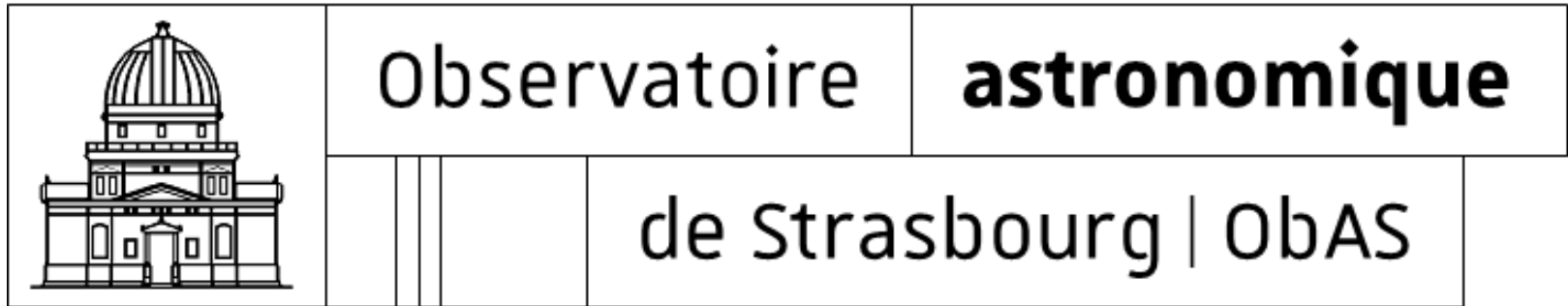


ObsCore extensions modeling and TAP serialization : application to High Energy data

VO Working groups involved :
Radio and HEIG interest groups, Data model WG.

Mireille Louys and HEIG Interest Group
CDS and Observatoire de Strasbourg
ICube Laboratory, Strasbourg University



□ ObsCore Model Extension in UML

- We have 2 initiatives going on to add new concepts and elements to ObsCore, the Observation Components data model.
- These belong to 2 different spectral regimes :
Radio domain : <https://github.com/ivoa-std/ObsCoreExtensionForRadioData>
High Energy domain : <https://github.com/ivoa/HighEnergyObsCoreExt>

