothed using a moving average filter with a time window of 1 minute. The identified peaks are marked with plus signs, and flare time ranges are colored in cyan.

As an example, Fig. [2](#fig:flare-det) shows the 4 – 10 keV QL light curve recorded from 2022-08-10T10:00:00Z to 2022-08-10T18:00:00Z, as well as the smoothed light curve and identified solar flares (marked with plus signs).

For each identified solar flare, start time, end time, peak time and peak counts are extracted from the selected light curve and assigned an 8-digit identification number of the format yymmddHHMM, which represents the solar flare peak time. For example, the identification number 2201010000 indicates that the solar flare peak time is at 2022-01-01T00:00:00 UT. The above steps are repeated for the QL light curves of other four higher energies in the same time frame. This can provide information on the upper limit of the X-ray energy produced by the flare, which is used to optimise the selection of scientific data to be downloaded. Then all extracted information as well as the ephemeris data calculated for the peak time are written to a dataset called solar flare list in the NoSQL database. The flare list database can be queried using web interface or stixpy.

## Estimation of background level

As mentioned earlier, a threshold value calculated using background data needs to be provided when performing flare identification. The background, although stable within a few days, can change over long periods of time, a suitable threshold is important for the accuracy of the identification. Therefore, QL light curves for quiet sun periods are selected by excluding flaring periods. Then median values and variances are calculated from the selected light curves and written to a dataset in the database. They are used as inputs of the next flare identification. Those processing steps are performed automatically after flare identification.

## Solar flare data analysis pipeline

### Flare GOES class determination

![Scatter plot of GOES low channel peak flux with respect to STIX 1-AU equivalent peak count rate in the 4 – 10 keV range for 717 solar flares observed by both GOES and STIX duration the commissioning phase. The solid line is a linear fit to the log-log graph. From the fit, we obtained the GOES flux (in units of W/m^2) f = 10^{0.622 -7.376 \log_{10} (X^{'})}, where X^{'} is STIX peak count rate corrected for the distance variations between the Sun and Solar Orbiter. ](data:application/pdf;base64,)

Scatter plot of GOES low channel peak flux with respect to STIX 1-AU equivalent peak count rate in the 4 – 10 keV range for 717 solar flares observed by both GOES and STIX duration the commissioning phase. The solid line is a linear fit to the log-log graph. From the fit, we obtained the GOES flux (in units of W/m) , where is STIX peak count rate corrected for the distance variations between the Sun and Solar Orbiter.

It is straightforward to calculate the GOES class of a flare observed by STIX from GOES flux if it is observed by GOES. However, most of the time Solar Orbiter is far away from Earth and looks at the sun from different angles. Therefore a considerable number of flares observed by STIX are not observed by near earth satellites (and vice verse). As already discussed in Ref. , it is possible to estimate GOES classes by using STIX count rates. In order to study the correction, we selected 717 solar flares observed by by both GOES and STIX during the commissioning phase. Fig. [3](#fig:goes-stix) shows the scatter plot of GOES low channel peak flux with respect to STIX peak count rate of the 4 – 10 keV QL light curves. STIX count rates have been subtracted for background, and corrected for the difference distance of Solar Orbiter to the Sun using , where is the count rate after background subtraction and the distance between Solar Orbiter and the Sun in units of AU. As can be seen in the figure, there is clear correlation between these two quantities. The wide spread at the low flux could be explained by the difference in the energy response of the two instruments and in the flare temperatures. A linear line was fitted to the correlation in the log-log scale. From the fit, we obtained the GOES flux in units of W/m . The formula is currently used to estimate GOES classes of flares observed by STIX. It is worthwhile to mention that more observations will be included into the fit. The estimated GOES classes as well as those from GOES measurements are stored in the flare list database.

