ObsCore Extensions

Mark Cresitello-Dittmar

IVOA Interop, College Park, MD

June 4, 2025

Session Overview

- This session is formally starting the process of moving the individual proposals for domain-specific ObsCore extensions (Radio, TimeDomain, High Energy) from the various Interest Groups, through the Working Groups to the appropriate Standards.

- Each of these proposals have had working group members actively involved in the production of the Notes from the beginning, and there has been significant cross-group collaboration over the past Semester to unify the content.

- We apologize to newcomers, in the interest of putting the focus on moving forward, we will not be providing much backstory on the various efforts. There have been several presentations in past Interops on the Notes that one can reference.

Session Structure

- General outline

  - Introductory overview of the state of each proposal

  - Working group statements of their review of the proposals

  - Discussion/Planning

Session Goals

- Gain Working Group acknowledgement of content under their umbrella

- Identify problem spots in proposals; conflicts with current design/definition

- Clarify any ambiguities raised by Working Groups reviewing proposals

- Provide input to Interest Group discussion points

- Identify impact on standards, defining roadmap forward

Introduction
Interest Group Statements

- We asked each interest group to provide:

  - A couple slides summarizing the state of their Note and any implementations

  - A statement on the following questions:

    - What are the features about your domain which require an extension to ObsCore in order to facilitate the discovery of your data?

    - Are there open questions which still need to be resolved within your group on your Note?

    - Are there particular questions which you feel need discussion at this session?

- I've included their full responses here, but will edit for time, highlighting significant points.

Radio Extension

Radio Extension

Unique features of the domain for data discovery

- There can be significant variation in some of the parameters within the spectra bounds of a single observation. This is especially true at low frequencies where FoV and resolution can vary by an order of magnitude between the lowest and highest frequency. This is true for many types of radio data.

- For visibility data, characterization of the UV coverage is important for discovery. High-fidelity imaging requires good UV coverage, but modeling can be done with data where the UV coverage is less good.

- For single-dish data, the use of "single-pixel cameras" means that the scanning mode and/or tracking mode have to match the science goals.

Content provide by Mark Kettenis

The Radio ObsCore extension
Status of the note

- Former proposed recommendation turned into note:

  - Repository: https://github.com/ivoa-std/ObsCoreExtensionForRadioData

  - Latest PDF draft: https://github.com/ivoa-std/ObsCoreExtensionForRadioData/suites/37685707812/artifacts/3009955098

- Some small open issues:

  - *s_resolution_min* and *s_resolution_max* already in ObsCore as optional

    - Broader discussion about mandatory and optional columns?

  - Vocabulary for *scan_mode* and *tracking_mode*

  - Registry aspects don't use table utype as proposed by Marcus Demleitner

JIVE

RadioIG: Mark Kettenis

The Radio ObsCore extension
Prototype implementations

- DaCHS implements an earlier working draft of the extension

  - DaCHS doesn't implement optional ObsCore columns

    - *s_resolution_min/_max* landed in extension table

  - Prototype service running at JIVE: with EVN VLBI data:

    - https://evn-vo.jive.eu/tap

  - Some prototyping in SKA context based on DaCHS as well

- CAOM-2.5 based implementation at CADC anticipating ObsCore-1.2

  - Still under development; no datalink integration

  - Prototype service at CADC with LOFAR LOTSS data:

    - https://ws.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/caom25/argus

- Another implementation that includes single-dish data would be good

  - Planned at INAF and Yebes, but timescale uncertain

Time Domain Extension

- What do I need in terms of time properties to query for a time series dataset?

- Obscore 1.1 : https://www.ivoa.net/documents/ObsCore/ contains many general metadata for discovery including time features .

