International
Virtual
Observatory
Alliance

IVOA ObsCore Extension for High Energy Astrophysics Data

Version 1.0

Proposed Endorsed Note 2025-04-24

Working Group
High Energy Interest Group
This version
https://www.ivoa.net/documents/HighEnergyObsCoreExt/20250424
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????

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 is an IVOA Proposed Endorsed Note for review by IVOA members and other interested parties. It is appropriate to reference this document only as a Proposed Endorsed Note that is under review and may change before it is endorsed or may not be endorsed.

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

DRAFT – please do not distribute

Acknowledgments

The authors would like to thank all the participants in the IVOA High Energy Interest Group and Data Model Working Group for their ideas, discussions, critical reviews, and contributions to this document.

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

The International Virtual Observatory Alliance (IVOA) High Energy Interest Group (HEIG) was formed in the Fall of 2024, and developed an IVOA Note (Servillat, M., et al, 2024) that explores the connections between the Virtual Observatory (VO) and High Energy Astrophysics (HEA). The HEIG Note includes an outline of several important topics that has formed a roadmap for the group. An ObsCore (Louys and Tody et al., 2017) extension for High Energy (HE) data is the first priority in order to meet the needs of HEA, and to coincide with similar work being carried out by the Radio IG, Time Domain IG, and discussions on DM standards.

The goal is to define an extension to ObsCore to enable a more complete description of HE data, and that will lead to better discovery and selection of HE data through IVOA interfaces. We explore elements needed to reliably discover HE data and highlight commonalities with proposed Radio and/or Time extensions. We suggest that if an attribute is unique to HE data then that element should appear in an HE ObsCore extension, whereas if an attribute makes sense for more that one domain and can be shared across those domains then that element should be added to the base ObsCore model. We will make recommendations in both of these categories. We also discuss enhancements to the data product vocabulary, UCDs, and MIME-types that are helpful (or in some cases necessary) to correctly represent HE data where these are related to ObsCore definitions or usage. Topics related to the Registry are currently outlined in the proposed Radio extension document and are not discussed here.

2 High Energy Astrophysics Data

HEA data are typically considered to include observations obtained using X-ray, gamma-ray, and cosmic ray detectors, as well as neutrino astronomy studies. The domain is now sufficiently developed to provide images, including full-sky surveys for some missions, spectra, and time series (light curves), as well as source catalogs with advanced, high-level data products such as aperture photometry probability densities, model spectral fits, extended source contours, and sensitivity maps.

Observations of the Universe at high energies are based on techniques that are radically different compared to the UV through radio domains. HE observatories are generally designed to detect particles (*e.g.*, individual photons, cosmic-rays, or neutrinos) with the ability to estimate multiple observables for those particles. This technique is generally 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 are first an instrumental signal, which is then calibrated and processed to estimate physical quantities such as a time of arrival, coordinates on the sky, and an energy proxy associated to the event. Several other intermediate and qualifying characteristics can be associated to a detected event. The list of events detected over a given time interval is termed an event list in the HE domain (**event-list** in this document). Though an **event-list** *may* include calibrated physical values, the data typically still have to be corrected for the photometric, spectral, spatial, and/or temporal responses of the telescope and detector combination to yield scientifically interpretable information. The mappings between the observables and physical measurements of the source properties are termed Instrument Response Functions (IRFs) that enable estimation of the latter (*e.g.*, 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) from the former. The IRFs are typically probabilistic in nature, and may depend on the set of events selected for analysis by the end user. They are usually not invertible, so methods such as forward fitting (model optimization) are needed to estimate physical properties given observables. Some missions split IRFs into multiple separate data products.

3 ObsCore Attribute Definitions for High Energy Astrophysics Data

The ObsCore representation of HEA **event-list** data products is described in terms of curation, coverage, and access. However, following the ObsCore Recommendation, several properties, including resolutions, observable axis descriptions, and polarization states are simply set to NULL, and data axis lengths are set to -1. Therefore, for these data products and associated IRFs, the definitions of some ObsCore attributes should be adjusted so that they better represent the content of the data from the perspective of data discovery. These adjustments will also typically apply to advanced, high-level data products derived from **event-list** data.

