

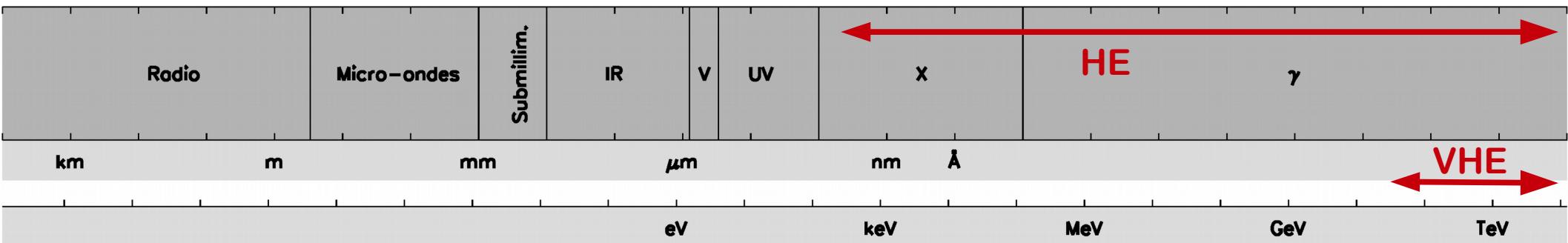
# VO Data Diffusion for IACTs

Mathieu Servillat

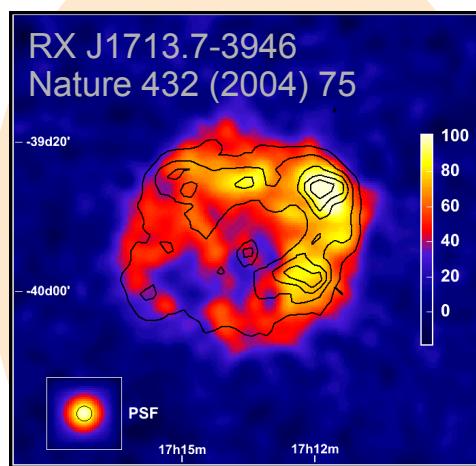
Observatoire de Paris  
Laboratoire Univers et Théories  
Paris Astronomical Data Centre



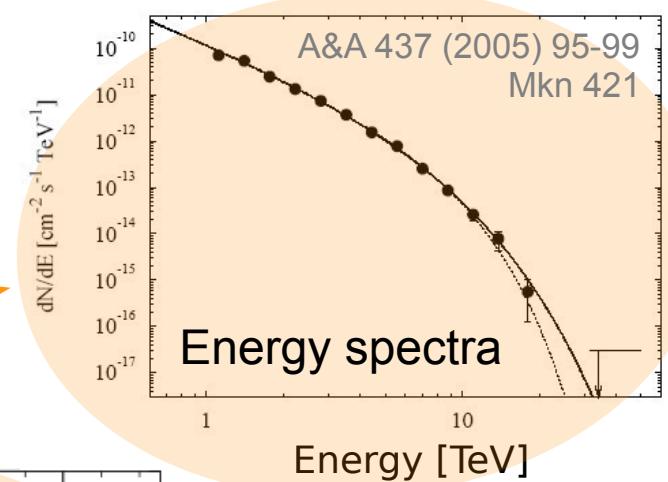
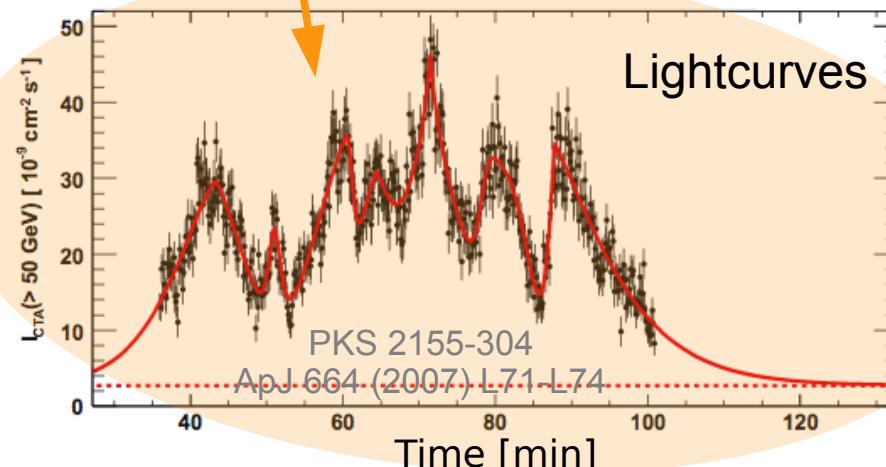
# Very high energy (VHE) data