Use case example :

Give me a light curve from an ObsTap service satisfying such constraints :

—> ivoa dataproductype vocabulary :
https://www.ivoa.net/rdf/product-type/2024-03-22/product-type.html

- data type is 'light curve'

- belongs to ZTF survey

- taken around position ra1, dec1

- taken between [date1 date2]

- contains more than 10 time stamps

- Time Interval > 1 week

- *dataproduct_type* == 'light-curve'

- *obs_collection* == 'ZTF'

- *s_region* contains (ra1,dec1)

- *obs_date* in [date1 date2]

- *t_xel* > 10

- *t_max -t_min* > 7 (days)

Status for ObsCore Time Extension, M.Louys Interop May 2025

☐ Extension for Time series

- What is missing for all time variant data sets ?
- Enhance the choice of data product_types
  initially *dataproduct_type == time series (Obscore 1.1)*
  - today ivoa.net/rdf/product-type has :
    light-curve, time-cube, dynamic-spectrum, etc.

- Details on time sampling : period, sampling extent
- Period folding
- Reference to time system

- Working draft in progress for this specification

https://github.com/ivoa-std/ObscoreTimeExtension

International
Virtual
Observatory
Alliance

ObsCore Metadata Extension for Time Properties

Version 1.0

IVOA Working Draft 2024-07-17

Working Group
   DM
This version
   https://www.ivoa.net/documents/ObscoreTimeExtension/20240717
Latest version
   https://www.ivoa.net/documents/ObscoreTimeExtension
Previous versions
   This is the first public release
Author(s)
   Mireille Louys, François Bonnarel, Vincenzo Galluzzi, Baptiste Cecconi, Ada Nebot
Editor(s)
   Mireille Louys

Time features available in ObsCore 1.1 (current version)

ObsCore metadata	Definition	Utype (Characterisation DM based)	UCD	Units	Mandatory/ optional
t_min	Time start of the sequence (MJD)	Char.TimeAxis.Coverage.Bounds.Limits.LoLim	time.start;obs.sequence	d	man
t_max	Time end of the sequence	Char.TimeAxis.Coverage.Bounds.Limits.HiLim	time.end;obs.sequence	d	man
t_exptime	Exposure time (sum of multiple exposures)	Char.TimeAxis.Support.Extent	time.duration;obs.exposure	s	man
t_resolution	Minimal interpretable time difference	Char.TimeAxis.Resolution.Refval	time.resolution	s	man
t_xel	Number of time stamps in the series	Char.TimeAxis.numBins	meta.number	null	man

Time sampling features to add to ObsCore

tObs metadata	Definition	Utype/datamodel path	UCD	Units	Mandatory/ optional
%minimal and maximal sample duration . The estimation of signal to noise signal (SNR) can be derived from this.					
t_exp_min **t_sample_extent_min**	minimal length of time sample (min integration time)	Char.TimeAxis.Sampling.Extent.lo Lim	time.duration;obs.sequence ;stat.min.	s	opt
t_exp_max **t_sample_extent_max**	maximal length of time sample (max integration time)	Char.TimeAxis.Sampling.Extent.hi Lim	time.duration;obs.sequence ;stat.max	s	opt
%time space between 2 time samples / cadence					
t_delta_min	minimal length of time interval between 2 observations / cadence (min)	Char.TimeAxis.Sampling.Period.lo Lim	time.interval;obs.sequence; stat.min.	s	opt
t_delta_max	maximal length of time interval between 2 observations / cadence (max)	Char.TimeAxis.Sampling.Period.hi Lim	time.interval;obs.sequence; stat.max	s	opt
t_fold	folding period length		time.period	s	opt
t_fold_phase_reference	Time stamp of folding start		meta.ref;time.phase	s	opt

- The datasets description in Obscore 1.1 does not constrain to use one specific TIME coordinate system for the data.
- In VOTable, the data representation in the time series data rely on a TIMESYS element.

  add to **time extension** table
- Is it useful to query on it ?
- At least it is useful to get in the query response and let a client application prepare time coordinates' conversion if needed

ivoa.t_obs	Definition	VODML-ID IN Coords DM and **TIMESYS attribute**	UCD	Units	Mandatory/ optional	Query setup
% Time Coordinate system						
t_origin	Time(frame origin)	TimeOffset.time0 (TBC) timeorigin	time.epoch	null	opt	
t_scale	Time frame scale	TimeFrame.timeScale timescale	time.scale	null	opt	
t_refposition (barycenter, heliocenter, ...)	Time reference position	TimeFrame.refPosition refposition	meta.id;time	null	opt	
t_refdirection (e.g. for solar observations)	Time reference direction	TimeFrame.refDirection refdirection	meta.id;time	null	opt	
%Time representation ISOtime , MJD, JD , ...						
t_format	Time representation	TimeMJD, TimeJD, IsoTime not part of TIMESYS	meta.id;time	null	opt	MJD