We note that many properties, including spatial and spectral coverage and resolution can vary strongly with energy and off-axis angle.

3.1 *dataproduct_type*

The attribute *dataproduct_type* provides a high level scientific classification of the data product and is of primary importance for data discovery, especially when there may be many different types of data product associated with an observation (as is often the case for HEA datasets). The current ObsCore Recommendation defines an **event** *dataproduct_type* as:

> **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.

We propose to add the following *dataproduct_type*s to better define a HE **event-list** and a bundle that includes the **event-list** and associated ancillary data:

> **event-list**: A collection of observed events, such as incoming HE particles, where an event is typically characterized by a spatial position, a time, and a spectral value (*e.g.*, an energy, a channel, a pulse height).
>
> **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 analyze the **event-list**. Data in an event-bundle may thus be used to produce higher level data products such as images or spectra.

An **event-bundle** might for example consist of an **event-list** and the associated IRFs used to calibrate the dataset; alternatively an **event-bundle** may include the **event-list** and associated ancillary data products necessary for the user to *create* the IRFs (for those cases where detailed knowledge of the scientific use case — for example, the user's selection of events — may be required to compute the responses).

In addition to *dataproduct_type*s that focus on event data, we note that existing ObsCore definitions do not adequately span the breadth of advanced data products (with *calib_level* ≥ 3 that may be generated from astronomical observations. The computational complexity of analyzing HEA data robustly in the extreme Poisson regime (*e.g.*, Bayesian X-ray aperture photometry applied simultaneously to multiple overlapping detections and observations) means that data providers may choose to provide such analysis products directly to the end user. For example, the *Chandra* Source Catalog includes 38 types of advanced data products (for a total of $\sim$90 million files) and $\sim 50\%$ of these data product types are not well represented by a *dataproduct_type* value that allows for meaningful data discovery. Users will certainly want to discover these data products independently from the associated observation data (and many of these data products combine data from multiple observations). We therefore propose the following additional *dataproduct_type*s for these advanced data products, and note that these *dataproduct_type*s will certainly be useful independent of waveband (*i.e.*, they can be equally applicable to UV/optical, IR, and radio datasets:

> **draws**: A dataset that represents draws computed from a probability distribution, for example the Markov chain Monte Carlo (MCMC) draws used when computing the Bayesian marginal probability density function for a random variable. The draws can be interpreted to provide a robust estimation of the probability distribution of variable, and correlations between the draws provide information about how well the draws converge to the parent probability distribution.
>
> **pdf**: A dataset that represents the probability density function of a quantity, for example the Bayesian marginal probability density function for a random variable. The probability density function provides a robust estimation of the variable and allows arbitrary confidence intervals to be computed directly from the distribution.
>
> **region**: A dataset that includes an encoding of (one or more) regions of parameter space, for example a spatial region or a region of phase space covered by a dataset. The set of dimensions represented by the region can be arbitrary.
>
> **response-function**: A dataset that represents a mapping between a physical quantity and an observable. For HEA this may be the components of an IRF such as an Auxiliary Response File (ARF), Redistribution Matrix File (RMF), the composite IRF. More general response-functions include a Point Spread Function (PSF). While these datasets may typically be represented as an N-dimensional cube (often a 2-dimensional image) designating them as **response-function**s enhances data discovery for a very common type of HE dataset.

The **measurements** data product type is quite useful for many different types of advanced data products (that may be derived from multiple observations) but users of those products often may not be interested the progenitor datasets, especially if many advanced data products are extracted from a single or a few progenitors (*e.g.*, **measurements** associated with sources detected in a single observation field). We propose to delete the caveat associated with *dataproduct_type* = **measurements** type in the ObsCore IVOA Recommendation (§ 4.1.1) that requires the derived data products be exposed "**together** with the progenitor observation dataset".