- ◆ Several orders of magnitude
- ◆ Photon counting
- ◆ Low count statistics, high background
- ◆ **Event lists**  
(coordinates, time, energy)



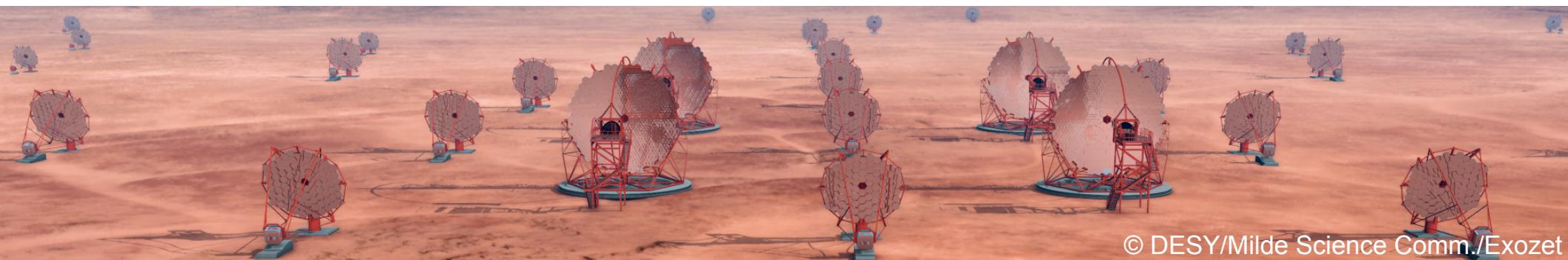
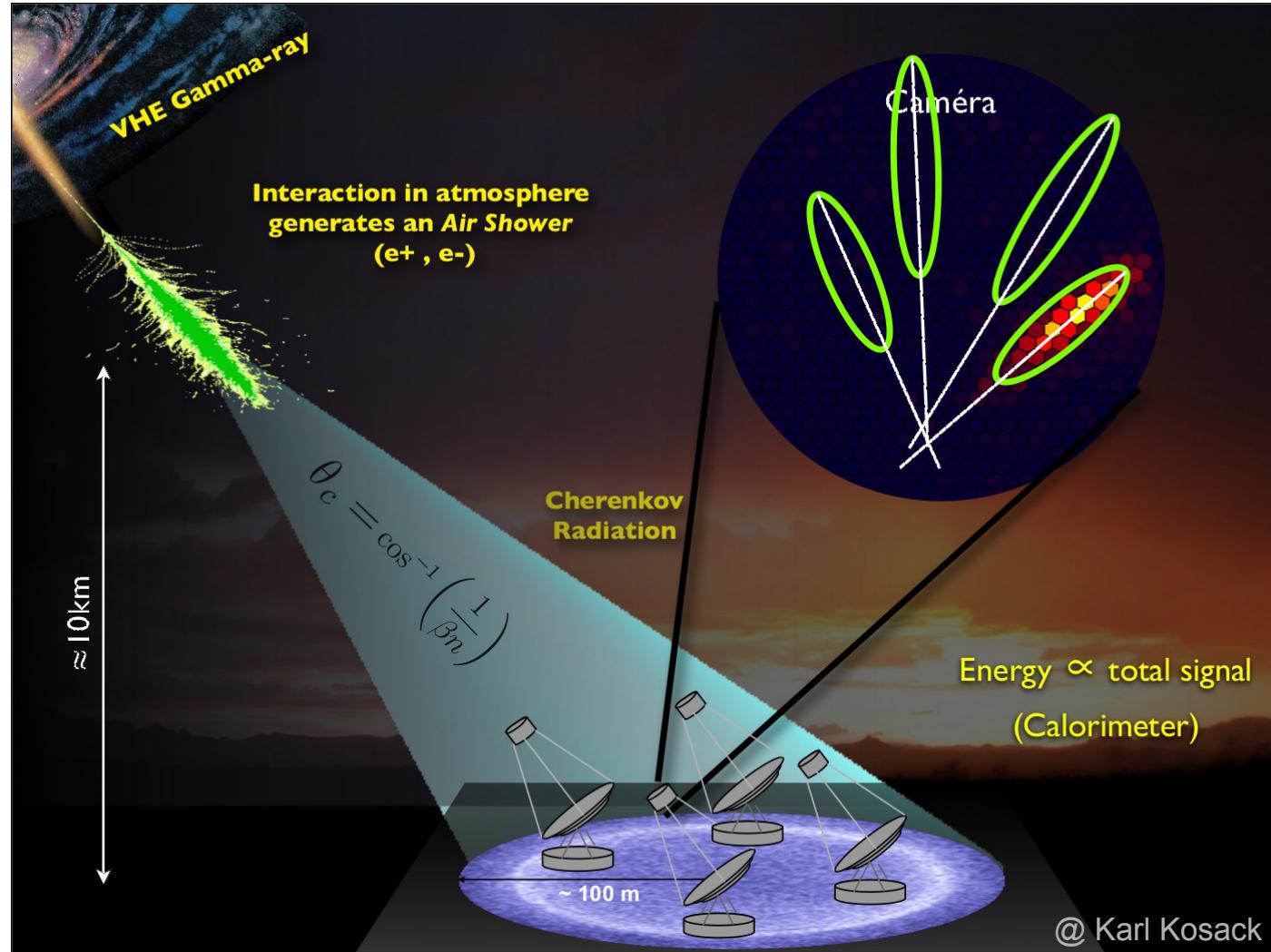
Images