- Should this extension be included in the next ObsCore 1.2 ?
  - 6 columns names to add only

- If we choose a dedicated Time extension:
  - Table name for the ObsTAP table :
    - ivoa.obs-time ? ivoa.obscore-time ?
  - Utype expression for this table :
    - ivo://ivoa.net/std/obscore#time-1.0?
    - Need to harmonise with TAP service declarations for Registry WG

  - In both cases, new implementations are needed.
  - Pulsar data distribution would be a good candidate to exercise time extension and radio extension at the same time.

High Energy Extension

High Energy Astrophysics (HEA) Data

- HEA instrumentation usually identifies individual particles (*e.g.*, photons) with energies ~0.1 keV and higher (up to TeV+)

  - The primary datasets are **event-lists** that record detected particle event properties (*e.g.*, spatial, spectral, time information) for each event

    $\Longrightarrow$ **there are multiple observables per dataset**

  - In some HEA energy regimes, event-lists may often include **calibrated spatial and temporal axes** but have an **uncalibrated spectral axis** with photometric units of counts; nevertheless, these event lists are typically considered to be ``**calibrated**'' in those regimes

HEA Response Functions

- Mappings from physical attributes (*e.g.*, real particle energy) to observables (*e.g.*, measured pulse height [PHA]) are often **probabilistic** and **not invertible**

- Instrument **response functions** provide the mappings and depend on specific observation details and conditions

$\Rightarrow$ **there must be a way to associate these data products with the event-lists**

- Evaluating response functions *may* require scientific input, potentially necessitating creation by the end user, using additional data products that similarly should be associated with the event-lists