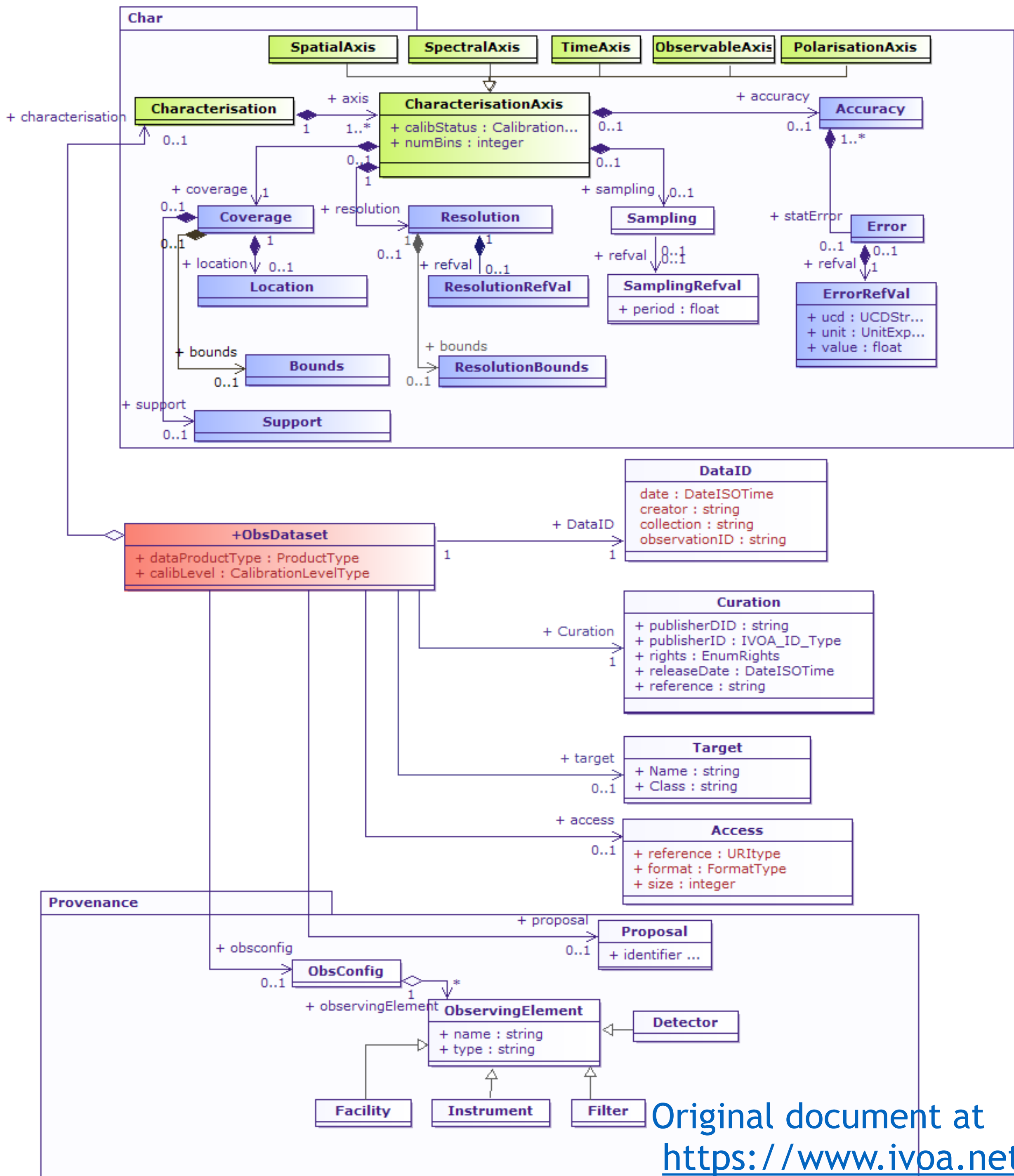
In this approach I try to evaluate which new classes could help to extend the ObsCore DM in order to support the needs expressed in these notes.

Here I focus on High energy data, where new data products are requested like [advanced data products](#) and [response data products](#).