3.2 *dataproduct_subtype*

The optional attribute *dataproduct_subtype* may be used by the data provider to specify additional information about the nature of the data product. For some datasets this attribute may specify the data format or data level (*e.g.*, DL3–6), or may be combined with *dataproduct_type* to more precisely define the content of the dataset (*e.g.*, *dataproduct_type* = **image** + *dataproduct_subtype* = **exposuremap**, or *dataproduct_type* = **response-function** + *dataproduct_subtype* = **rmf**), which is particularly useful and important for discovery of advanced data products. A vocabulary of such data product (sub-)types should be developed to support discovery of advanced data products.

3.3 *calib_level*

ObsCore defines calibration **Level 1** as "Instrumental data in a standard format (FITS, VOTable, SDFITS, ASDM, etc.) which could be manipulated with standard astronomical packages." and **Level 2** as "Calibrated, science ready data with the instrument signature removed." However, many HEA **event-list**s include spatial and time axes that are calibrated physical quantities, but the spectral axis is instrumental and requires application of the IRFs to remove this signature. This is typically done because the IRFs can depend on the choice of region (spatial/time) from which the events are extracted (especially for telescope/detector combinations where the telescope position dithers on the sky during the exposure), which depends on the specific science case and therefore cannot be determined *a priori*. Such **event-list**s fall "between" *calib_level* 1 and 2. However, other **event-list**s may not have any calibrated axes or may have all axes calibrated, and it is important to be able to differentiate between these for data discovery. While the value for *calib_level* for any data product is left for the data provider to determine, we *suggest* that individual data providers set *calib_level* = 1 if an **event-list** is considered to be "uncalibrated" according to normal usage for their data products and set *calib_level* = 2 if an **event-list** is considered to be "calibrated" according to normal usage for their data products.

For **event-list**s, we propose that the calibration status of the spatial/spectral/time data axes be identified using the appropriate axis *calib_status* keyword (*s_calib_status* for the spatial axes, *em_calib_status* for the spectral axis, and *t_calib_status* for the time axis).

3.4 *obs_collection*

3.5 *access_format*

The *access_format* attribute specifies the format of the data product when downloaded as a file from the *access_url*. The analysis of HE data often requires use of multiple, related data products, for example an **event-list** combined with associated IRFs or ancillary files that can be employed by the user to create IRFs. These associated products are often bundled together with the **event-list** and we proposed in § 3.1 to assign such bundles *dataproduct_type* = **event-bundle**. While these bundles are typically not standardized across different projects, knowledge of the bundle content is useful for client applications to properly handle the bundles (for example to send the data to an appropriate visualization tool). This is readily achieved by encoding an appropriate MIME-type using the *access_format* attribute. In Appendix D we propose additional MIME-types for some common **event-bundle**s.

3.6 *s_ra/s_dec*

We propose that the attributes *s_ra*/*s_dec* be redefined to be the ICRS right ascension and ICRS declination of "a reference position (typically the center)" of an observation on the sky, rather than the ICRS right ascension and ICRS declination of "the center" of the observation. The center of an observation often is not useful for advanced data products that may be extracted from a cut-out from the progenitor observation, and many facilities allow an instrument to be displaced from the optical axis of the telescope, which means that the definition of "the center" of an observation may be unclear (especially for facilities for which the PSF varies strongly across the telescope field of view).

3.7 *s_calib_status*

We propose that *s_calib_status* encode the calibration status of an **event-list** dataset's spatial axes. Where multiple spatial axes are included in a dataset (*e.g.*, physical detector pixel coordinates, virtual detector coordinates corrected for distortions, world coordinates) then we recommend that the data provider use the coordinate system that is most likely to be preferred by the end user (typically the most fully calibrated spatial axes) to define *s_calib_status*.

3.8 *t_calib_status*

We propose that *t_calib_status* encode the calibration status of an **event-list** dataset's time axis. Where multiple time axes are included in a dataset (*e.g.*, instrument counter, absolute time) then we recommend that the data provider use the coordinate system that is most likely to be preferred by the end user (typically the most fully calibrated time axis) to define *t_calib_status*.