Lightcurves



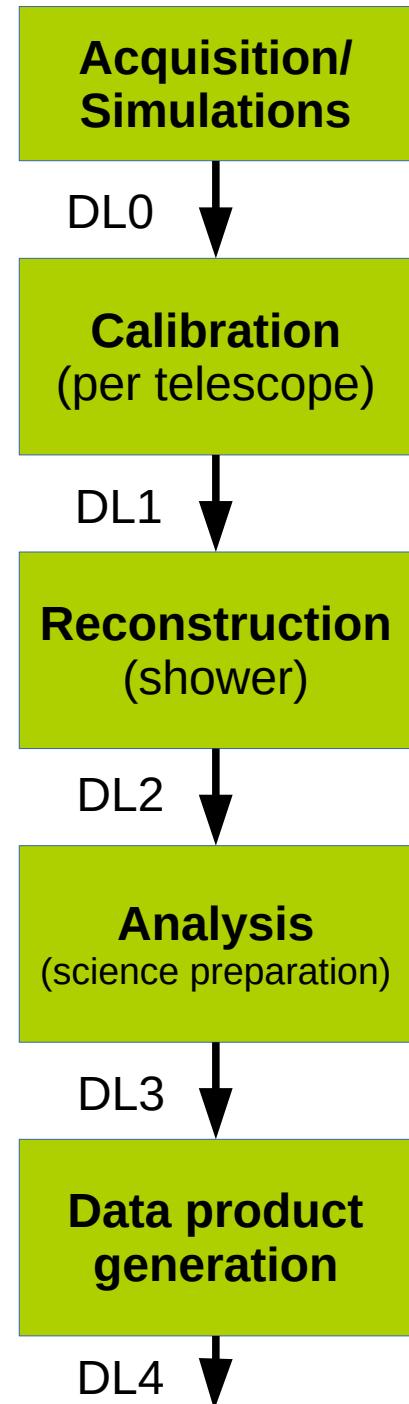
- ◆ **Dark nights** (small duty cycle)
- ◆ Field of view: 5-8 degrees
- ◆ **Event Reconstruction:**  
photon, particle shower,  
Cherenkov light  
(faint, few nanoseconds)
- ◆ **Atmosphere = calorimetre**  
Simulations, assumptions
- ◆ **Complex Metadata**,  
need to be structured



