Virtual Observatory and High Energy Astrophysics

Version 1.0

IVOA Note 2024-11-12

Working Group
DM
This version
https://www.ivoa.net/documents/Notes/VOHE-Note/20241112
Latest version
https://www.ivoa.net/documents/Notes/VOHE-Note
Previous versions
This is the first public release
Author(s)
Mathieu Servillat (Obs Paris)
Catherine Boisson (Obs Paris)
François Bonnarel (CDS)
Mark Cresitello-Dittmar (CfA)
Pierre Cristofari (Obs Paris)
Ian Evans (CfA)
Janet Evans (CfA)
Matthias Fuessling (CTAO)
Tess Jaffe (HEASARC)
Bruno Khélifi (APC)
Karl Kosack (CEA)
Mireille Louys (CDS)
Laurent Michel (Obs Strasbourg)
Ada Nebot (CDS)
Jutta Schnabel (FAU)
Fabian Schussler (CEA)
and the HE discussion group at IVOA
Editor(s)
Mathieu Servillat

Version Control

Revision e4af419, 2024-11-13 22:14:32 +0100

Abstract

This note explores the connections between the Virtual Observatory (VO) and High Energy (HE) astrophysics. Observations of the Universe at high energies are based on techniques that are different compared to the optical, or radio domain. We describe several HE observatories, then detail the specificities of the HE data and its processing, and derive typical HE use cases relevant for the VO. A HE group has been federated over the years and this note reports on several topics that could constitute an initial roadmap to a HE interest group within the International Virtual Observatory Alliance (IVOA).

Status of this document

This is an IVOA Note expressing suggestions from and opinions of the authors. It is intended to share best practices, possible approaches, or other perspectives on interoperability with the Virtual Observatory. It should not be referenced or otherwise interpreted as a standard specification.

A list of current IVOA Recommendations and other technical documents can be found in the IVOA document repository[1].

Contents

[1] https://www.ivoa.net/documents/

Acknowledgments

We acknowledge support from the ESCAPE project funded by the EU Horizon 2020 research and innovation program (Grant Agreement n.824064). Additional funding was provided by the INSU (Action Spécifique Observatoire Virtuel, ASOV), the Action Fédératrice CTA and the Action Pluriannuelle Incitatrice Astrophysique des processus de Hautes Énergies at the Observatoire de Paris, and the Paris Astronomical Data Centre (PADC).

Conformance-related definitions

The words "MUST", "SHALL", "SHOULD", "MAY", "RECOMMENDED", and "OPTIONAL" (in upper or lower case) used in this document are to be interpreted as described in IETF standard RFC2119 (Bradner, 1997).

The *Virtual Observatory (VO)* is a general term for a collection of federated resources that can be used to conduct astronomical research, education, and outreach. The International Virtual Observatory Alliance (IVOA) is a global collaboration of separately funded projects to develop standards and infrastructure that enable VO applications.

1 Introduction

HE astronomy typically includes X-ray astronomy, gamma-ray astronomy, neutrino astronomy, and studies of cosmic rays. This domain is now sufficiently developed to provide advanced data products such as catalogs, images, including full-sky surveys for some missions, and sources properties in the shape of spectra and time series. Some HE observations have been included in the VO, via data access endpoints provided by observatories or by agencies and indexed in the VO Registry.

However, after browsing those data products, users may want to reapply data reduction steps relevant to their Science objectives. A common scenario is to download HE "event" lists, i.e. list of detected events on a HE detector that are expected to be the detection of particles (e.g. a HE photon, or a neutrino), and the corresponding calibration files, including instrument response function (IRF)s. The findability and accessibility of this data via the VO is the focus of this note. We note that some existing IVOA recommendations are of interest to the domain. These should be further explored and tested by HE observatories.

We first identify and expose the specificities of HE data as provided by several HE observatories. We report typical use cases for data access and analysis of data from current HE observatories. We then intend to illustrate how HE data is or can be handled using current IVOA standards. We also discuss how IVOA standards could evolve to better integrate specific aspects of HE data, and if new standards should be developed.

1.1 Objectives of the document

The main objective of the document is to analyse how HE data can be better integrated to the VO.

A related objective is to provide a context and a list of topics to be further discussed within the IVOA by a dedicated HE Interest Group (HEIG).

1.2 Scope of the document

This document mainly focuses on HE data discovery through the VO, with the identification of common use cases in the HE astrophysics domain, which provides an insight of the specific metadata to be exposed through the VO for HE data.

Some of the current existing IVOA recommendations are discussed in this document within the HE context and will be in-depth studied in the HEIG.

2 High Energy observatories and experiments

There are various observatories, either ground, space or deep-sea based, that distribute HE data with different levels of involvement in the VO. We list here the observatories currently represented in the VO HE group. There are also other observatories that are connected to the VO in some way, and may join the group discussions at IVOA.

2.1 Gamma-ray programs

2.1.1 H.E.S.S

The High Energy Stereoscopic System (H.E.S.S.) experiment is an array of imaging atmospheric Cherenkov telescopes (IACT) located in Namibia that investigates cosmic Very High Energy (VHE) gamma rays in the energy range from 10s of GeV to 100 of TeV. It is comprised of four telescopes officially inaugurated in 2004, and a much larger fifth telescope operational since 2012, extending the energy coverage towards lower energies and further improving sensitivity.

The H.E.S.S. collaboration operates the telescopes as a private experiment and publishes mainly high level data, i.e. images, time series and spectra in scientific publications after dedicated analyses. Using complex algorithms, private software process the raw data by applying calibration, reconstructing event properties from their Cherenkov images and cleaning the event list by removing as much as possible events induced by atmospheric cosmic rays. Even after cleaning, events are largely generated by cosmic rays and statistical analyses are required to derive the astrophysical source properties.

Models of background due to the remaining cosmic rays (generally generated from real observations) are used with the gamma-ray IRFs, i.e. point spread function (PSF), Energy Dispersion and Collection Area, that are generated by extensive Monte Carlo simulations. These 4 IRFs (background, PSF, Energy Dispersion and Collection Area) are computed for each observation of $\sim$ 30min and are valid for the field of view. They depend on true energies, positions in the field of view and sometimes from event classification types. The derivation of astrophysical quantities from the event lists are now using open libraries, in particular the reference library Gammapy (Donath and Terrier et al., 2023).

In September 2018, the H.E.S.S. collaboration has, for the first time and unique time, released a small subset of its archival data using the GADF format (see 3.3.2) serialised into the Flexible Image Transport System (FITS) format, an open file format widely used in astronomy. The release consists of Cherenkov event-lists and IRFs for observations of various well-known gamma-ray sources (H.E.S.S. Collaboration, 2018).

This test data collection has been registered in the VO via a table access protocol (TAP) service hosted at the Observatoire de Paris, with a tentative ObsCore description of each dataset (see section 5.1.1). In the future, the H.E.S.S. legacy archive will possibly be published in a similar way and accessible through the VO.

2.1.2 CTAO

The Cherenkov Telescope Array Observatory (CTAO) is the next generation ground-based IACT instrument for gamma-ray astronomy at very high energies. With tens of telescopes located in the northern (La Palma, Canary Island) and southern (Chili) hemispheres, CTAO will be the first open ground-based VHE gamma-ray observatory and the world's largest and most sensitive instrument to study HE phenomena in the Universe. Built on the technology of current generation ground-based gamma-ray detectors - e.g. H.E.S.S., Major Atmospheric Gamma-ray Imaging Cherenkov (MAGIC) and Very Energetic Radiation Imaging Telescope Array System (VERITAS) - CTAO will be between five and 10 times more sensitive and have unprecedented accuracy in its detection of VHE gamma rays.

CTAO will distribute data as an open observatory, for the first time in this domain, with calls for proposals and publicly released data after a proprietary period. CTAO will ensure that the provided data will be FAIR: Findable, Accessible, Interoperable and Reusable, by following the FAIR Principles for data management (Wilkinson and Dumontier et al., 2016). In particular, because of the complex data processing and reconstruction steps, the provision of provenance metadata for CTAO data has been a driver for the development of a provenance standard in astronomy.

CTAO will also ensure VO compatibility of the distributed data and access systems. CTAO participated to the ESCAPE European Project, and is now part of the ESCAPE Open Collaboration to face common challenges for Research Infrastructures in the context of cloud computing, including data analysis and distribution.

A focus of CTAO is to distribute in this context the event list datasets, that correspond to lists of Cherenkov events detected by the telescopes along with the associated IRFs. CTAO is planning an internal and a public Science Data Challenges, which represent opportunities to build "VO inside" solutions.