3.9 *em_calib_status*

We propose that *em_calib_status* encode the calibration status of an **event-list** dataset's spectral axis. Where multiple spectral axes are included in a dataset (*e.g.*, PHA, PI, energy) then we recommend that the data provider use the coordinate system that is most likely to be preferred by the end user (typically the most fully calibrated spectral axis) to define *t_calib_status*.

3.10 *o_ucd*

For an **event-list** we can consider that all measures stored in column values are observables. This is *the* fundamental difference between HEA **event-list**s and typical pixelated datasets. The current ObsCore Recommendation suggests that *o_ucd* be set to NULL for event lists; however this significantly hampers data discovery for HE datasets: since the data content of **event-list**s may vary significantly from facility to facility, meaningful discovery of HEA datasets *requires* the user be able to query the UCDs of the set of observables included in an **event-list**.

A natural way of doing this that is consistent with current usage would be to extend *o_ucd* to allow specification of *multiple* observables for **event-list**s, for example, *o_ucd* = *pos.eq,time,phys.pulseHeight*. Note that real **event-list**s may include an extensive set of columns (*e.g.*, a *Chandra* ACIS Level 1 **event-list** includes ∼20 columns, depending on observing mode) and several columns may represent similar (but not identical) observables (*e.g.*, event position in detector pixel coordinates, projected onto the focal surface, corrected for geometric distortions, corrected for spacecraft dither motion, mapped to world coordinates). Currently defined UCDs are not sufficiently fine-grained to be able to differentiate between these various cases but that may not be necessary since for data discovery purposes the user is typically interested in the "most calibrated" properties (*i.e.*, world coordinates in this example).

In the example *o_ucd* above, the example UCD *phys.pulseHeight* is used to represent the detector Pulse Height Amplitude (PHA). There is currently no UCD defined for a raw measure like PHA, but we propose the addition of *phys.pulseHeight* to the UCDList vocabulary, together with other UCDs that are relevant for HEA data, in Appendix C.

Finally, we note that extending *o_ucd* to allow specification of multiple observables would require similar adjustments to the other observable axis attributes *o_unit*, *o_calib_status*, and *o_stat_err*.

3.11 *facility_name*

3.12 *instrument_name*

3.13 *proposal_id*

To support advanced data products that may be constructed using data from multiple progenitor observations, we propose to modify the ObsCore Recommendation for *proposal_id* to allow multiple values, similar to *facility_name* and *instrument_name*.

4 Extensions to ObsCore Specific to High Energy Astrophysics Data

4.1 *ev_xel*

The lengths of each data axis (spatial, spectral, time, polarization) captured in attributes *s_xel1*, *s_xel2*, *em_xel*, *t_xel*, *pol_xel* do not apply non-pixelated data including **event-list**s, and ObsCore recommends that these attributes be set to -1. However, the dimensionality of an event list is an important property for data discovery, as the number of events often scales with signal-to-noise (and also data volume scales with number of events). We propose to add a new, optional attribute *ev_xel* that records the number of events in an **event-list** (effectively, the length of the "events" axis in the **event-list**s table).

4.2 *s_ref_energy*/*em_ref_energy*/*s_ref_oaa*/*em_ref_oaa*

For HE datasets that typically span decades of energy, both spatial resolution and sky coverage, and spectral resolution, can be strongly dependent on particle energy. The ObsCore Recommendation suggests that in such circumstance a *characteristic* value be specified for the spatial and spectral characterization attributes (*e.g.*, *s_fov*, *s_region*, *s_resolution*, *em_res_power*, *em_resolution*). We propose adding optional attributes (*s_ref_energy* for spatial characterization attributes and *em_ref_energy* for spectral characterization attributes) that define the energy (in units of eV) at which these characteristic values are specified.

For some HE datasets these attributes vary strongly with position in the field of view, typically as a function of off-axis angle (*i.e.*, the angular separation of the target or source from the telescope optical axis). We similarly propose adding optional attributes (*s_ref_oaa* for spatial characterization attributes and *em_ref_oaa* for spectral characterization attributes) that define the off-axis angle (in units of arcmin) at which these characteristic values are specified.