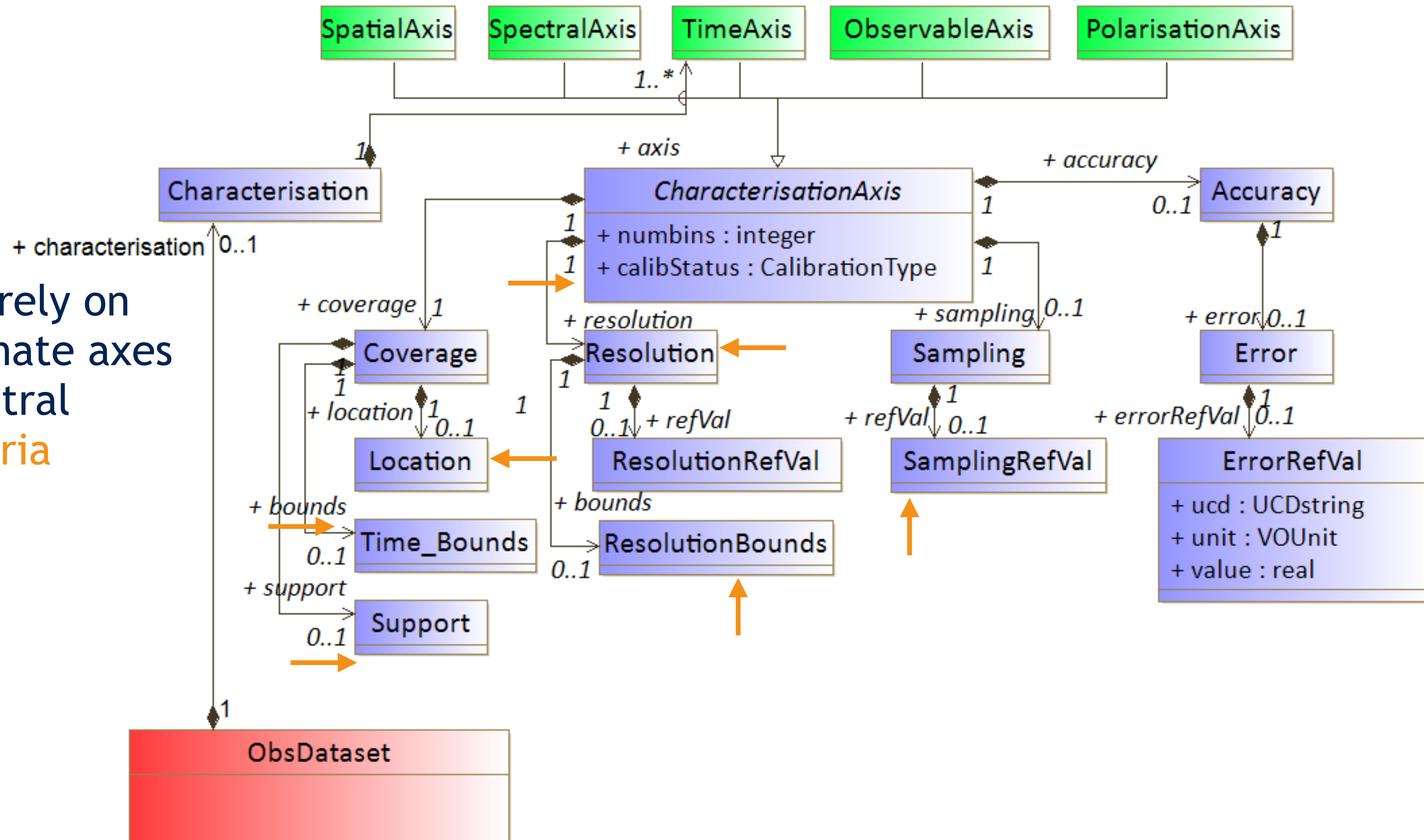
Obscore 1.1

Ivoatex transcription going on : see <https://github.com/ivoa-std/ObsCore>



Original document at <https://www.ivoa.net/documents/ObsCore/20170509/REC-ObsCore-v1.1-20170509.pdf>

Characterisation of physical axes



Many data search rely on
dependent coordinate axes
Space, Time, Spectral
—> **discovery criteria**

Data management metadata

Many data search rely on data set identification as collection name, dataID, observationID, proposal ID, release date, rights, etc.