### Coarse flare location estimation

Flare centroid location can constraint flaring region when reconstructing flare images. STIX uses a dedicated sub-collimator called Coarse Flare Locator (CFL) to estimate flare locations. The CFL consists of a front grid with a distinctive pattern which selectively illuminates pixels of a dedicated detector based on the source location. Flare location is estimated onboard by maximizing the correlation between observed CFL pixels counts with expected counts using a look-up table. However, due to the constraints of onboard computation, the onboard flare location is only calculated for intermediate flares and has an accuracy of about 2 arcmin. With the downloaded measured counts in each pixel, the coarse flare location can be reconstructed to on the ground as well. This allows for more sophisticated algorithms, greater flexibility of selecting time and energy range to be integrated, and more careful background subtraction.

![image](data:application/pdf;base64,)

The coarse location of an observed flare is estimated on ground after its pixel data (onboard level-1) is parsed. The procedure consists of the following steps:

* Selection of pixel data based on the flare time range and energy range information stored in the flare list dataset;
* Subtraction of background. The most recent level-1 dataset for the quiet sun period is used for background subtraction;
* Calculation of the mean fluence except for the CFL and the background detectors; Except for the those two detectors, the ratio between the illuminated area and total area only depends on the slit width and the pitch width. It doesn’t vary with the position of the flare in the STIX FOV. Therefore, the fluence of a detector can be given by $F=c/(r S\_{\rm d})$, where is the registered count and the sum of the sensitive areas of twelve pixels.
* Calculation of the areas of illuminated regions on CFL pixels using , where is an array of counts recorded by 12 pixels. Their errors of the areas are also calculated;
* Coarse flare location is estimated by minimizing the weighted sum of squared deviations (i.e. weighted chi-squares) between the calculated illuminated areas and expectations simulated for potential flare locations in a 400 400 grid, whose locations are separated by 10 arcsec.
* Saving the calculated flare location to the flare list dataset.

As an example, the left panel of Fig. [[fig:cfl]](#fig:cfl) shows the calculated and best-fit illuminated areas of the CFL pixels measure observed for STIX flare 2105071900 (GOES class M3.9); the middle panel shows the best-fit flare centroid location, as well as its , and contours. The simulated shadow of the CFL sub-collimator is shown in the right panel.

### Image reconstruction and spectral analysis

![image](data:application/pdf;base64,)

STIX detects thousands of solar flares each year. However, only a part of them are analyzed in detail by solar physicists. To facilitate the selection of flares of interest, a flare imaging reconstruction and spectral fitting pipeline has been developed and integrated into the main data processing pipeline on the platform. For each flare, pixel counts recorded within one-minute time frame around the flare peaks are selected. Reconstruction of an image require inputs such as background data in the same level, STIX pointing information, spacecraft orientation. They are prepared based on the knowledge in the NoSQL. After background subtraction, transmission and dead time corrections, the selected counts are used for calculation of visibilities for two energy ranges 4 – 10 keV (thermal emission) and 16 – 28 (non-thermal) keV. Then images are reconstructed with four different algorithms: Back-projection, CLEAN, MEM and VIS\_FWDFIT . Reconstructed images are further corrected for STIX off-pointing and spacecraft rotations with the auxiliary data. As an example, the first panel of Fig.  [[fig:imaging]](#fig:imaging) shows the light curves and time range selected for a flare occurred at about 2022-03-08T08:55:17Z . The rest of the panels show the reconstructed images. The final by the pipeline are written to files in FITS and PNG formats, while their metadata are written to the NoSQL database.

### Spectral analysis

Energy spectra provide direct information on electron acceleration in solar flares. The x-ray spectral fitting package OSPEX in SSWIDL includes a wide range of commonly used functions for parametrising the thermal and non-thermal components as well as an interface to perform the fits [[1]](#footnote-52). It reads pixel count data from L1 FITS files, corrections for dead time, transmission and energy binning for the counts, and calls OSPEX routines for spectral fitting.