4.3 *t_coverage*

The global time bounds described by *t_min*/*t_max* in general are not sufficiently flexible when representing HEA datasets and advanced data products from any waveband. The former are typically composed of many stable time intervals (STIs)/good time intervals (GTIs), where data are only valid during the stable or good intervals, while advanced data products may be constructed from multiple progenitor observations that can span decades from the start time of the first observations to the stop time of the last observation. For the latter, data queries using *t_min*/*t_max* are useless, while for the former if data queries require precise observation timing then more detailed knowledge of the observation time coverage is necessary. We propose to add a new optional attribute *t_coverage* that would contain the list of observation intervals or STIs/GTIs as a TMOC 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 TMOC representation.

4.4 *energy_min/energy_max*

The existing attributes *em_min* and *em_max* that define the coverage of the spectral axis (defined as wavelength expressed in units of m) are not user friendly for HEA where datasets are generally selected according to an energy range (*i.e.*, inverse wavelength) in units of eV (or scaled units of eV, for example keV, MeV, GeV, TeV). Unlike the radio domain where $\lambda = c/\nu$, where c is an almost universally remembered physical constant, the conversion $\lambda = hc/E$ is not simple for the user to express. As the spectral range covered by HE data is many decades larger than for other wavebands, the accurate numerical representations of typical HE spectral ranges as *em_min*/*em_max* requires quantities with many digits of precision and exponents ranging from $\sim 10^{-5}$–10^{-19}, which are not easily comparable to current usage. Since specification of the spectral range is largely fundamental to data discovery in the HE regime, we propose to add attributes *energy_min* and *energy_max* that specify the minimum and maximum spectral range values in units of eV. Note that the sense of these attributes is opposite that of *em_min* and *em_max* because of the inverse wavelength relationship between energy and wavelength. An alternate approach would be to add attributes *em_min_energy* and *em_max_energy* that represent the energies corresponding to *em_min* and *em_max* in units of eV. This is less desirable since queries on an energy would need to be specified as $em\_max\_energy \leq E < em\_min\_energy$, which is likely confusing.

4.5 *obs_mode*

4.6 *analysis_mode*

References

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

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.

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.

A Detailed Science Use Cases for ObsCore

A.1 Event-List Data and Responses

A.1.1 Use Case — Search for event lists surrounding Sgr A* for morphological studies

Identify all HEA event lists encompassing Sgr A for initial selection for subsequent morphological studies. Since a large number (at least several hundred) of observations are expected to match this request and the focus is on morphological studies, only the event lists and not the event bundles are desired.*

Find all datasets satisfying:

(i) Target name = "Sgr A*" or position inside 30 arcmin from (266.4168, −29.0078).

(ii) dataproduct_type = "event-list".

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
(target_name = 'Sgr A*' OR
CONTAINS(POINT(s_ra, s_dec), CIRCLE, 266.4168, -29.0078, 0.5) = 1)
AND (dataproduct_type EQ 'event-list')
```

A.1.2 Use Case — Search for event bundles that include Cas A for spectrophotometric evolution studies

Identify all event bundles that include the Cas A SNR and have at least 1 million events for subsequent spectrophotometric studies of the SNR expansion. Since only a few observations are expected to match this request and because the focus is on spectrophotometric studies, the event bundles that include the responses or the ancillary products used to make the responses are required.

Find all datasets satisfying:

(i) Target name = "Cas A" or position inside 10 arcmin from (350.8584, +58.8113).

(ii) dataproduct_type = "event-bundle".

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
(target_name = ‘Cas A’ OR
CONTAINS(POINT(s_ra, s_dec), CIRCLE, 350.8584, +58.8113, 0.16667) = 1)
AND (dataproduct_type EQ ‘event-bundle’)
AND (ev_xel >= 1000000)
```

A.1.3 Use Case — Identify PSF response-functions for further analysis of previously downloaded data products

Identify all Chandra Source Catalog point spread functions for source detections that fall within 2 arcmin radius of (83.84358, -5.43639*) in the Orion star-forming complex for Chandra observation 4374. These PSFs will be used to analyze previously downloadedcatalog data products for the same field.*

Find all datasets satisfying:

(i) Position inside 3 arcmin from (83.84358, -5.43639).