The CTAO is complementary to other gamma-ray instruments observing the sky up to ultra high energies (ie PeV). Detecting directly from ground secondary charged particles of extensive air showers initiated by gamma rays, Water Cherenkov Detector (WCD) survey the whole observable sky above the TeV/tens of TeV energy range. The High Altitude Water Cherenkov Experiment (HAWC) and Large High Altitude Air Shower Observatory (LHAASO) detectors are running in the northern hemisphere and the future SWGO observatory will be installed in the southern hemisphere. Such instruments have similar high-level data structures and it has been already demonstrated that joined analyses with Gammapy of data from IACTs and WCDs using the GADF format are very powerful (Albert and Alfaro et al., 2022).

2.2 X-ray programs

2.2.1 Chandra

Part of National Aeronautics and Space Administration (NASA)'s fleet of "Great Observatories", the Chandra X-ray Observatory (CXO) was launched in 1999 to observe the soft X-ray universe in the 0.1 to 10 keV energy band. Chandra is a guest observer, pointed-observation mission and obtains roughly 800 observations per year using the Advanced CCD Imaging Spectrometer (ACIS) and High Resolution Camera (HRC) instruments. Chandra provides high angular resolution with a sub-arcsecond on-axis PSF, a field of view up to several hundred square arcminutes, and a low instrumental background. The Chandra PSF varies with X-ray energy and significantly with off-axis angle, increasing to R50 ∼25 arcsec at the edge of the field of view. A pair of transmission gratings can be inserted into the X-ray beam to provide dispersed spectra with E/DeltaE ∼1000 for bright sources. The Chandra spacecraft normally dithers in a Lissajous pattern on the sky while taking data, and this motion must be removed from the time-resolved X-ray event lists when constructing X-ray images using the motion of optical guide stars tracked by the Aspect camera.

The Chandra X-ray Center (CXC) processes the spacecraft data through a set of Standard Data Processing Level 0 through Level 2 pipelines. These pipelines perform numerous steps including decommutating the telemetry data, applying instrument calibrations (e.g., detector geometric, time- dependent gain, and CCD charge transfer efficiency (CTI) corrections, bad and hot pixel flagging), computing and applying the time-resolved Aspect solution to de-dither the motion of the telescope, identifying good time interval (GTI)s, and finally filtering out bad times and X-ray events with bad status. All data products are archived in the Chandra Data Archive (CDA) in FITS format following OGIP standards; see also § 3.3.1. The CDA manages the proprietary data period (currently 6 months, after which the data become public) and provides dedicated interactive and IVOA-compliant interfaces to locate and download datasets.

The CXC also provides the Chandra Source Catalog, which in the latest release (2.1) includes data for ∼407K unique X-ray sources on the sky and more than 2.1 million individual detections and photometric upper limits. For each X-ray source and detection, the catalog provides a detailed set of more than 100 tabulated positional, spatial, photometric, spectral, and temporal properties. An extensive selection of individual observation, stacked-observation, detection region, and master source FITS data products are also provided that are directly usable for further detailed scientific analysis.

Finally, the CXC distributes the CIAO data analysis package to allow users to recalibrate and analyse their data. A key aspect of CIAO is to provide users the ability to create instrument responses for their observations, i.e. redistribution matrix file (RMF)s, auxiliary response file (ARF)s, PSFs, etc. The Sherpa modeling and fitting package supports N-dimensional model fitting and optimisation in Python, and supports advanced Bayesian Markov chain Monte Carlo analyses.

2.2.2 XMM-Newton

The European Space Agency (ESA)'s X-ray Multi-Mirror Mission (XMM-Newton)[2] was launched by an Ariane 504 on December 10th 1999. XMM-Newton is ESA's second cornerstone of the Horizon 2000 Science Programme. It carries 3 high throughput X-ray telescopes with an unprecedented effective area, 2 reflexion grating spectrometers and an optical monitor. The large collecting area and ability to make long uninterrupted exposures provide highly sensitive observations. The XMM-Newton mission is helping scientists to solve a number of cosmic mysteries, ranging from the enigmatic black holes to the origins of the Universe itself. Observing time on XMM-Newton is being made available to the scientific community, applying for observational periods on a competitive basis.

One of the mission's ground segment modules, the Survey Science Centre (SSC)[3], is in charge of maximising the scientific return of this space observatory by exhaustively analyzing the content of the instruments' fields of view. During the development phase (1996-1999), the SSC, in collaboration with the Science Operations Centre (SOC) at European Space Astronomy Centre (ESAC), designed and produced the scientific analysis software (SAS). Since then, it has contributed to its maintenance and development. This software is publicly available.

The general pipeline is operated as ESAC since 2012, except for the part concerning cross-correlation with astronomical archives which runs in Strasbourg. The information thus produced is intended for the guest observer and, after a proprietary period of one year, for the international community. In parallel, the SSC regularly compiles an exhaustive catalog of all X-ray sources detected by European Photon Imaging Camera (EPIC) cameras.

[2] https://www.cosmos.esa.int/web/xmm-newton
[3] http://xmmssc.irap.omp.eu/

The SSC validates these catalogs, enriches them with multi-wavelength data and exploits them in several scientific programs.

The XMM-Newton catalog is published through various web applications: XSA[4], XCatDB[5], IRAP[6] and HEASARC[7]. It is also published in the VO, mainly as TAP services. It is to be noted that the TAP service operated in Strasbourg (https://xcatdb.unistra.fr/xtapdb - to be deployed in 10/2024) returns responses where data is mapped on the MANGO model with MIVOT (see section 5.1)

2.2.3 SVOM

Space-based multi-band astronomical Variable Objects Monitor (SVOM)[8] is a Sino-French mission dedicated to the study of the transient HE sky, and in particular to the detection, localisation and study of Gamma Ray Bursts (GRBs). The special feature of the SVOM mission is that it combines ground-based and space-based observations, providing a spectral bandwidth from the visible to the HE range.

The SVOM spacecraft carries four multi-wavelength instruments: ECLAIRs(4-250keV), GRM (15-5000 keV), MXT (0.3 - 10 keV) and VT (optical Blue and Red broadband filters). ECLAIRs and GRM can detect gamma-ray transient sources in real-time with localisation capabilities for ECLAIRs. An autonomous slew of the platform can be requested (only by ECLAIRs) to perform X-ray and optical follow-up of the source with the smaller field of view instruments: MXT and VT. SVOM also transfers alerts data of potential GRBs detection in near real-time to the ground with a typical latency of less than 30 seconds. The most valuable information (e.g. localisation, SNR, energy range and more) are then automatically shared to the worldwide community within the form of Notices. They will be broadcasted to the worldwide community using the NASA's General Coordinates Network (GCN) system both in VOEvent and in JSON format. Public access to the dedicated Kafka streams are planned to be opened at the beginning of 2025.

All data related to GRB detections will be public and can be retrieved through the SVOM portal (not deployed at the time of writing). All these science products, in FITS format, do conform to a global data model based on JSON descriptors. Pipeline modules are able to extend the data products they deliver with a list of keywords that carry most of the Obscore quantities. This feature will facilitate their publication in ObsTAP services.

SVOM has been successfully launched on June 22 2024 from Xichang lauchpad. As early as the commissionning phase, it has detected numer-

[4] https://www.cosmos.esa.int/web/xmm-newton/xsa
[5] https://xcatdb.unistra.fr/4xmm
[6] http://xmm-catalog.irap.omp.eu/
[7] http://heasarc.gsfc.nasa.gov/db-perl/W3Browse/w3browse.pl
[8] https://www.svom.eu/en/home/

ous interesting GRBs and triggered follow-up campaigns with very different facilities such as SWIFT, Einstein Probe or even the VLT.

2.3 KM3Net and neutrino detection

The Cubic Kilometre Neutrino Telescope (KM3NeT) is an array of WCDs currently under construction in the deep Mediterranean Sea. With its two sites off the French and Italian coasts, the KM3NeT collaboration aims at single particle neutrino detection for neutrino physics with the more densely instrumented Oscillation Research with Cosmics in the Abyss (ORCA) detector in the GeV to TeV range, and VHE astrophysics with the Astroparticle Research with Cosmics in the Abyss (ARCA) detector in the TeV range and above.

Using Earth as a shield from atmospheric particle interference by searching for upgoing particle tracks in the detectors, the measurement of astrophysical neutrinos can be performed almost continuously for a wide field of view that covers the full visible sky. For these particle events, extensive Monte Carlo simulations are performed to evaluate the statistical significance towards the various theoretical assumptions for galactic or cosmic neutrino signals and extensive filtering of the events dominated by the atmospheric particle background by about $1 : 10^6$ is required.

During the construction phase, the KM3NeT collaboration develops its interfaces for open science and builds on the data gathered by its predecessor Astronomy with a Neutrino Telescope and Abyss Environmental Research (ANTARES), from which neutrino event lists have already been published on the KM3NeT VO server as TAP service. However, for full reproducibility of searches for point-like astronomical sources as well as wider scientific use of dedicated neutrino selections, information derived from simulations like background estimate, PSF and detector acceptance are required and should be linked to the actual event list and interpolation for a given observation.