—> **discovery criteria**

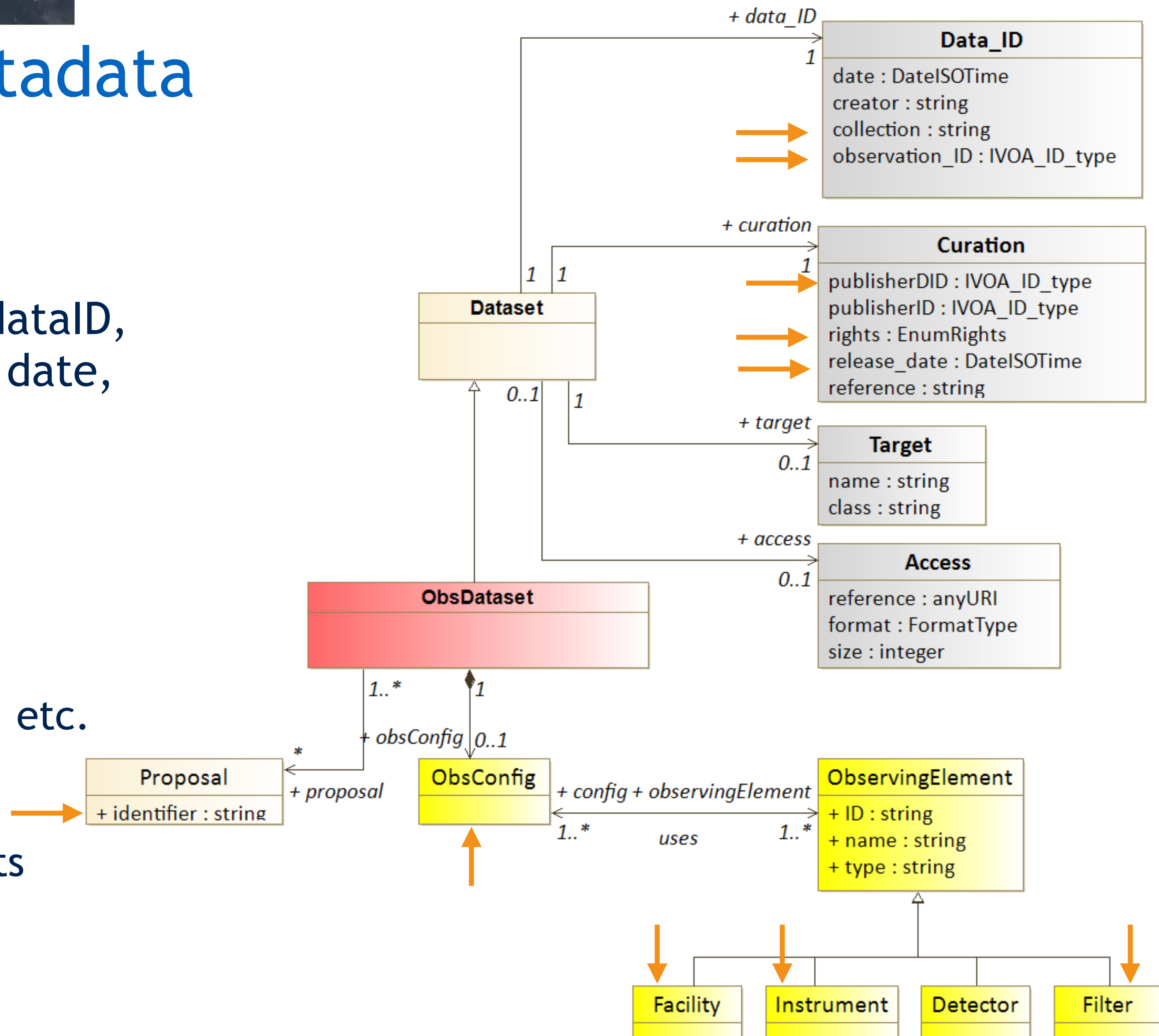
Observing configuration

Telescope, instrument, observing mode, etc.

