International
Virtual
Observatory
Alliance

IVOA Obscore Extension for HE data
Version 1.0

Endorsed Note ???

Working Group
High Energy Interest Group
This version
https://www.ivoa.net/documents/HighEnergyObsCoreExt/???
Latest version
https://www.ivoa.net/documents/HighEnergyObsCoreExt
Previous versions
This is the first public release
Author(s)
????Fred Offline????
Editor(s)
????Alfred Usher Thor????
Version Control
Revision 8ebe0dc-dirty, 2025-03-06 18:52:56 +0100

Abstract

This is a proposed extension to the ObsCore specification for data description, discovery and selection of High Energy data.

Status of this document

This document is an IVOA Endorsed Note. It has been reviewed and endorsed by the IVOA Technical Coordination Group as a stable, citable document which constitutes valuable information for the IVOA community and beyond.

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

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

Contents

Acknowledgments

???? Or remove the section header ????

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 (**?**).

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

The High Energy Interest Group was formed in the IVOA in Fall of 2024. As part of that process for acceptance, the group wrote an International Virtual Observatory Alliance (IVOA) note that explores the connections between the Virtual Observatory (Virtual Observatory (VO)) and High Energy (High Energy (HE)) astrophysics (Servillat, M., et al, 2024). The High Energy Interest Group (HEIG) Note includes an outline of several important topics that has formed a roadmap for the group. An ObsCore extension for High Energy data is the first priority in order to meet a needs for HE, and to

coincide with work being carried out by the Radio IG, Time Domain IG, and discussions on DM standards (e.g., Comon Archive Observation Model (CAOM)).

The goal is to define an extension to Obscore to enable a more complete description of High Energy data, and that will lead to better discovery and selection of High Energy data through IVOA interfaces. We explore what is needed for High Energy and we highlight commonality with the proposed Radio and/or Time extension. If 1) an item is unique for HE, it will appear in the HE extension, if 2) an item makes sense for more that one domain, the keyword will need to be named unique for each domain, if 3) an item can be shared, then it needs to be added to the base Obscore model. We will make recommendations in all three categories. Topics related to the Registry are currently outlined in the Radio document and not discussed here.

2 Obscore section from VOHE Note

The following section is a placeholder and has been copied from section 6.2 of the HE Note (section with Obscore extension details). Seeding it here to collect information on Obscore written by HEIG member to one spot. Expect it to be edited/replaced - whatever makes sense for this doc.

2.1 ObsCore description of an event-list

ObsCore (**?**) 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).

2.1.1 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, point spread function (PSF).

Terms in ObsCore may be filled in the following way for example, considering a Cherenkov Telescope Array Observatory (CTAO) DL3 dataset:

- dataproduct_type = event
- dataproduct_subtype = DL3, maybe specific data format (e.g. Very-high-energy Open 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.

2.1.2 Metadata re-interpretation for the HE context

obs_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 (Office of Guest Investigator Programs (OGIP), Gamma-ray Astronomy Data Format (GADF), tar ball, ...) or as different files available independently in various urls, a datalink pointer may be used for accessing the various parts of instrument response function (IRF)s, 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 **??**.

o_ucd For the even-list table, we can consider that all measures stored in column 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[2].

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.

2.1.3 Possible additions

ev_number The event-list contains a number of rows, representing detections candidates, that have no metadata keyword yet in Obscore. We could 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 **??** 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.

[2]See VEP-UCD-15_pulseheight.txt proposed at 'https://voparis-gitlab.obspm.fr/vespa/ivoa-standards/semantics/vep-ucd/-/blob/master/'

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 good time interval (GTI) as a T_MOC description following the Multi-Order-Coverage (MOC) IVOA standard (**?**). 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.

2.1.4 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 table access protocol (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

3 Obscore table from Ian's Malta Obscore presentation

The following section is a placeholder and has been copied from Ian Evans Malta Interop presentation. We are seeding it here to collect information on Obscore written by HEIG members to one spot. Expect it to be edited/replaced - whatever makes sense for this doc. See Table 1 below.

Table 1: Elements Already Present in Core ObsCore

Element	**HEA-specific Issue?**	**Description**	**Suggested Recommendation**
o_ucd/o_unit/ o_calib_status/ o_stat_error	Yes?	Most HEA event lists have multiple observables; these quantities should support multiple observables; setting o_ucd = NULL for event lists as recommended is obscuring and does not facilitate data discovery	Modify these elements to support meaningful representation of DPs that include multiple observables
s_fov/ s_resolution/ em_resolution	Maybe	These properties may be (strongly) dependent on particle energy, location within the FoV, ... (*e.g.*, *Chandra* spatial resolution varies by a factor 50× across the FoV)	Consider how to represent properties that are strongly dependent on other quantities and how to enable meaningful data discovery for these cases
calib_level	No?	Many HEA event lists fall between calib_levels 1 and 2 (spatial and temporal axes are calibrated physical quantities, but spectral axis is instrumental and requires application of responses); this is not always the case so it is beneficial to be able to differentiate "1.5" and 2	Consider how to capture the calibration level of DPs that have the instrument signature partially removed (*e.g.*, only for some axes)
dataproduct_type	No	Limited set of dataproduct_types don?t represent some common types of HEA or Advanced DPs; limitation on use of "measurements" means that value can?t be used for most ADPs; lack of a broad set of dataproduct_types limits ability to perform meaningful data discovery	Add a wider set of dataproduct_types in consultation with the multi-waveband community; remove caveat on use of measurements
t_min/t_max	No	Not useful for ADPs that combine multiple observations; see also t_gti in Table 2	Modify these elements to support multiple disjoint time intervals
proposal_id	No	Single valued proposal_id may not work for ADPs that combine multiple observations	Modify proposal_id to allow multiple values similar to other provenance properties (facility_name, instrument_name)

Table 2: Additional Elements

Element	**HEA-specific Issue?**	**Description**	**Suggested Recommendation**
ev_number	Yes	Number of events in an event list is a useful HEA dimensionality for data discovery	Add as HEA extension
energy_min/ energy_max	Yes	em_min/em_max in units of m do not work well for HEA, where the natural units are energy (*i.e.*, inverse wavelength); there are additional usability concerns for VHEA that may make em_min/em_max unusable	Consider adding energy_min/ energy_max as HEA extension
t_gti	Yes?	t_min/t_max do not allow representation of multiple GTIs/STIs so queries on time may not be accurate	Consider solving as part of support for multi-valued t_min/t_max rather than adding a separate HEA-specific concept
irf_type/ irf_description etc.	No	HEA event list data products typically require associated DPs such as instrument response functions (*e.g.*, IRF, RMF, ARF) for analysis; HOWEVER, identification of associated DPs required to enable meaningful further analysis NOT a HEA-specific issue	May be soluble using datalink; alternatively consider adding (multi-valued) assocproduct_type, assocproduct_description etc. to core
access_format	No	Additional access_format MIME-types may be needed to support standardized HEA formats	Identify and request addition of appropriate MIME-types
UCDs	No	Additional UCDs may be needed to support some HEA observables; (*e.g.*, no UCD is defined for PHA)	Identify and request addition of appropriate UCDs

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 was initiated and elaborated in several dedicated workshops, Interop meetings, 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:

- 2024-11-16: IVOA Malta meeting, DM session with 2 High Energy presentations (B. Khelifi/I. Evans), https://wiki.ivoa.net/twiki/bin/view/IVOA/InterOpNov2024DM
- 2024-11-15: IVOA Malta Plenary, CSP Plenary session, https://wiki.ivoa.net/twiki/bin/view/IVOA/InterOpNov2024CSPPlenary
- 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/