With multiple detectors targeting HE neutrinos like IceCube, ANTARES, KM3NeT, Baikal and future projects, the chance to detect a significant amount of cosmic and galactic neutrinos increases, requiring an integrated approach to link event lists with instrument responses and to correctly interpret observation time and flux expectations. As observations generally encompass large continously taken data sets covering a large area of the sky for multiple years, with very low statistical expectations for actual neutrino observation, especially correctly interpreting the observation time interval and re-weighting and limiting any probabilistic measures to a dedicated study must be facilitated for proper use of neutrino data.

2.4 Gravitional wave experiments

Gravitational wave (GW) astronomy is a subfield of astronomy concerned with the detection and study of GWs emitted by astrophysical sources. GWs are generally produced by cataclysmic events such as the merger of binary black holes, the coalescence of binary neutron stars, or supernova explosions. Those cataclysmic events may also be related to emission of HE radiations.

As of 2012, the LIGO and VIRGO observatories were the most sensitive detectors. The Japanese detector KAGRA was completed in 2019; its first joint detection with LIGO and VIRGO was reported in 2021. Another European ground-based detector, the Einstein Telescope, is under development. A space-based observatory, the Laser Interferometer Space Antenna (LISA), is also being developed by the European Space Agency.

Observations of GWs may be called GW events, though they are not related to HE events that are detections of HE particles. However, GW astronomy produces alerts and regions of interest that are relevant for HE observatories to follow-up on GW detections.

3 Common practices in the High Energy community

3.1 Data specificities

3.1.1 Event-counting

Observations of the Universe at high energies are based on techniques that are different compared to the optical, or radio domain. HE observatories are generally designed to detect particles, e.g. individual photons, cosmic rays, or neutrinos, with the ability to estimate several characteristics of those particles. This technique is named **event counting**, where an event has some probability of being due to the interaction of an astronomical particle with the detectors.

The data corresponding to an **event** is first an instrumental signal, which is then calibrated and processed to estimate event characteristics such as a time of arrival, coordinates on the sky, and the energy proxy associated to the event. Several other intermediate and qualifying characteristics can be associated to a detected event.

When observing during an interval of time, the data collected is a list of the detected events, named an **event list** in the HE domain, and event-list in this document.

3.1.2 Data levels

After detection of events, data processing steps are applied to generate data products. We typically distinguish at least 3 main data levels.

1 An event-list with calibrated temporal and spatial characteristics, e.g. sky coordinates for a given epoch, event arrival time with time reference, and a proxy for particle energy.

2 Binned and/or filtered event-list suitable for preparation of science images, spectra or light-curves. For some instruments, corresponding instrument responses associated with the event-list, calculated but not yet applied (e.g, exposure maps, sensitivity maps, spectral responses).

3 Calibrated maps, or spectral energy distributions for a source, or light-curves in physical units, or adjusted source models.

Those data products may be found in catalogs, e.g. a source catalog pointing to several data products for each source (e.g. collection of high-level products), or a catalog of source models generated with an uniform analyse.

The definitions of these data levels can vary from facility to facility. For example, in the VHE Cherenkov astronomy domain (e.g. CTAO), the data levels listed above are labelled DL3[9] to DL5. For Chandra X-ray data, the first two levels correspond to L1 and L2 data products (excluding the responses), while transmission-grating data products are designated L1.5 and source catalog and associated data products are all designated L3.

3.1.3 Background signal

Observations in HE may contain a high background component, that may be due to instrument noises, or to unresolved astrophysical sources, emission from extended regions or other terrestrial sources producing particles similar to the signal. The characterisation and estimation of this background may be particularly important to then apply corrections during the analysis of a source signal.

In the VHE domain with the IACT, WCD and neutrino techniques, the main source of background is generated by cosmic-ray induced events. The case of unresolved astrophysical sources, emission from extended regions are treated as models of gamma-ray or neutrino emission.

In the X-ray domain, contributions to background can include an instrumental component, the local radiation environment (i.e. space weather) which can change dynamically, and may include the cosmological background

[9] lower level data (DL0–DL2), that are specific to the used instrumentation (IACT, WCD), are reconstructed and filtered, which constitute the events lists called DL3.

due to unresolved astrophysical sources, depending on the spatial resolution of the instrument.

3.1.4 Time intervals

Depending on the stability of the instruments and observing conditions, a HE observation can be decomposed into several intervals of time that will be further analysed.

For example, stable time interval (STI)s are defined in Cherenkov astronomy to characterise periods of time during which the instrument response is stable. In the X-ray domain, GTIs are computed to exclude time periods where data are missing or invalid, and may be used to reject periods impacted by high radiation, e.g. due to space weather. In contrast, for neutrino physics, relevant observation periods can cover up to several years due to the low statistics of the expected signal and a continuous observational coverage of the full field of view.

3.1.5 Instrument Response Functions

Though an event-list can contain calibrated physical values, the data typically still has to be corrected for the photometric, spectral, spatial, and/or temporal responses of the instruments used to yield scientifically interpretable information. The IRFs provide mappings between the physical properties of the source and the observables, and so enable estimation of the former (such as the real flux of particles arriving at the instrument, the spectral distribution of the particle flux, and the temporal variability and morphology of the source).

The instrumental responses typically vary with the true energy of the event, the arrival direction of the event into the detector. A further complication of ground-based detectors like IACTs and WCTs is that the instrumental responses also vary with:

- The horizontal coordinates of the atmosphere, i.e. the response to a photon at low elevation is different from that at zenith due to a larger air column density, and different azimuths are affected by different magnetic field strengths and directions that modify the air-shower properties.
- The atmosphere density, which can have an effect on the response that changes throughout a year, depending on the site of observation.
- The brightness of the sky (for IACTs), i.e. the response is worse when the moon is up, or when there is a strong night-sky-background level from e.g. the Milky Way or Zodiacal light.

Since these are not aligned with a sky coordinate system, field-rotation during an observation must also be taken into account. Therefore the treatment of the temporal variation of IRFs is important, and is often taken into account in analysis by averaging over some short time period, such as the duration of the observation, or intervals within.

3.1.6 Granularity of data products

The event-list dataset is generally stored as a table, with one row per candidate detection (event) and several columns for the observed and/or estimated physical parameters (e.g. arrival time, position on detector or in the sky, energy or pulse height, and additional properties such as errors or flags that are project-dependent).

The list of columns in the event-list is for example defined in the data format, such as OGIP or GADF as introduced further below (3.3). The data formats in use generally describe the event-list data together with the IRFs (Effective Area, Energy Dispersion, Point Spread Function, Background) and other relevant information, such as: Stable and/or Good Time Interval, dead time, ...

Such time intervals may be used to define the granularity of the data products, e.g. it may be practical to list all events that will be analysed with the same IRFs over a given stable time interval. In H.E.S.S., such event-list correspond to a run of 30min of data acquisition.

Where feasible, the efficient granularity for distributing HE data products seems to be the full combination of data (event-list) and associated IRFs, packed or linked together, with further calibration files, so that the package becomes self-described.

3.2 Statistical challenges

In order to produce advanced astrophysics data products such as light curves or spectra, assumptions about the noise, the source morphology and its expected energy distribution must be introduced. This is one of the main drivers for enabling a full and well described access to event-list data, as HE scientific analyses generally start at this data level.

3.2.1 Low count statistics

Low count statistics are common for sources detected in HE astrophysics observations. For detectors with low intrinsic backgrounds, limiting source detection thresholds may be in the range 3–5 counts, *i.e.*, in the Poisson regime. Even for observations with more counts, many detectors have sufficient spatial and spectral channels (and observations are typically time-resolved) so that the number of counts per spatial pixel/spectral channel/temporal bin will often be very low, and so appropriate extreme Poisson statistical methods must be used to analyze the data (*e.g.*, using the C-statistic when analyzing low-count Poisson data that may include bins with no counts). This implies that measurements may require representations that are more robust than a mean value with Gaussian distributed errors.

3.2.2 Event selection

When analyzing an event-list, optimal selection of the events according to the science analysis use case is essential. While appropriately selecting data from an observation (*e.g.*, selecting a region surrounding the target source) is a common practice, for HE observations spatial, spectral, and temporal selection is typically necessary because of the large ranges covered by these dimensional axes. Selections may be performed on the event characteristics such as time, energy, or more specific indicators (*e.g.*, patterns, shape, IRFs properties).

3.2.3 Event binning

Binning together events in any of the spatial/spectral/temporal axes is commonly used when analyzing HE astrophysics data to increase the number of counts per bin (at the expense of reduced resolution along the given axis). For example, binning spatially can increase the S/N of faint extended emission. For the spectral and temporal axes, binning to achieve a minimum number of counts per bin may be used to facilitate data modeling while still preserving the highest possible resolution in regions with more counts. After binning, this means that spectra and light curves with variable bin widths may be commonly encountered when dealing with HE datasets.