Spectrum

$PHA \rightarrow E$

Light Curve

$t_{obs} \rightarrow t_{TT}$

y

Image

x

$(xraw, yraw) \rightarrow (\alpha, \delta)$
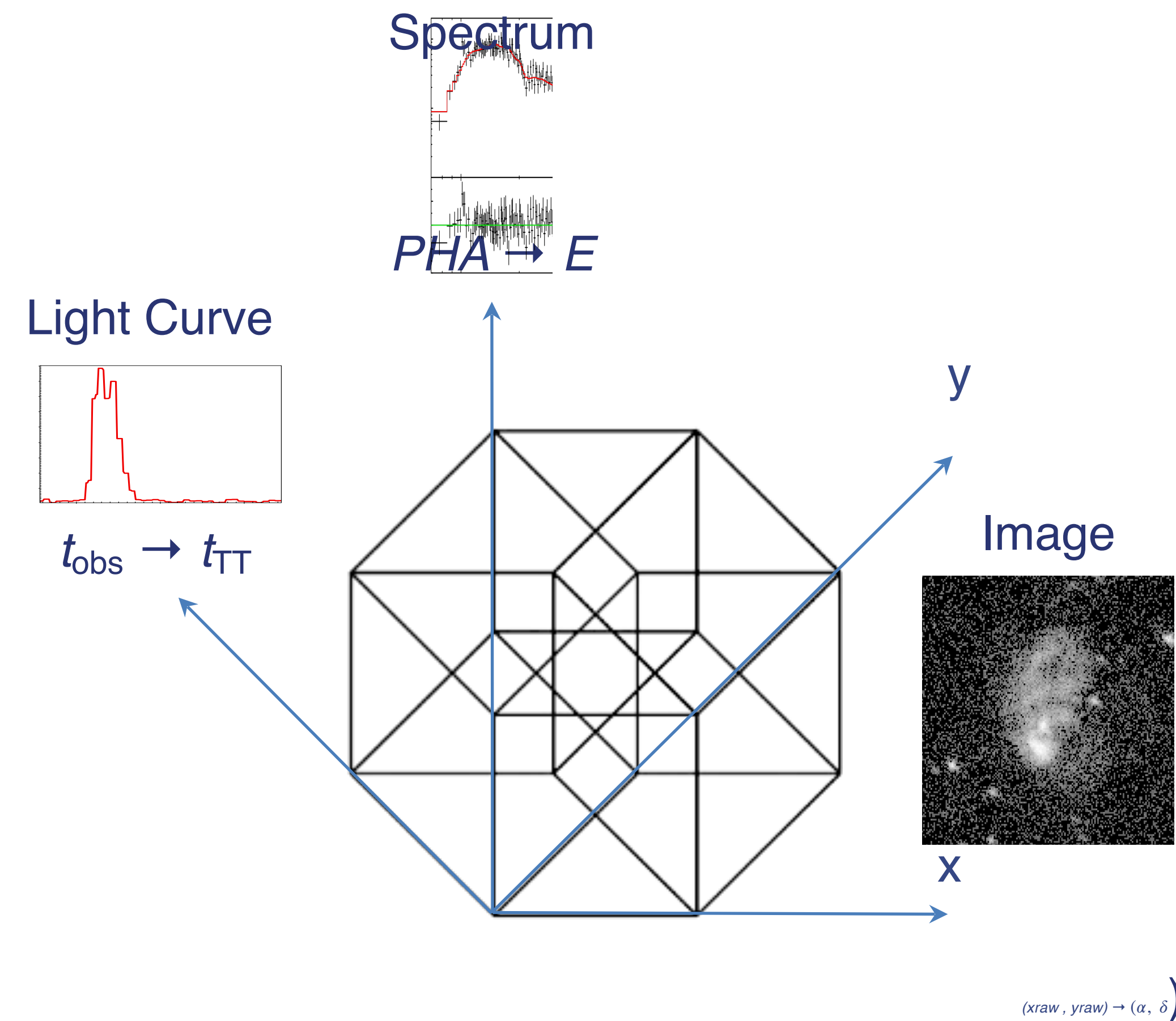
The HEA tesseract

HEIG ObsCore Extension Note Status

- A draft note is available that includes suggested HEA-specific extensions and updates to core ObsCore definitions (such as dataproduct_type)

- Some sections contain only placeholders based on earlier feedback

- A limited number of use cases are currently provided

- The note requires broader examination by the IVOA HEA community, with further contributions needed to develop placeholders and create more use cases

- Considerable additional input is necessary for the sections on vocabulary, UCD, and MIME-type enhancements

- There are currently no implementations underway or planned (the work is still in the early stages)

Open Questions

- How do we best coordinate across different wavebands for updates that extend beyond just the HEA waveband, even if HEA "arrives first"?

- What is the plan for converging ObsCore dataproduct_types and the data product type vocabulary?

- We have some concerns regarding the absence of clear definitions (for instance, in vocabulary entries) that may impact HEA data products, as prevailing assumptions from other wavebands may not apply

    - For example, the messenger may not always be photons, and "calibrated" data often lacks calibrated spectral axes or photometry

    - What is the most effective approach to tackle these cross-waveband issues?

Discussion Points

Discussion Points

Address or Identify players taking actions to resolve

- Optional vs Mandatory columns in ObsCore: implementation support is lacking for optional columns.

- Migration of ObsCore lists/enums to Vocabularies - reconciling the content, and backward compatibility issues

- Properties included for Query Response only: not something one would query on, but useful/ important for user to have in response to assess the usefulness of, or compare data. Is this in keeping with ObsCore design and current practice?

- scan_mode/tracking_mode vocabularies: proper definitions and alignment with existing content in ObsLocTap and CAOM

- Handling vocabulary definitions which have waveband bias.

- Extension table details (name, utype expression). Proposal by Marcus D.

Discussion Points

Address or Identify players taking actions to resolve

- From SphereX presentation this interop

  - CAOM/ObsCore 'monolithic' and they are experiencing latency/ performance challenges. Adding more content will only make that worse.. is there something to do or consider regarding this in ObsCore? Or is their solution (locally dividing into smaller tables) the correct answer?

  -