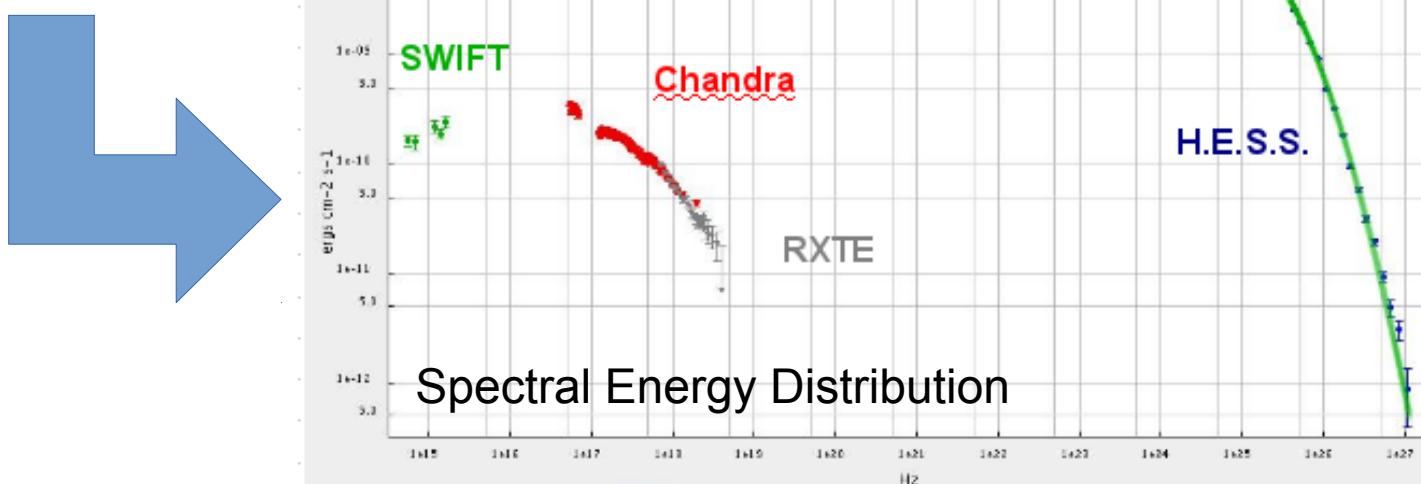
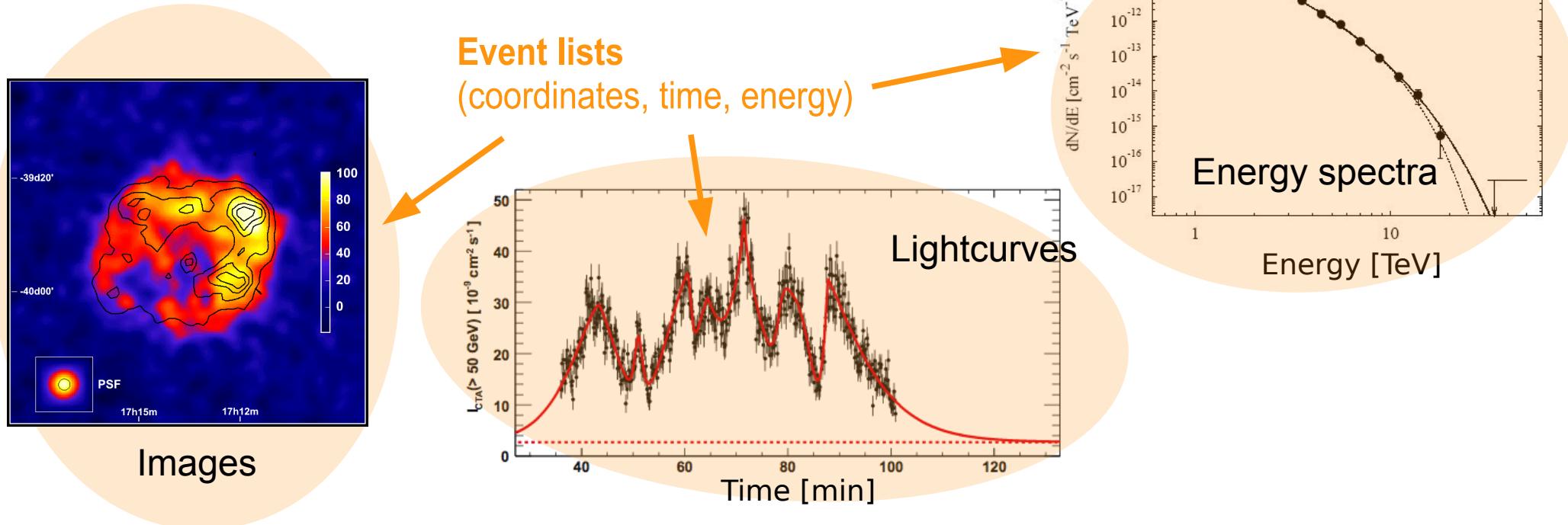
© DESY/Milde Science Comm./Exozet