3.2.4 The "unfolding" problem

HE and VHE astrophysics experiments are using complex detection techniques from the interaction of the radiation and the matter. For X-rays, photons interact with the materials of the telescope and detector (*e.g.*, by exciting K-shell electrons). Very high energy gamma-rays or neutrinos are interacting first with the atmosphere or the Earth to create particle cascades, whose secondaries radiate Cherenkov light. These complex interactions render the relationship between the detector observables and the source's physical properties of interest very complex. Recovering the physical properties from the observables is sometimes termed "the unfolding problem."

Most of the time, the detected number of expected counts can be related to the physical source spectrum as follows:

$$M(E',\hat{p}',t) = \int_{E'} dE\, d\hat{p}\, R(E';E,\hat{p},t)A(E,\hat{p},t)P(\hat{p}';E,\hat{p},t)S(E,\hat{p},t) + B(E',\hat{p}',t) \tag{1}$$

where $M(E',\hat{p}',t)$ is the detected source counts per bin in apparent energy E', apparent location $\hat{p}'$ and arrival time t, $R(E';E,\hat{p},t)$ is the redistribution matrix that defines the probability that a photon with actual energy E, location $\hat{p}$, and arrival time t will be observed with apparent energy E', $A(E,\hat{p},t)$ is the instrumental effective area (sensitivity), $P(\hat{p}';E,\hat{p},t)$ is the photon spatial dispersion transfer function (*i.e.*, the instrumental point spread function), $S(E,\hat{p},t)$ is the physical model that describes the physical energy spectrum, spatial morphology, and temporal variability of the source, and $B(E',\hat{p}',t)$ the number of expected background[10].

Missions that follow the OGIP standards (see section 3.3.1) generally record the redistribution matrix using the RMF format and the instrumental effective area using the ARF format. For VHE experiments, R, P, A and B form the four instrument response functions (IRFs) that are described into the Gamma-ray Astronomy Data Format (GADF) format.

Low count statistics implies that the mapping from S to M is typically not invertible (*i.e.*, one cannot simply derive S given M). Methods such as forward-folding fitting (Mattox and Bertsch et al., 1996) (*i.e.*, proposing a model for S, folding the model through equation (1) to derive M and optimizing the model parameters to minimize the deviations between M and the actual observed data) are needed to estimate the physical properties of the source from the observables. A further added complexity is that the redistribution matrix and the photon spatial dispersion transfer function can not be factorised in some cases.

3.3 Data formats

3.3.1 OGIP

NASA's High Energy Astrophysics Science Archive Research Center (HEASARC) FITS Working Group was part of the Office of Guest Investigator Programs (OGIP), and created in the 1990's the multi-mission standards for the format of FITS data files in NASA HE astrophysics. Those so-called OGIP recommendations[11] include standards on keyword usage in metadata, on the storage of spatial, temporal, and spectral (energy) information, and representation of response functions, etc. These standards predate the IVOA but include such VO concepts as data models, vocabularies, provenance, as well as the corresponding FITS serialisation specification.

[10] It can originate from the intrument, atmospheric cosmic-rays, terrestrial phenomena, etc

[11] https://heasarc.gsfc.nasa.gov/docs/heasarc/ofwg/ofwg_recomm.html

The purpose of these standards was to allow all mission data archived by the HEASARC to be stored in the same data format and be readable by the same software tools. § 2.2.1 above, for example, describes the Chandra mission products, but many other projects do so as well. Because of the OGIP standards, the same software tools can be used on all of the HE mission data that follow them. There are now some thirty plus different mission datasets archived by NASA following these standards and different software tools that can analyse any of them.

As IVOA is defining data models for spectra and time series, we should be careful to include the existing OGIP standards as special cases of what are developed to be more general standards for all of astronomy. Standards about source morphology should also be introduced.

3.3.2 GADF and VODF

The GADF[12] is a community-driven initiative for the definition of a common and open high-level data format for gamma-ray instruments (Deil and Boisson et al., 2017) starting at the reconstructed event level. GADF is based partially on the OGIP standards and is specialised for VHE data. It was originally developed in 2011 for CTAO during it's prototyping phase, and was further tested on data from the H.E.S.S. telescope array. This format is now used as a standard for VHE gamma-ray data. The project was made open-source in 2016, and became the base format for the Gammapy software.

The Very-high-energy Open Data Format (VODF)[13] (Khélifi and Zanin et al., 2023) will build upon and be the successor to GADF. It is intended to address some of the short-comings of the GADF format, to provide a properly documented and consistent data model, to cover use cases of both VHE gamma-ray and neutrino astronomy, and to provide more support for validation and versioning. VODF will provide a standard set of file formats for data starting at the reconstructed event level (event list, i.e. first item in the section 3.1.2) as well as higher-level products (i.e. sky images, light curves, and spectra) and source catalogues (see section 3.1.2), as well as N-dimensional binned data cubes. With these standards, common science tools can be used to analyse data from multiple HE instruments, including facilitating the ability to do combined likelihood fits of models across a wide energy range directly from events or binned products. VODF aims to follow or be compatible with existing IVOA standards as much as possible.

3.4 Data extraction and visualisation

HE data are typically multi-dimensional (*e.g.*, 2 spatial dimensions, time, energy, possibly polarisation) and may be complex and diverse at lower levels.

[12] https://gamma-astro-data-formats.readthedocs.io/
[13] https://vodf.readthedocs.io/

Therefore one may commonly find specific tools to process the data for a given facility, *e.g.*, CIAO for Chandra, SAS for XMM-Newton, or Gammapy for gamma-ray data, with a particular focus on Cherenkov data as foreseen by CTAO.

However, many tools in a high energy astrophysics data analysis package may perform common tasks in a mission-independent way and can work well with similar data from other facilities. For example, one commonly needs to be able to filter and project the multi-dimensional data to select specific data subsets with manageable sizes and eliminate extraneous data. Some tool sets include built-in generic filtering and binning capabilities so that a general purpose region filtering and binning syntax is available to the end user. Examples include the HEASoft package[14] (enabled by the OGIP standards mentioned above), Gammapy[15], Gamma-ray Data Tools (GDT)[16], etc.

A high energy astrophysics data analysis package typically includes tools that apply or re-apply instrumental calibrations to the data, and as described above these may be observatory-specific. More general algorithms (*e.g.*, source detection) and utility tools (*e.g.*, extract an observed spectrum from a region surrounding a source) are applied to calibrated data to extract data subsets that can then be fed into modeling tools (*e.g.*, Xspec, Sherpa, or Gammapy) together with the appropriate instrumental responses (IRFs, or RMFs and ARFs) to derive physical quantities. Since instrumental responses are often designed to be compliant with widely adopted standards, the tools that apply these responses in many cases will interoperate with other datasets that use the same standards.

Most data analysis packages provide a visualisation capability for viewing images, interacting with astronomy databases, overlaying data, or interacting via SAMP to tie several application functions together *(e.g.*, TopCat, Aladin, ds9, ESASky, Firefly) to simultaneously support both analysis and visualisation of the data at hand. In addition, many packages offer a scripting interface (*e.g.*, Python, Jupyter notebooks) that enable customised job creation to perform turn-key analysis or process bulk data in batch mode.

To allow users of data to use pre-existing tools, often packages will support file I/O using several formats, for example, including FITS images and binary tables (for event files), VO formats, and several ASCII representations (*e.g.*, space, comma, or tab-separated columns).

We do note that currently high energy astrophysics data and analysis systems are not created equally and there are a number of nuances with some of the data formats and analysis threads for specific instrument and projects.

[14] https://https://heasarc.gsfc.nasa.gov/docs/software/heasoft/
[15] https://gammapy.org/
[16] https://astro-gdt.readthedocs.io/

4 Use Cases

Given the variety of HE observatories (see section 2) and the specificities of HE data (see section 3), we list in this section some use cases that are typical to the search and handling of HE data.

4.1 UC1: re-analyse event-list data for a source in a catalog

After the selection of a source of interest, or a group of sources, one may access different high level HE data products such as images, spectra and light-curves. To further study the HE data, users genrally download the corresponding event-lists and calibration files to performe a new analyse of the data, with their specific science case in mind.

Users will thus access those event-list and retrieve or regenerate the related calibration files. They will also install and run dedicated tools to reprocess this low-level data.

One of the characteristics of the HE data is that end data products depend strongly on assumptions taken when processing raw data. This preprocessing of the data conducting to event lists has a strong impact on the results and different data manipulations may be applyed depending on the science case.

The record of provenance information during data preparation is particularly important for HE/VHE data. Their optimal use requires providing users with a view of the processing that generated the data. This implies providing ancillary data, products with different calibration levels, and possibly linking together products issued by the same processing.

4.2 UC2: observation preparation

When planning for new HE/VHE observations, one needs to search for any existing event-list data already available in the targeted sky regions, and assess if this data is enough to fulfill the science goals.

For this use case, one needs first to obtain the stacked exposure maps of past observations. This quantity is energy-dependent for VHE data can be derived from pointing position and effective areas that are position- and energy- dependent associated to each observation.