—> focus on **observation process**

Still simple in Obscore 1.1

Difficult to generalize among all projects

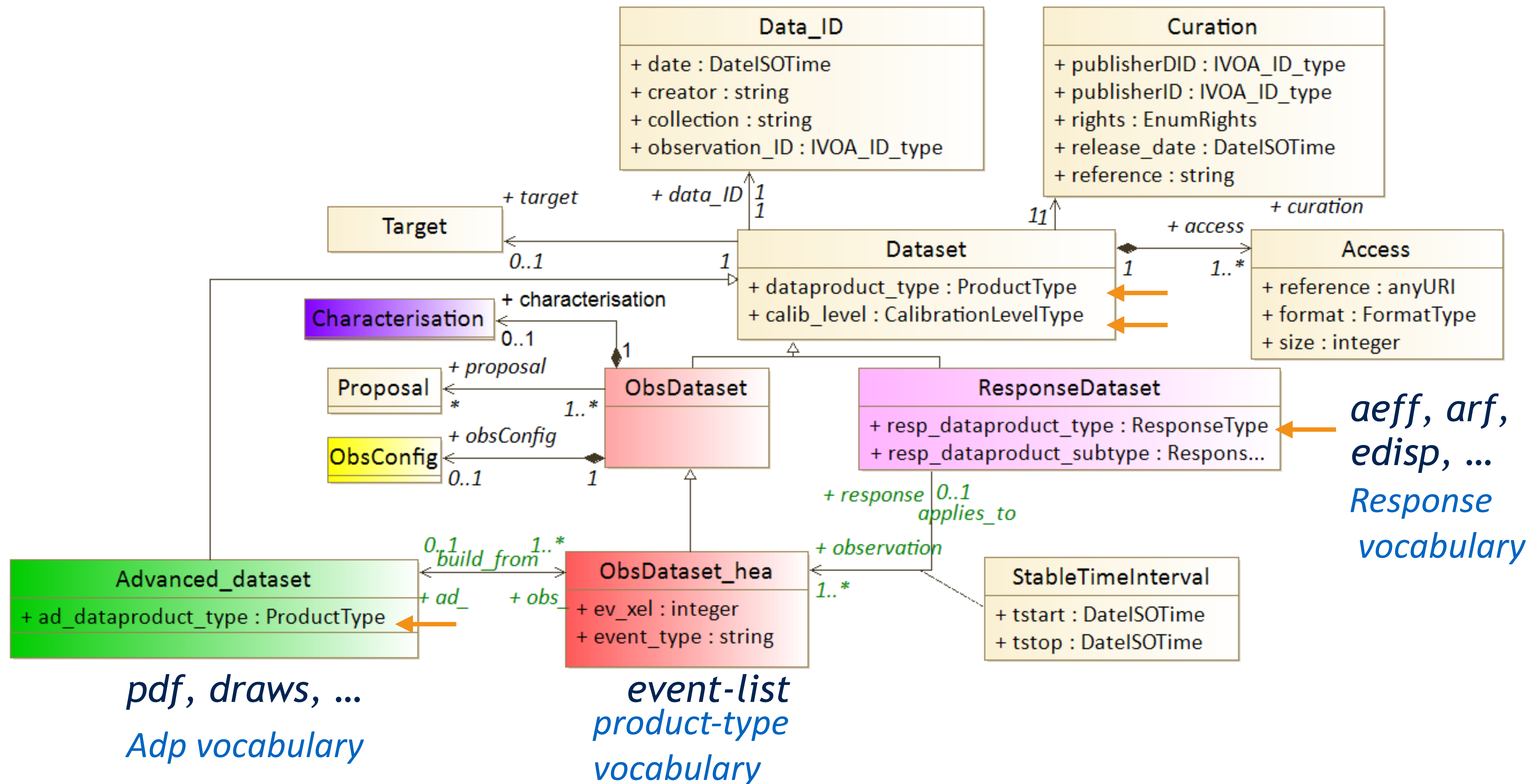


□ Serialising new classes

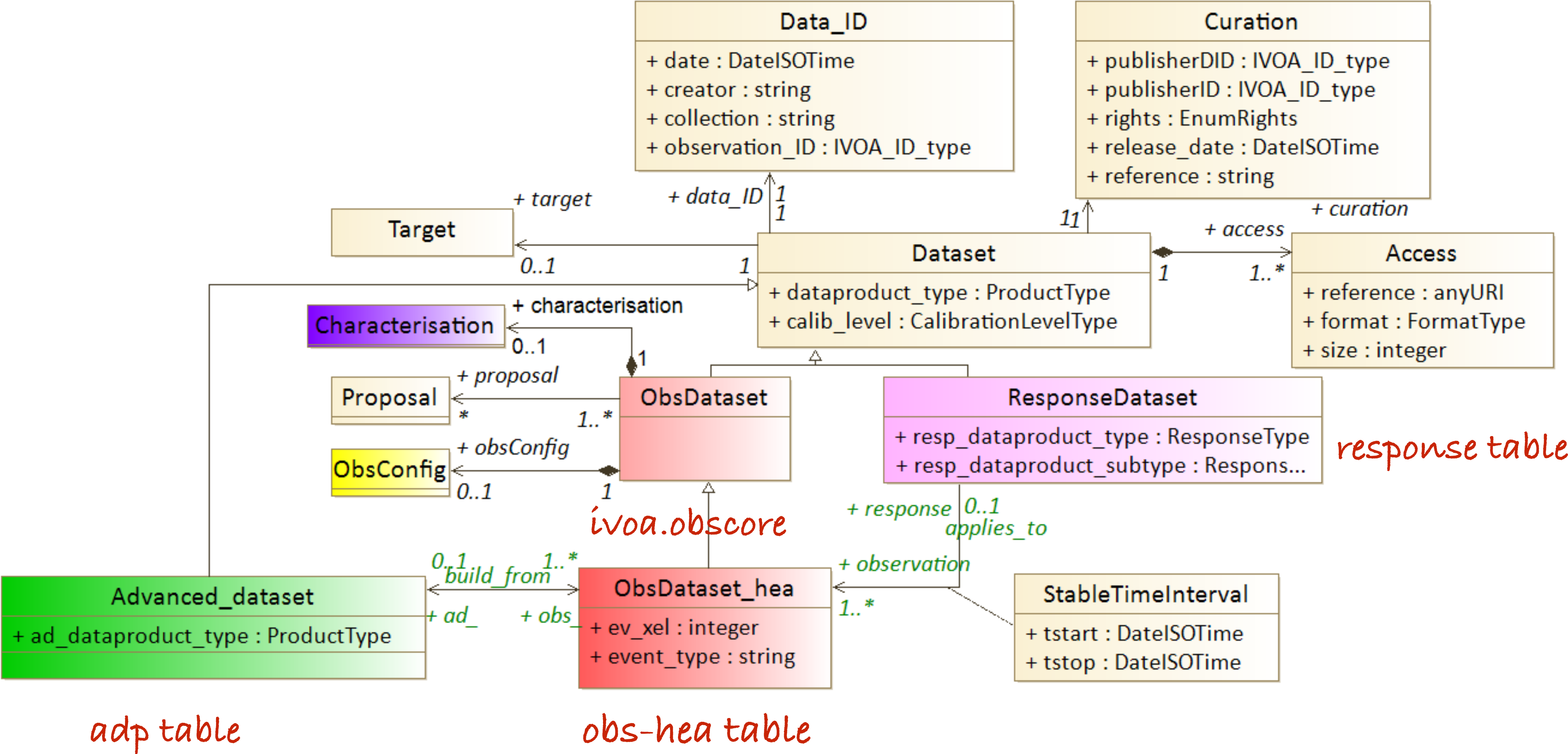
- Class reuse:
event-list datasets can be covered by [ObsDataset-hea](#), derived from ObsDataset with new attributes about events .
- [Response datasets](#)
 - are new data products : vocabulary defined in the note : *aeff* , *arf* , *edisp* , etc .
 - do not have Characterisation features compliant to an ObsDataset.
 - many attributes would be null in the ObsTAP mandatory columns.
 - can reuse the basic data management classes : Dataset, Access, DataId, etc.
 - they are linked to an observation (or a set of) via the '*applies_to*' relation.
—> New classes required
- [Advanced data products](#)
 - are new data products : vocabulary defined in the note : *draws* , *pdf* , etc .
 - derive from an ObsDataset but have observables like probabilities, statistical parameters, etc .
This is already interpretation material, and no longer pure sky measures.
 - they are linked to an observation (or a set of) via the '*build_from*' relation.



Dataset categories



Dataset categories



□ TAP tables proposal

- **ivoa.obscore** figures general metadata for obscore
- **ivoa.ext-he** figures added metadata on observation conditions + event specific properties
A row represents an instance of a **ObsDataset-hea** class .
- **ivoa.response** for response functions and binding to applicable observation (obs_id, obs_publisher_did)
A row represents an instance of a **Response dataset** class .
- **ivoa.adp** description and properties of **Advanced_dataset** class
A row represents an **Advanced_dataset** class .
The *'build_from'* relation uses (*obs_id*, *obs_publisher_did*) from the progenitor ObsDataset.