(ii) dataproduct_type = "response-function".

(iii) dataproduct_subtype = "psf".

(iv) obs_id = "4374".

(v) obs_collection = "CSC2".

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
```

```
(CONTAINS(POINT(s_ra, s_dec), CIRCLE, 83.84358, -5.43639, 0.033333) = 1)
AND (dataproduct_type EQ ‘response-function’)
AND (dataproduct_subtype EQ ‘psf’)
AND (obs_id = ‘4374’)
AND (obs_collection = ‘CSC2’)
```

A.1.4 Use Case — Search for event lists that include a fully calibrated spectral axis for BL Lac for rapid spectrophotometric evaluation

Identify all event lists that include the BL Lac, have a fully calibrated spectral axis (i.e., spectral responses have already been applied), and have at least 10,000 events. These data will be used to prepare slides for a presentation. Note that since calib_status = 2 may not specify that the spectral axis is fully calibrated in physical units (HEA event lists are often considered "calibrated" even if the spectral axis is in pulse height units) the calibration status of the spectral axis must be checked explicitly.

Find all datasets satisfying:

(i) Target name = "BL Lac" or position inside 5 arcmin from (330.680338, +42.27777).

(ii) dataproduct_type = "event-list".

(iii) calib_level $\geq$ 2.

(iv) em_calib_status = "calibrated".

(v) ev_xel $\geq$ 10000.

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
(target_name = ‘BL Lac’ OR
CONTAINS(POINT(s_ra, s_dec), CIRCLE, 330.680338, +42.27777, 0.083333) = 1)
AND (dataproduct_type EQ ‘event-list’)
AND (calib_level >= 2)
AND (em_calib_status = ‘calibrated’)
AND (ev_xel >= 10000)
```

A.1.5 Use Case — Search for spatially resolved spectropolarimetric observation of the Crab with spectral resolution R > 100

Identify all event bundles for observations of the Crab that intersect the 1.0–100.0 keV energy range, have calibrated spatial and time axes, are spatially resolved in 2 dimensions in equatorial coordinates, have spectral resolution

R > 100, and include polarimetry measurements. Note that ObsCore specifies that the axes lengths — s_xel1, s_xel2, em_xel, t_xel, pol_xel — should be set to −1 for non-pixelated data like event lists, so these quantities are not useful for this query.

Find all datasets satisfying:

(i) Target name = “Crab” or position inside 5 arcmin from (83.6324, +22.0174).

(ii) dataproduct_type = “event-bundle”.

(iii) calib_level ≥ 2.

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
(target_name = ‘Crab’ OR target_name = ‘M1’ OR
CONTAINS(POINT(s_ra, s_dec), CIRCLE, 83.6324, +22.0174, 0.083333) = 1)
AND (dataproduct_type EQ ‘event-bundle’)
AND (calib_level >= 2)
AND (s_resolution > 100)
AND (energy_min < 100000.0) AND (energy_max >= 1000.0)
AND (o_ucd LIKE ’%phys.polarization%’)
AND (o.ucd LIKE ’%pos.eq%’)
```

A.2 High-Level Data Products

A.2.1 Use Case — Search for Chandra Source Catalog position error MCMC draws for X-ray detections in the vicinity of Gaia DR3 486718823701242368

Identify all Chandra Source Catalog position error MCMC draws for source detections that fall within 5 arcsec radius of (54.036061, +61.907633). The MCMC draws will be evaluated to establish whether there are potentially unresolved X-ray sources that may conincide with the white dwarf for observation planning.

Find all datasets satisfying:

(i) Position within 5.0 arcsec from (54.036061, +61.907633).

(ii) dataproduct_type = “draws”.

(iii) dataproduct_subtype = “poserr”.

(iv) obs_collection = “CSC2”.

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
(CONTAINS(POINT(s_ra, s_dec), CIRCLE, 54.036061, +61.907633, 0.0013888) = 1)
AND (dataproduct_type EQ 'draws')
AND (dataproduct_subtype EQ 'poserr')
AND (obs_collection = 'CSC2')
```

A.2.2 Use Case — Search for M31 source light curves and aperture photometry probability density functions that intersect a specific time interval

Identify all light curves and aperture photometry probability density functions of sources detected in the field of M31 covering the energy range 0.3–7.0 keV that include observation data in the interval MJD 56320–56325 TT during which interval a transient event was thought to have occurred. Because the data products are expected to include extremely sparse time axes, the t_coverage TMOC must be used for the query.