4.3 UC3: transient or variable sources

HE sources can be variable at different time scales. Observations of those sources may be triggered when sources enter a particular state, or when a transient source or phenomena appears.

Use cases for the study of those sources imply the emission of alerts from one observatory to the others, with relevant content to describe the source and its variability to organise further observtions. Otherwise, archived HE data may be reprocessed to explore the past variability at different wavelengths.

4.4 UC4: Multi-wavelength and multi-messenger science

Though there are scientific results based on HE data only, the multi-wavelength and multi-messenger approach is particularly developed in the HE domain. An astrophysical source of HE radiations is indeed generally radiating energy in several domains across the electromagnetic spectrum and may be a source of other particles, in particular neutrino. It is not rare to observe a HE source in radio and to look for counterparts in the infrared, optical or UV domain and either in X-rays or VHE band. Spectroscopy and spatially-resolved spectroscopy are also widely used to identify HE sources.

The HE domain is thus confronted to different kinds of data types and data archives, which leads to interesting use cases for the development of the VO.

One use case is associated to independent analyses of the multi-wavelength and multi-messenger data for a given source. Each kind of data product has to be retrieved, and all the datasets have to be associated to realise astrophysical interpretations, requiring some level of compatibility.

The other growing use case is associated to joint statistical analyses of multi-instrument data using adapted open science analysis tools.

For both use cases, any type of data should be findable on the VO and retrievable. And the data should have a standardised open format (OGIP, GADF, VODF).

Such use cases are already common with many examples in the X-rays in the decades that missions have been contributing to the standardized HEASoft package. Other examples include small data sets shared by VHE experiments. In (Nigro and Deil et al., 2019; Albert and Alfaro et al., 2022), groups of astronomers working on the Gammapy library had successfully analysed data taken on the Crab nebula by different facilities (MAGIC, H.E.S.S., VERITAS, Fermi and HAWC). A real statistical joint analysis has been performed to derive an emitting model of the Crab pulsar wind nebula over more than five decades in energy. Such analysis types can be now retrieved in the literature. One can also find joint analyses using X-ray and VHE data (Giunti, L. and Acero, F. et al., 2022). A proof of concept of joint analysis of VHE gamma-ray and VHE neutrino, using simulated data, has been also published (Unbehaun and Mohrmann et al., 2024).

4.5 UC5: Extended source searches

Beyond the multimessenger approach towards a specific source type, an extension of this approach can be seen in the analysis of long-term and wide-angle observations for extended sky regions in the multimessenger domain. For these analyses, extensive filtering and statistical analyses of the datasets is required. This approach is especially dominant in low-countrate experiments like neutrinos, where former analyses included the mapping of neutrino emissions in the galactic plane to gamma-ray emissions (IceCube Collaboration, 2023) or search for neutrino emission from the fermi bubbles with ANTARES data (Adrián-Martínez and Albert et al., 2014).

5 IVOA standards of interest for HE astrophysics

5.1 IVOA Recommendations

5.1.1 ObsCore and TAP

Event-list datasets can be described in ObsCore using a dataproduct_type set to "event", and distributed via a TAP service. However, this is not widely used in current services, and we observe only a few services with event-list datasets declared in the VO Registry, and mainly the H.E.S.S. public data release (see 2.1.1).

As services based on the Table Access Protocol (Dowler and Rixon et al., 2019) and ObsCore are well developed within the VO, it would be a straightforward option to discover HE event-list datasets, as well as multi-wavelength and multi-messenger associated data.

Extension of ObsCore are proposed for some astronomy domains (radio, time), which is also relevant for the astronomy domain. The ObsCore description of HE datasets is further discussed in section 6.2.

5.1.2 DataLink

The DataLink specification (Bonnarel and Dowler et al., 2023) defines a {links} endpoint providing the possibility to link several access items to each row of the main response table. These links are described and stored in a second table. In the case of an ObsCore response each dataset can be linked this way (via the access_url FIELD content) to previews, documentation pages, calibration data as well as to the dataset itself. Some dynamical links to web services may also be provided. In that case the service input parameters are described with the help of a "service descriptor" feature as described in the same DataLink specification.

5.1.3 HiPS

Several HE observatories are well suited for sky survey, and the Hierarchical Progressive Survey (HiPS) standard is well suited for sky survey exploration.

We note that the Fermi facility provides a useful sky survey in the GeV domain using this standard.

5.1.4 MOCs

Cross-correlation of data with other observations is an important use case in the HE domain. Using the Multi-Order Coverage map (MOC) standard, such operations become more efficient. Distribution of MOCs associated to HE data should thus be encouraged and especially ST-MOCs (space + time coverage) that make easier the study of transient phenomena.

5.1.5 MIVOT

Model Instances in VOTables (MIVOT, Michel and Cresitello-Dittmar et al. 2023) defines a syntax to map VOTable data to any model serialised in VO-DML. The annotation operates as a bridge between the data and the model. It associates the column/param metadata from the VOTable to the data model elements (class, attributes, types, etc.) [...]. The data model elements are grouped in an independent annotation block complying with the MIVOT XML syntax. This annotation block is added as an extra resource element at the top of the VOTable result resource. The MIVOT syntax allows to describe a data structure as a hierarchy of classes. It is also able to represent relations and composition between them. It can also build up data model objects by aggregating instances from different tables of the VOTable.

In the case of HE data, this annotation pattern, used together with the MANGO model, will help to make machine-readable quantities that are currently not considered in the VO, such as the hardness ratio, the energy bands, the flags associated with measurements or extended sources.

5.1.6 Provenance

Provenance information of VHE data product is crucial information to provide, especially given the complexity of the data preparation and analysis workflow in the VHE domain. Such complexity comes from the specificities of the VHE data as exposed in sections 3.

The develoment of the IVOA Provenance Data Model (Servillat and Riebe et al., 2020) has been conducted with those use cases in mind. The Provenance Data Model proposes to structure this information as activities and entities (as in the W3C PROV recommendation), and adds the concepts of descriptions and configuration of each step, so that the complexity of provenance of VHE data can be exposed.

5.1.7 VOEvent

Source variability and observations of transient are common in the HE domain, and as such, handling of alerts is generally included in the requirements of HE observatories. Alerts are both sent and received by HE observatories. The IVOA recommendation VOEvent (Swinbank and Allan et al., 2017) is thus of interest to the HE domain. This standard has been part of the decades-long success of of the General Coordinates Network (GCN)[17], an alert system first created in the 1990's for BATSE (Barthelmy and Butterworth et al., 1995) that has been through a number of technology and standards refreshes. See also § 6.5.

5.1.8 Measurements

The Measurements model (Rots and Cresitello-Dittmar, 2022) describes measured or determined astronomical data and their associated errors. This model is highly compatible with the primary measured properties of HE data (Time, Spatial Coordinates, Energy).

However, since HE data is typically very sparse, derived properties are often expressed as probability distributions, which are not well represented by the IVOA models. This is one area where input from the HE community can help to improve the IVOA models to better represent HE data.

5.1.9 Photometry

Flux density measurements are commonly performed in the HE domain, e.g. from images with various photometry techniques. The Photometry Data Model (PhotDM, Salgado and Louys et al. 2022) could be of interest to obtains such measurements in HE as well as at other wavelength, in order to compute Spectral Energy Distribution for a given source. PhotDM is particularly developed with an attention to optical photometry, but may be adapted to HE needs.

5.1.10 Object visibility and scheduled observations

HE observatories have similar needs on the topic of observation preparation and scheduling. As suchs, standards like ObsLocTAP (Salgado and Ibarra et al., 2021) and ObjVisSAP[18] are relevant and may be of interest in the HE domain.

[17] https://gcn.nasa.gov/
[18] https://www.ivoa.net/documents/ObjVisSAP/

5.2 Data Models in working drafts

The HE domain and practices could serve as use cases for the development of data models, such as Dataset DM, Cube DM or MANGO DM.

5.2.1 Dataset

The Dataset Metadata model[19] provides a specification of high-level metadata to describe astronomical datasets and data products. One feature of this model is that it describes a Dataset as consisting of one or more associated data products. This feature is not well fleshed out in the model. The HE use cases provide examples where it may be necessary to associate multiple data products (e.g. an event-list and its associated IRFs) as a single entity to form a useful dataset.

5.2.2 Cube

The Cube model[20] describes multi-dimensional sparse data cubes and images. This submodel is specifically designed to represent event-list data and provides the framework to represent data products such as Spectra and Time Series as slices of a multi-dimensional cube. The image modeling provides the structure necessary to represent HE image products.

5.2.3 MANGO

MANGO[21] is a model that has been developed to reveal and describe complex quantities that are usually distributed in query response tables. The use cases on which MANGO is built were collected in 2019 from different scientific fields, including HE. The model focuses on the case of the epoch propagation, the state description and photometry.