TAP table for advanced data products

Fkey{

Column Name	Datatype	Size	Units	ObsCoreDM Utype	UCD	
ad_dataproduct_type	adql:VARCHAR	TBD	NULL	Ad_Dataset.dataProductType	meta.code.class	Pdf
ad_dataproduct_subtype	adql:VARCHAR	TBD	NULL	Ad_Dataset.dataProductSubtype	meta.code.class	Probability density
calib_level	adql:INTEGER	NULL	NULL	ObsDataset.calibLevel	meta.code;obs.calib	4 or 5?
obs_collection	adql:VARCHAR	TBD	NULL	DataID.collection	meta.id	Chandra
obs_id	adql:VARCHAR	TBD	NULL	DataID.observationID	meta.id	ID of the progenitor
obs_publisher_did	adql:VARCHAR	TBD	NULL	Curation.publisherDID	meta.ref.ivoid	ID of the progenitor
ad_publisher_did	adql:VARCHAR	TBD	NULL	Curation.publisherDID	meta.ref.ivoid	Publisher did of this dataset
ad_release_date	adql:TIMESTAMP	NULL	NULL	Curation.releaseDate	time.release	Release_date of this dataset
ad_bib_reference	adql:VARCHAR	TBD	NULL	Curation.reference	meta.bib	Reference to the publication
ad_creator_name	adql:VARCHAR	TBD	NULL	DataID.creator	meta.id	Creator of this advanced
ad_access_url	adql:CLOB	NULL	NULL	Access.reference	meta.ref.url	As in ObsDataset
ad_access_format	adql:VARCHAR	NULL	NULL	Access.format	meta.code.mime	As in ObsDataset
ad_access_estsize	adql:BIGINT	NULL	kbyte	Access.size	phys.size;meta.file	As in ObsDataset
target_name???	adql:VARCHAR	TBD	NULL	Target.name	meta.id;src	Source name for pdf ?
ad_o_ucd	adql:VARCHAR	TBD	NULL	Char.ObservableAxis.ucd	meta.ucd	Stat.probability

See Google doc
advanced data products [here](#)



□ Query examples

*Give me all **advanced data products** that are derived from Obsdataset_healpix with obs_publisher_id= 'Y', i.e that have a build_from relation to the datapoint with obs_publisher_id='Y'.*

```
SELECT ad_publisher_id, ad_access_url, ad_datapoint_type, ad_datapoint_subtype,  
ad_release_date ,ad.obs_publisher_id  
FROM ivoa.adp as ad  
JOIN ivoa.obscore as o  
WHERE  
ad.obs_publisher_id = o.obs_publisher_id  
AND o.obs_publisher_id='Y'
```

This would give all advanced products stemming from Y .

NB1 : here we do not need items from ivoa.obscore-healpix and only ivoa.obscore

NB2: one can also add a constraint on *ad_release_date* if needed : AND ad_release_date > 2010

In case more columns are requested they can be queried to the original dataset it belongs to from the **ivoa.obscore** table or **obscore.ext-he** table.