![ An example of spectral fitting results. The imaging and spectroscopy pipeline performs spectral analysis for all flares in the flare list. Spectral fitting results are written to both FITS files and the NoSQL database. ](data:application/pdf;base64,)

An example of spectral fitting results. The imaging and spectroscopy pipeline performs spectral analysis for all flares in the flare list. Spectral fitting results are written to both FITS files and the NoSQL database.

# Science data request strategy

As mentioned earlier, STIX only down-links low-latency data automatically due to the data rate constraints. Pixel data with higher time and energy resolution contain much richer information. They persist in the onboard archive memory for about five to six months before they are overwritten. They can be compressed and down-linked after receiving data request telecommands initiated by the STIX operation team. A data request telecommand needs to provide information on the time range, minimal time binning, energy range, energy binning and selected detectors (or pixels), based on the knowledge from low-latency quick-look light curves. Requesting science data from STIX is a tedious task, as it must take into account many factors, such as constraints on the allocated telemetry data rate, the daily maximum number of requests that can be submitted, the onboard time binning, solar flare fluxes and the scientific value of the data. After two years of operation, the following data selection strategy has been adopted:

* Level-1 pixel data of the flaring time period is requested for detected flares with a total number signal counts greater than 1000, which is the minimal counts to reconstruct an image. In order to minimize the telemetry data size, the energy range of the request is limited to the range with signal counts based on the knowledge from light curves. If the pixel-summed peak count rate is less than 125 counts/sec (equivalent to the count rate observed for a B3 flare at 1 au), a single time-bin is requested. Otherwise, a finer time resolution of science data is requested. The actual time resolution is adjusted based on the amount of data allocated.
* Spectrograms (onboard level-4) require relative small amount data, therefore the highest possible time resolution (  0.5 seconds depending on the configuration) and energy resolution (no rebinning of science energy bins) are requested for all time periods that STIX is in NOMINAL mode.
* At least one level 1 scientific data background data is requested per day. Background period selection is done by excluding periods from the list of flares stored in the database. If a quiet solar period is not found, a relatively quiet solar period will be selected. To reduce the amount of telemetry data, background data request counts are time-integrated over the requested time period, i.e., a single time-bin.

A routine has been developed to create data requests mentioned above automatically. An unique ID is assigned to each request. Unique IDs have a format of *yymmddxxxx*, where the first eight digits indicate the two-digit year, month, day of the data start time, respectively, and the last four digits are a random number, which is unique for the day. Information of created data requests is stored in the NoSQL database. After being checked by STIX team, data requests are flagged as pending requests. Then they are selected compiled to instrument operation requests (IORs), which are used to create the final telecommands to be executed by the instrument. Apart from those automatically created data requests, data requests are also created by the instrument team for special requirements.

# STIX data platform user interfaces

## Interactive web pages

![ Data flow at STIX data center. ](data:application/pdf;base64,)

Data flow at STIX data center.

![image](data:application/pdf;base64,)

To facilitate the access of the data, a series of web APIs have been developed. The web APIs accept HTTP requests from the client side, then the requests are processed on the server-side and the results are returned to the client side. The web APIs provide access to all STIX data products and the metadata stored in the NoSQL database.

Based on the web APIs, we have built dozens of different web applications to manage and visualize different data products. The reason that we chose web techniques for the data browsers is that web applications provide many advantages, such as clear cross-platform usability, wide access with browsers, allowing rapid development and easy maintenance.

As an example, Fig. [[fig:qlbrowser]](#fig:qlbrowser) shows a screenshot of the interactive web-based STIX Quick-look data browser hosted at STIX data center website. It allows users to browse any historical STIX Quick-look data. On the browser, users can specify time ranges of Quick-look data to be loaded. After the web API on the server receives a request, it reads level-1 packets from the NoSQL database measured in the given time frame, and sends the compiled and serialized data back to the web browser. Then the data are parsed, and used to create interactive light curve plot on the client-side by using JavaScript. Thanks to the use of state-of-art web technologies, the interactive plot allows users to rebin integration times, correct light travel times, estimate GOES fluxes, and export data to local files. Apart from STIX quick-looks, quick-looks of simultaneous measurements performed by other Sun observing instruments can also be loaded on the same page, which greatly facilitates finding of the events of interest for joint analysis.