Some features of MANGO are useful for the HE domain:

- Hardness ratio support
- Energy band description
- Machine-readable description of state values
- Ability to group quantities (e.g., position with detection likelihood)
- MANGO instance association (e.g., source with detections)

[19] https://www.ivoa.net/documents/DatasetDM
[20] https://www.ivoa.net/documents/CubeDM
[21] https://github.com/ivoa-std/MANGO

6 Topics for discussions in an Interest Group

6.1 Definition of a HE event in the VO

6.1.1 Current definition in the VO

The IVOA standards include the concept of event-list, for example in ObsCore v1.1 (Louys and Tody et al., 2017), where event is a dataproduct_type with the following definition:

> **event**: an event-counting (e.g. X-ray or other high energy) dataset of some sort. Typically this is instrumental data, i.e., "event data". An event dataset is often a complex object containing multiple files or other substructures. An event dataset may contain data with spatial, spectral, and time information for each measured event, although the spectral resolution (energy) is sometimes limited. Event data may be used to produce higher level data products such as images or spectra.

More recently, a new definition was proposed in the product-type vocabulary[22] (draft):

> **event-list**: a collection of observed events, such as incoming HE particles. A row in an event list is typically characterised by a spatial position, a time and an energy.

Such a definition remains vague and general, and could be more specific, including a definition for a HE event, and the event-list data type.

6.1.2 Proposed definition to be discussed

A first point to be discuss would be to converge on a proper definition of HE specific data products:

- Propose definitions for a product-type **event-list**: A collection of observed events, such as incoming HE particles, where an event is generally characterised by a spatial position, a time and a spectral value (e.g. an energy, a channel, a pulse height).
- Propose definitions for a product-type **event-bundle**: An event-bundle dataset is a complex object containing an event-list and multiple files or other substructures that are products necessary to analyse the event-list. Data in an event-bundle may thus be used to produce higher level data products such as images or spectra.

[22] https://www.ivoa.net/rdf/product-type

An ObsCore erratum could then propose to change event for event-list and event-bundle.

The precise content of an event-bundle remains to be better defined, and may vary significantly from a facility to another.

For example, Chandra primary products distributed via the Chandra Data Archive include around half a dozen different types of products necessary to analyse Chandra data (for example, L2 event-list, Aspect solution, bad pixel map, spacecraft ephemeris, V&V Report).

For VHE gamma rays and neutrinos, the DL3 event-lists should mandatory be associated to their associated IRFs files. The links between the event-list and these IRFs should be well defined in the event-bundle.

6.2 ObsCore description of an event-list

ObsCore (Louys and Tody et al., 2017) can provide a metadata profile for a data product of type event-list (event) and a qualified access to the distributed file using the Access class from ObsCore (URL, format, file size).

6.2.1 Usage of the mandatory terms in ObsCore

In the ObsCore representation, the event-list data product is described in terms of curation, coverage and access. However, several properties are simply set to NULL following the recommendation: Resolutions, Polarisation States, Observable Axis Description, Axes lengths (set to -1).

We also note that some properties are energy dependent, such as the Spatial Coverage, Spatial Extent, PSF.

Mandatory terms in ObsCore may be for example:

- dataproduct_subtype = DL3, maybe specific data format (VODF)
- calib_level = between 1 and 2
- obs_collection could contain many details : obs_type (calib, science), obs_mode (subarray configuration), pointing_mode, tracking_type, event_type, event_cuts, analysis_type. . .
- s_ra, s_dec = maybe telescope pointing coordinates
- target_name : several targets may be in the field of view
- s_fov, s_region, s_resolution, em_resolution... all those values are energy dependent, one should specifiy that the value is at a given energy, or within a range of values.
- em_min, em_max : add fields expressed in energy (e.g. eV, keV or TeV)

- t_exptime : ontime, livetime, stable time intervals... maybe a T-MOC would help
- facility_name, instrument_name : minimalist, would be e.g. CTAO and a subarray.

6.2.2 Metadata re-interpretation for the HE context

observation_id In the current definition of ObsCore, the data product collects data from one or several observations. The same happens in HE context.

access_ref, access_format The initial role of this metadata was to hold the access_url allowing data access. Depending on the packaging of the event bundle in one compact format (OGIP, GADF, tar ball, ...) or as different files available independently in various urls, a datalink pointer can be used for accessing the various parts of IRFs, background maps, etc. Then in such a case the value for access_format should be "application/x-votable+xml;content=datalink". The format itself of the data file is then given by the datalink parameter "content-type". See next section 6.4.

o_ucd For the even-list table, we can consider all measures stored in columns values have been observed . The nature of items along time, position and energy axis are identifed in Obscore with ucd as 'time', 'pos.eq.*', 'em.*' and counted as t_xel, s_xel1, s_xel2, em_xel which correspond to the number of rows/events candidates observed.

The signal observed is the result of event counting and would be PHA (Pulse height amplitude at detector level) or a number of counts for photons or particles, or a flux, etc.., depending on the data calibration level considered. ObsCore uses o_ucd to characterise the nature of the measure. Various UCDs are used for that: o_ucd=phys.count, phot.count, phot.flux, etc. there is currently no UCD defined for a raw measure like PulseHeightAmplitude, but if needed this can be requested for addition in the UCDList vocabulary. See VEP-UCD-15_pulseheight.txt proposed at 'https://voparis-gitlab.obspm.fr/vespa/ivoa-standards/semantics/vep-ucd/-/blob/master/'.

Note that these parameters vary between the dataset of calib_level of 1 (Raw) to the a more advanced data products (calib_level 2 or 3), which are filtered and rebinned from the original raw event-list.

6.2.3 Proposed additions

ev_number The event-list contains a number of rows, representing detections candidates, that have no metadata keyword yet in Obscore. We propose 'ev_number' to record this. In fact the t_xel, s_xel1 and s_xel2, em_xel elements do not apply for an event-list in raw count as it has not been binned yet.

Adding MIME-type to access_format table As seen in section 3.3 current HE experiments and observatories use their community defined data format for data dissemination. They encapsulate the event-list table together with ancillary data dedicated to calibration and observing configurations and parameters. Even if the encapsulation is not standardised between the various projects, it is useful for a client application to rely on the access_format property in order to send it to an appropriate visualising tool.

Therefore these can be included in the MIME-type table of ObsCore section 4.7. suggestion for new terms like :

- application/x-fits-ogip ...
- application/x-gadf ...
- application/x-vodf ...

energy_min, energy_max It is not user-friendly for the user to select dataset according to an energy range when the spectral axis is expressed in wavelength and meters. The units and quantities are not familiar to this community. Moreover the numerical representation of the spectral range in em_min leads to quantities with many figures and a power as -18 not easily comparable with the current usage.

t_gti The searching criteria in terms of time coverage require the list of stable/good time intervals to pick appropriate datasets. t_min, t_max is the global time span but t_gti could contain the list of GTI as a T_MOC description following the Multi-Order-Coverage (MOC) IVOA standard (Fernique and Nebot et al., 2022). This element could then be compared across data collections to make the data set selection via simple intersection or union operations in T_MOC representation. On the data provider's side, the T-MOC element can be computed from the GTI table in OGIP or GADF to produce the ObsCore t_gti field.

6.2.4 Access and Description of IRFs

Each IRF file can have an Access object from ObsCore DM to describe a link to the IRF part of the data file. This can be reflected in an extension of ObsTAP TAP_SCHEMA.

In the TAP service we could add an IRF Table, with the following columns:

- event-list datapublisher_id

- irf_type, category of response: EffectiveArea, PSF, etc.
- irf_description, one line explanation for the role of the file
- Access.url, URL to point to the IRF
- Access.format, format of IRF
- Access.size, size of IRF file

6.3 Event-list Context Data Model

The event-list concept may include, or may be surrounded by other connected concepts. Indeed, an event-list dataset alone cannot be scientifically analysed without the knowledge of some contextual data and metadata, starting with the good/stable time intervals, and the corresponding IRFs.

The aim of an Event-list Context Data Model is to name and identify the relations between the event-list and its contextual information. A first attempt is presented in Figure 1.

Such a model could help to define specific HE data attributes, that could be relevant for an ObsCore description of HE dataset, and thus incuded in a proposed extension.

6.4 Use of Datalink for HE products

There are two options to provide an access to a full event-bundle package.

In the first option, the "event-bundle" dataset (6.1) exposed in the discovery service contains all the relevant information, e.g. several frames in the FITS file, one corresponding to the event-list itself, and the others providing good/stable time intervals, or any IRF file. This is what was done in the current GADF data format (see 3.3.2). In this option, the content of the event-list package should be properly defined in its description: what information is included and where is it in the dataset structure? The Event-list Context Data Model (see 6.3) would be useful to provide that information.