Find all datasets satisfying:

(i) Position within 1.5 degrees from (10.6847, +41.2688).

(ii) dataproduct_type = "timeseries" *or* "pdf".

(iii) calib_level = 4.

(iv) energy_min ≤ 0.3 *and* energy_max ≥ 7.0.

(v) t_coverage TMOC intersects MJD 56320–56325 TT.

```
SELECT * FROM ivoa.obscore
NATURAL JOIN  ivoa.obscore-hea
WHERE
(CONTAINS(POINT(s_ra, s_dec), CIRCLE, 10.6847, +41.2688, 1.5) = 1)
AND ((dataproduct_type EQ 'timeseries') OR (dataproduct_type EQ 'pdf'))
AND (calib_level = 4)
AND (energy_min <= 300.0) AND (energy_max >= 7000.0)
AND (INTERSECTS(TMOC(17, t_coverage), ...) = 1)
```

B Vocabulary Enhancements

While the IVOA Data Product Type Vocabulary (http://www.ivoa.net/rdf/product-type) provides terms, labels, and descriptions for many types of astronomical data products, there are some additions and changes that are appropriate to better support HE datasets.

We propose to add the data product types outlined in § 3.1 (*i.e.*, **event**, **event-bundle**, **draws**, **pdf**, **region**, **response-function**) and slightly modify the existing definition of **event-list** so that it aligns more accurately with the definition in § 3.1.

Additionally, we propose to clarify the terms that use the word "flux" in the description (currently **light-curve**, **polarization-resolved-dataset**, **polarized-spectrum**, and **spectrum**) to determine whether they are applicable to HE data. The issue is that the term "flux" is not defined, and the standard astronomical definition of "flux" for at least the last century has been an energy flux (with SI units $\mathrm{W\,m^{-2}}$). This interpretation is bolstered by the statements "flux or magnitude" applied to several of the descriptions since optical/IR magnitude and energy flux density are tightly related. However, with this definition many HE data products (that typically have units of counts) would not satisfy the descriptions of a light-curve or a spectrum (for example), event though common usage in the HEA community would term the corresponding products "light-curve" or "spectrum". Restating the descriptions to explicitly state "Particle or energy flux or magnitude" where appropriate would resolve this ambiguity.

C UCD Enhancements

C.1 Pulse Height

For HEA, the signal observed in a given detector spectral channel is the result of event counting and would typically be recorded as a Pulse Height Amplitude (PHA), or perhaps a Pulse Invariant (PI) value that is calculated from PHA by applying an appropriate gain calibration. The PHA (or PI) can be related to the incident particle energy by applying the appropriate IRF, and higher data calibration level products may replace or augment these values with quantities such as energy, or perhaps particle or energy flux.

The is currently no UCD defined for a raw pulse height amplitude measure like PHA (or PI). PHA is such an important quantity to HEA datasets that we propose adding a new UCD *phys.pulseHeight* for these raw data values. We note that the background signal (both of instrumental and cosmological origin) may be significant for many HE detectors and so the detected events are unrelated to any observed source on the sky. A current proposed solution suggests using *src.var.amplitude;src.var.pulse;stat.uncalib* for PHA, but this is not really appropriate since the connection to *src* ("observed source viewed on the sky") is misleading and *src.var.amplitude* is defined as the "amplitude of variation" of the source which is a completely separate concept from an astronomical perspective.

D MIME-type Enhancements

D.1 application/x-gadf

E Changes from Previous Versions

No previous versions yet.

F 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/