Based on similar concepts, tens of web tools have been developed for browsing other STIX data products. The four most used by STIX data users are:

* Science data browsering and interactive analyis tool: It provide users tools to search for science data, download science data, perform common data analysis tasks and visualize science data and data analysis results. On the web app, users can also select select data of interest for common analysis tasks, such as background subtraction, energy and time rebinning, and estimation of coarse flare locations, which on the client-side using JavaScript. Moreover, Users can also submit imaging and spectrocopy requests to the cloud with the interactive analysis tools on the page without the need of installing software. Interactive plots are created using the results. Users can also export data from the interactive plots or download the science data in the FITS format to the local disk for further analysis. The browser greatly reduces the barrier to explore STIX science data for new commers. It is also facilitated for experience user.
* Preview images and spectroscopy product browser: The tool provides an web-based imaging and spectroscopy manager. With the tool, users can view the the reconstructed solar flare images and fitting results from automated imaging/spectroscopy runs and those submited by register users. It also allow users to plot time evolution of emission of emission measures and temperatures as well as animation of x-ray images for the selected runs. Users can also create idl or python templates, allowing the same results to be reproduced on their local machine.
* Ancilary data viewer: The auxiliary data viewer allows users to view any historical auxilary data of Solar Orbiter. After reciving the time range information from the client-side, the server-side uses the SPICE kernels and SPICE toolkit (or STIX apsection solutions when they are available) to calculate auxiliary data commonly used in data analysis for user-selected time ranges, such as spacecraft location, velocity, light time difference, STIX pointing, STIX FOVs, and angles from differet observers, and sends the results to the client side. Then the results are charted on the web page using JavaScript.
* Housekeeping data browser: The housekeeping browser allows users to browse any historical HK data from stix. After the hk browser receives the user’s request, it sends the request to the server. The server directly uses the nosql data package to scale and serialize the data and return it to the browser. The browser then uses the data to generate interactive graphics (such as sensors). temperature, voltage, working status, memory status, etc.). which provide great convenience for load operation control and for understanding instrument status in data analysis.
* STIX data access page: STIX adopts an open data policy. STIX data products are published on the data access page once they are generated. On the page, user can search for data products by providing the data type and the observation time range , or download data products to local disks for further analysis.

## STIX data center interface: stixdcpy

*stixdcpy* is a python package that facilitates access and analysis of STIX data. It provides APIs to query and download data from STIX data center and a set of tools for visualizing data and performing common analysis tasks. With stixdcpy, users can query and download the following almost all data products available at STIX data center.

similar to the web tools, stixdcpy also provide common data analysis algorithm, such as live time correction, transmission correction, data clipping and merging, auxilary data tools, creating quick-looks for products stixdcpy is still under development The source code of stixdcpy is hosted at the github repo at <https://github.com/i4Ds/stixdcpy>.

# Summary

STIX is one of ten instruments onboard Solar Orbiter. It measures the spectrum and takes X-ray images of solar flares in the energy range 4 - 150 keV. Solar Orbiter was launched into space on February 10, 2020. During nominal operations, STIX continuously generates telemetry data. To process and archive data as well as to support the operation of instruments and scientific activities using STIX data, automated data processing pipelines and data platforms have been developed for STIX at FHNW. The pipelines generates telemetry in different levels and perform common scientific analysis. The platform provides all STIX data products of different levels and also provide users with various web-based tools to search for, browser STIX data products. It also provides web-based tools to perform common analysis tasks with STIX data. The data center is designed to work in a fully automatic mode with minimal human intervention. The concept has proven successful and has been running continuously for over two years. The platform not only facilitates the operations of the instrument, but also provide great support to STIX data users.

1. <https://hesperia.gsfc.nasa.gov/ssw/packages/spex/doc/> [↑](#footnote-ref-52)