In the second option, we would provide links to the relevant information from the base "event-list" (6.1) or product exposed in the discovery service. This could be done using Datalink and a list of links to each contextual information such as the IRFs. The Event-list Context Data Model (see 6.3) would provide the concepts and vocabulary to characterise the IRFs and other information relevant to the analysis of an event-list. These specific concepts and terms describing the various flavors of IRFs and GTIs will be given in the semantics and content_qualifier FIELDS of the DataLink response to qualify the links. The different links can point to different dereferencable URLs or alternatively to different fragments of the same dereferencable URL as stated by the DataLink specification.

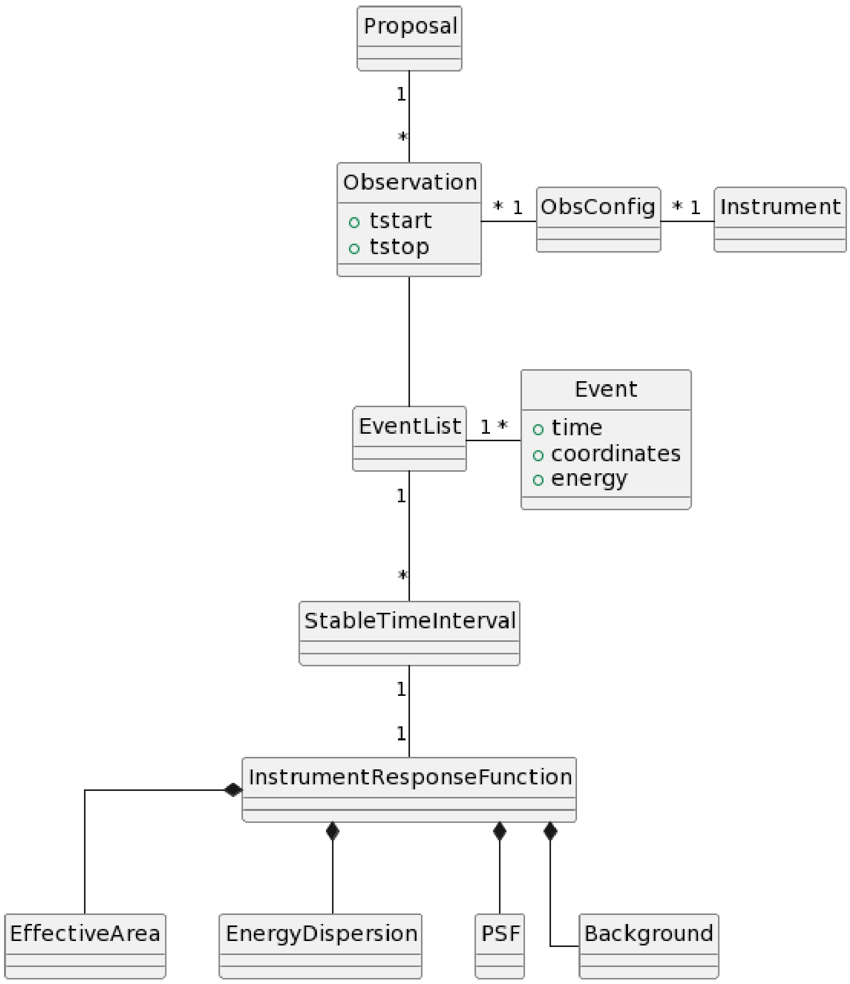

Figure 1: event-list Context Data Model. Notes: STIs and GTIs are slightly different concepts, and multiplicities should be adapted, energy is to specific for an event (intensity?), more products may be attached to a STI/GTI or to IRF.

6.5 VOEvent modernization

The modern GCN is exploring new standards more compatible with modern event brokers such as Kafka. This includes defining metadata serialization in formats such as JSON and new schema[23] to describe how to use them. This interest group should actively engage the growing community of projects publishing transient alerts through systems that have outgrown the VOEvent standard.

[23] https://gcn.nasa.gov/docs/notices/schema

Glossary

ACIS Advanced CCD Imaging Spectrometer. 8

ANTARES Astronomy with a Neutrino Telescope and Abyss Environmental Research. 11, 21

ARCA Astroparticle Research with Cosmics in the Abyss. 11

ARF auxiliary response file. 9, 17, 19

CDA Chandra Data Archive. 8

CTAO Cherenkov Telescope Array Observatory. 7, 13, 18, 27

CTI charge transfer efficiency. 8

CXC Chandra X-ray Center. 8, 9

EPIC European Photon Imaging Camera. 9

ESA European Space Agency. 9

ESAC European Space Astronomy Centre. 9

FITS Flexible Image Transport System. 6, 8, 17, 19, 31

GADF Gamma-ray Astronomy Data Format. 17, 18, 21, 28, 29, 31

GTI good time interval. 8, 14, 29–31

GW Gravitational wave. 11, 12

H.E.S.S. High Energy Stereoscopic System. 6, 7, 15, 18, 21, 22

HE High Energy. 2, 5–7, 10–13, 15–28, 30, 31, 37

HEASARC High Energy Astrophysics Science Archive Research Center. 17

HRC High Resolution Camera. 8

IACT imaging atmospheric Cherenkov telescopes. 6–8, 13, 14

IRF instrument response function. 5–7, 14–16, 19, 24, 26, 28–31

IVOA International Virtual Observatory Alliance. 2, 5, 6, 8, 17, 18, 23–25, 29, 37

KM3NeT Cubic Kilometre Neutrino Telescope. 11

NASA National Aeronautics and Space Administration. 8, 17, 18

OGIP Office of Guest Investigator Programs. 17, 18, 21, 28, 29

ORCA Oscillation Research with Cosmics in the Abyss. 11

PSF point spread function. 6, 8, 9, 11, 27, 29

RMF redistribution matrix file. 9, 17, 19

SAS scientific analysis software. 9, 18

SOC Science Operations Centre. 9

SSC Survey Science Centre. 9

STI stable time interval. 14, 30

SVOM Space-based multi-band astronomical Variable Objects Monitor. 10

TAP table access protocol. 7, 10, 11, 22, 29

VHE Very High Energy. 6, 7, 11, 13, 16–18, 20, 21, 23, 26

VO Virtual Observatory. 2, 5–7, 10, 11, 17, 19–23

VODF Very-high-energy Open Data Format. 18, 21, 27

WCD Water Cherenkov Detector. 7, 8, 11, 13

XMM-Newton X-ray Multi-Mirror Mission. 9, 18

References

Adrián-Martínez, S., Albert, A., Al Samarai, I. et al. (2014), 'A search for neutrino emission from the fermi bubbles with the antares telescope', *The European Physical Journal C* **74**(2). doi:10.1140/epjc/s10052-013-2701-6, http://dx.doi.org/10.1140/epjc/s10052-013-2701-6.

Albert, A., Alfaro, R., Arteaga-Velázquez, J. C. et al. (2022), 'Validation of standardized data formats and tools for ground-level particle-based gamma-ray observatories', A&A **667**, A36, arXiv:2203.05937. doi:10.1051/0004-6361/202243527, https://ui.adsabs.harvard.edu/abs/2022A&A...667A..36A.

Barthelmy, S. D., Butterworth, P., Cline, T. L., Gehrels, N., Fishman, G. J., Kouveliotou, C. and Meegan, C. A. (1995), 'BACODINE, the Real-Time BATSE Gamma-Ray Burst Coordinates Distribution Network', Ap&SS **231**(1-2), 235–238. doi:10.1007/BF00658623, https://ui.adsabs.harvard.edu/abs/1995Ap&SS.231..235B.

Bonnarel, F., Dowler, P., Michel, L., Demleitner, M. and Taylor, M. (2023), 'IVOA DataLink Version 1.1', IVOA Recommendation 15 December 2023. doi:10.5479/ADS/bib/2023ivoa.spec.1215B, https://ui.adsabs.harvard.edu/abs/2023ivoa.spec.1215B.

Bradner, S. (1997), 'Key words for use in RFCs to indicate requirement levels', RFC 2119. http://www.ietf.org/rfc/rfc2119.txt.

Deil, C., Boisson, C., Kosack, K., Perkins, J., King, J., Eger, P., Mayer, M., Wood, M., Zabalza, V., Knödlseder, J., Hassan, T., Mohrmann, L., Ziegler, A., Khelifi, B., Dorner, D., Maier, G., Pedaletti, G., Rosado, J., Contreras, J. L., Lefaucheur, J., Brügge, K., Servillat, M., Terrier, R., Walter, R. and Lombardi, S. (2017), Open high-level data formats and software for gamma-ray astronomy, *in* '6th International Symposium on High Energy Gamma-Ray Astronomy', Vol. 1792 of *American Institute of Physics Conference Series*, AIP, p. 070006. doi:10.1063/1.4969003, https://ui.adsabs.harvard.edu/abs/2017AIPC.1792g0006D.