Response table

Fkey{

Column Name	Datatype	Size	Units	ObsCoreDM Utype	UCD	Response dataset	Psf example as in lan example in VOTable
resp_dataproduct_type	adql:VAR CHAR	TBD	NULL	ResponseDataset.data ProductType	meta.code.class	edisp	response-fonction
resp_dataproduct_subtype	adql:VAR CHAR	TBD	NULL	ResponseDataset.data ProductSubtype	meta.code.class	?	psf
obs_id	adql:VAR CHAR	TBD	NULL	DataID.observationID	meta.id	ID of the progenitor	4374
obs_publisher_did	adql:VAR CHAR	TBD	NULL	Curation.publisherDID	meta.ref.void	ID of the Obsdatase	ivo://cxc.cfa.harvard.edu/cda?10.25574/4374
resp_publisher_did	adql:VAR CHAR	TBD	NULL	Curation.publisherDID	meta.ref.void	Publisher did of this	??
resp_release_date	adql:TIME STAMP	NULL	NULL	Curation.releaseDate	time.release	Release_d ate of this	2024-10-18T09:25:13
resp_description	adql:VAR CHAR	TBD	NULL	Curation.reference	meta.bib	Reference to the	
resp_creator_name	adql:VAR CHAR	TBD	NULL	DataID.creator	meta.id	Creator of this	CXC
STI_tstart	adql:TIME STAMP		?	Char.TimeAxis.Coverag e.Support.STI.tstart	time.start;obs.ex posure	Stable time interval	52655.01173
STI_tstop	adql:TIME STAMP	TBD	?	Char.TimeAxis.Coverag e.Support.STI.tstop	time.stop;obs.ex posure	Stable time interval	52656.99358
resp_access_url	adql:CLO B	NULL	NULL	Access.reference	meta.ref.url	As in ObsDatase	https://cda.cfa.harvard.edu/csccli/retrieveFile?
resp_access_format	adql:VAR CHAR	TBD	NULL	Access.format	meta.code.mime	As in ObsDatase	image/x-fits
resp_access_estsize	adql:BIGI NT	NULL	kbyte	Access.size	phys.size;meta.fil e	As in ObsDatase	
resp_ucd	adql:VAR CHAR	TBD		Observable.ucd	meta.ucd		Phot.count ??

□ Query examples

*1. Give me all response data products that are applicable to Obsdataset_healpix with obs_publisher_id= 'Y',
i.e that have a “applies_to” relation to the datapoint with obs_publisher_id='Y' and more recent than 2010.*

Query :
SELECT resp_publisher_id, access_url, resp_dataproduct_type, resp_release_date ,resp.obs_publisher_id
FROM resp_table as r
JOIN ivoa.obscore as o
ON r.resp_obs_publisher_id = o.obs_publisher_id
AND o.obs_publisher_id='Y'
AND resp_release_date > 2010

*2. Give me all response data products of type **psf** that are applicable to ivoa.obscore-healpix with obs_publisher_id= 'Y',
i.e that have a “applies_to” relation to the datapoint with obs_publisher_id='Y'
Have ev_xel> 10000
Have been released after 2015.*

SELECT resp_publisher_id, resp_access_url, resp_dataproduct_type, resp_release_date ,resp.obs_publisher_id
FROM resp_table as r
JOIN ivoa.obscore-healpix as o ON r.resp_obs_publisher_id LIKE o.obs_publisher_id
AND o.obs_publisher_id='Y'
AND o.ev_xel> 10000
AND r.resp_dataproduct_type LIKE 'psf'
AND resp_release_date > 2015



□ Questions to solve

1. Can we define a better word for “[advanced data product](#)” ? Brainstorm needed.
In the attempt to standardise services for IVOA, this label is too context dependent.
2. Which Time Stamp representation in [Stable interval](#). MJD?
3. Do we need *target_name* for an ADP ?
Can we just use the target from the progenitor dataset if present?
4. Relation *Advanced_dataset* $\longleftrightarrow$ *Obsdataset-hea* :
How many instances would be bound through the *build_from* relation in general?
5. Are there use cases where one needs to select an advanced product/response on its own data axes?
spatial bounds, energy bounds, resolution, etc.
7. Would the binding Plane $\longleftrightarrow$ Artefact from CAOM be useful here for the response and advanced datasets



□ Conclusion

The modeling helps to clarify the relations between various types of data sets .

Implementing in TAP needs decisions on how we gather the various properties of these datasets.

Practical experience on TAP joins and TAP views will help to set up a TAP implementation .
Refer to DAL expertise.

Comments ? Questions ? Ideas ?

Thanks



Response dataset linkage