Donath, A., Terrier, R., Remy, Q., Sinha, A., Nigro, C., Pintore, F., Khélifi, B., Olivera-Nieto, L., Ruiz, J. E., Brügge, K., Linhoff, M., Contreras, J. L., Acero, F., Aguasca-Cabot, A., Berge, D., Bhattacharjee, P., Buchner, J., Boisson, C., Carreto Fidalgo, D., Chen, A., de Bony de Lavergne, M., de Miranda Cardoso, J. V., Deil, C., Füßling, M., Funk, S., Giunti, L., Hinton, J., Jouvin, L., King, J., Lefaucheur, J., Lemoine-Goumard, M., Lenain, J.-P., López-Coto, R., Mohrmann, L., Morcuende, D., Panny, S., Regeard, M., Saha, L., Siejkowski, H., Siemiginowska, A., Sip"ocz, B. M., Unbehaun, T., van Eldik, C., Vuillaume, T. and Zanin, R. (2023), 'Gammapy: A python package for gamma-ray astronomy', A&A **678**, A157. doi:10.1051/0004-6361/202346488.

Dowler, P., Rixon, G., Tody, D. and Demleitner, M. (2019), 'Table Access Protocol Version 1.1', IVOA Recommendation 27 September 2019. doi:10.5479/ADS/bib/2019ivoa.spec.0927D, https://ui.adsabs.harvard.edu/abs/2019ivoa.spec.0927D.

Fernique, P., Nebot, A., Durand, D., Baumann, M., Boch, T., Greco, G., Donaldson, T., Pineau, F.-X., Taylor, M., O'Mullane, W., Reinecke, M. and Derrière, S. (2022), 'MOC: Multi-Order Coverage map Version 2.0', IVOA Recommendation 27 July 2022. doi:10.5479/ADS/bib/2022ivoa.spec.0727F, https://ui.adsabs.harvard.edu/abs/2022ivoa.spec.0727F.

Giunti, L., Acero, F., Khélifi, B., Kosack, K., Lemière, A. and Terrier, R. (2022), 'Constraining leptonic emission scenarios for the pevatron candidate hess j1702-420 with deep xmm-newton observations', A&A **667**, A130. doi:10.1051/0004-6361/202244696.

H.E.S.S. Collaboration (2018), 'H.e.s.s. first public test data release'. doi:10.5281/ZENODO.1421098, https://zenodo.org/record/1421098.

IceCube Collaboration (2023), 'Observation of high-energy neutrinos from the galactic plane', *Science* **380**(6652), 1338–1343. doi:10.1126/science.adc9818, https://www.science.org/doi/abs/10.1126/science.adc9818.

Khélifi, B., Zanin, R., Kosack, K., Olivera-Nieto, L. and Schnabel, J. (2023), The very-high-energy open data format: towards a shared, open data format in very-high-energy astronomy. https://arxiv.org/abs/2308.13385.

Louys, M., Tody, D., Dowler, P., Durand, D., Michel, L., Bonnarel, F., Micol, A. and IVOA DataModel Working Group (2017), 'Observation Data Model Core Components, its Implementation in the Table Access Protocol Version 1.1', IVOA Recommendation 09 May 2017. doi:10.5479/ADS/bib/2017ivoa.spec.0509L, https://ui.adsabs.harvard.edu/abs/2017ivoa.spec.0509L.

Mattox, J. R., Bertsch, D. L., Chiang, J., Dingus, B. L., Digel, S. W., Esposito, J. A., Fierro, J. M., Hartman, R. C., Hunter, S. D., Kanbach, G., Kniffen, D. A., Lin, Y. C., Macomb, D. J., Mayer-Hasselwander, H. A., Michelson, P. F., von Montigny, C., Mukherjee, R., Nolan, P. L., Ramanamurthy, P. V., Schneid, E., Sreekumar, P., Thompson, D. J. and Willis, T. D. (1996), 'The Likelihood Analysis of EGRET Data', *Astrophysical Journal* **461**, 396. doi:10.1086/177068, https://ui.adsabs.harvard.edu/abs/1996ApJ...461..396M.

Michel, L., Cresitello-Dittmar, M., Bonnarel, F., Landais, G., Lemson, G., Louys, M. and Salgado, J. (2023), 'Model Instances in VOTables Version 1.0', IVOA Recommendation 20 June 2023. doi:10.5479/ADS/bib/2023ivoa.spec.0620M, https://ui.adsabs.harvard.edu/abs/2023ivoa.spec.0620M.

Nigro, C., Deil, C., Zanin, R., Hassan, T., King, J., Ruiz, J. E., Saha, L., Terrier, R., Brügge, K., Nöthe, M., Bird, R., Lin, T. T. Y., Aleksić, J., Boisson, C., Contreras, J. L., Donath, A., Jouvin, L., Kelley-Hoskins, N., Khelifi, B., Kosack, K., Rico, J. and Sinha, A. (2019), 'Towards open and reproducible multi-instrument analysis in gamma-ray astronomy', A&A **625**, A10, arXiv:1903.06621. doi:10.1051/0004-6361/201834938, https://ui.adsabs.harvard.edu/abs/2019A&A...625A..10N.

Rots, A. and Cresitello-Dittmar, M. (2022), 'Astronomical Measurements Model Version 1.0', IVOA Recommendation 04 October 2022. doi:10.5479/ADS/bib/2022ivoa.spec.1004R, https://ui.adsabs.harvard.edu/abs/2022ivoa.spec.1004R.

Salgado, J., Ibarra, A., Ehle, M., Gabriel, C., Kretschmar, P., Kuulkers, E., Merín, B., Ness, J.-U., Salazar, E., Saxton, R., Cecconi, B., Foster, K., Demleitner, M., Dempsey, J., Molinaro, M., Sánchez, C., Taylor, M., Díaz Trigo, M., Fernández, M., Kennea, J., Kettenis, M., Matt, G., Osborne, J., de Oña Wilhelmi, E., Salbol, E. J., Sivakoff, G., Tao, L., Tohuvavohu, A., Tibbetts, M. and Workman, B. (2021), 'Observation Locator Table Access Protocol Version 1.0', IVOA Recommendation 24 July 2021. doi:10.5479/ADS/bib/2021ivoa.spec.0724S, https://ui.adsabs.harvard.edu/abs/2021ivoa.spec.0724S.

Salgado, J., Louys, M., Michel, L., Rodrigo, C., Osuna, P., Allen, M., McDowell, J., Baines, D., Maíz Apellániz, J., Hatziminaoglou, E., Derriere, S. and Lemson, G. (2022), 'IVOA Photometry Data Model Version 1.1', IVOA Recommendation 01 November 2022. doi:10.5479/ADS/bib/2022ivoa.spec.1101S, https://ui.adsabs.harvard.edu/abs/2022ivoa.spec.1101S.

Servillat, M., Riebe, K., Boisson, C., Bonnarel, F., Galkin, A., Louys, M., Nullmeier, M., Renault-Tinacci, N., Sanguillon, M. and Streicher, O. (2020), 'IVOA Provenance Data Model Version 1.0', IVOA Recommendation 11 April 2020. doi:10.5479/ADS/bib/2020ivoa.spec.0411S, https://ui.adsabs.harvard.edu/abs/2020ivoa.spec.0411S.

Swinbank, J. D., Allan, A. and Denny, R. B. (2017), 'VOEvent Transport Protocol Version 2.0', IVOA Recommendation 20 March 2017. doi:10.5479/ADS/bib/2017ivoa.spec.0320S, https://ui.adsabs.harvard.edu/abs/2017ivoa.spec.0320S.

Unbehaun, T., Mohrmann, L., Funk, S. et al. (2024), 'Prospects for combined analyses of hadronic emission from gamma-ray sources in the milky way with cta and km3net', *The European Physical Journal C* **84**(2). doi:10.1140/epjc/s10052-023-12279-z, http://dx.doi.org/10.1140/epjc/s10052-023-12279-z.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B. (2016), 'The fair guiding principles for scientific data management and stewardship', *Scientific Data* **3**(1), 160018. doi:10.1038/sdata.2016.18.

A Changes from Previous Versions

No previous versions yet.

B Contributions to the Note

The authors of this Note contributed to write and structure the text. However, the note has been intitiated and elaborated in several dedicated workshops and in specific IVOA HE group meetings, involving more people. The IVOA HE group keeps track of its activities on an IVOA web page: https://wiki.ivoa.net/twiki/bin/view/IVOA/HEGroup.

Further material can be found with those links:

- 2024-05-21: IVOA Sydney meeting, DM Session High Energy focus, https://wiki.ivoa.net/twiki/bin/view/IVOA/InterOpMay2024DM
- 2023-06-28: IVOA standards for High Energy Astrophysics (French VO Workshop), https://indico.obspm.fr/event/1963/
- 2023-05-11: IVOA Bologna meeting: presentation ("DM for High Energy astrophysics", M. Servillat) and first IVOA HE group meeting, https://wiki.ivoa.net/internal/IVOA/IntropMay3023DM/2023-05-11_IVOA_meeting_-_VOHE.pdf
- 2022-10-11: Virtual Observatory and High Energy Astrophysics (French VO Workshop), https://indico.obspm.fr/event/1489/