# Data levels and workflow

Data Level	Short Name	Description	Data reduction factor
Level 0 (DL0)	DAQ-RAW	Data from the Data Acquisition hardware/software.	
Level 1 (DL1)	CALIBRATED	Physical quantities measured in each separate camera: photons, arrival times, etc., and per-telescope parameters derived from those quantities.	1-0.2
Level 2 (DL2)	RECONSTRUCTED	Reconstructed shower parameters (per event, no longer per-telescope) such as energy, direction, particle ID, and related signal discrimination parameters.	$10^{-1}$
Level 3 (DL3)	REDUCED	Sets of selected (e.g. gamma-ray-candidate) events, along with associated instrumental response characterizations and any technical data needed for science analysis.	$10^{-2}$
Level 4 (DL4)	SCIENCE	High Level binned data products like spectra, sky maps, or light curves.	$10^{-3}$
Level 5 (DL5)	OBSERVATORY	Legacy observatory data, such as CTA survey sky maps or the CTA source catalog.	$10^{-5} - 10^{-3}$



# Multiwavelength analysis



Compatible data  
at other wavelength?

Simultaneous  
Calibrated  
Specific Processing?  
Context?

# International Virtual Observatory Alliance

- ◆ Science driven
- ◆ Information technology developments for astronomical science
- ◆ Fosters interoperability
- ◆ Standardisation process a la W3C
  - ◆ Working groups
  - ◆ Technical coordination
  - ◆ Science priority committee
  - ◆ Executive Board representing all national projects
- ◆ Interop meeting every 6 months

[www.ivoa.net](http://www.ivoa.net)



**LEVEL 0**

**USERS**



**COMPUTERS**

**USER LAYER**

**USING**

F  
I  
N  
D  
I  
N  
G

VO  
CORE

G  
E  
T  
T  
I  
N  
G

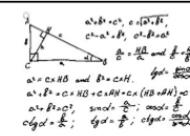
**SHARING**

**RESOURCE LAYER**

20101004  
IVOA Architecture



**PROVIDERS**

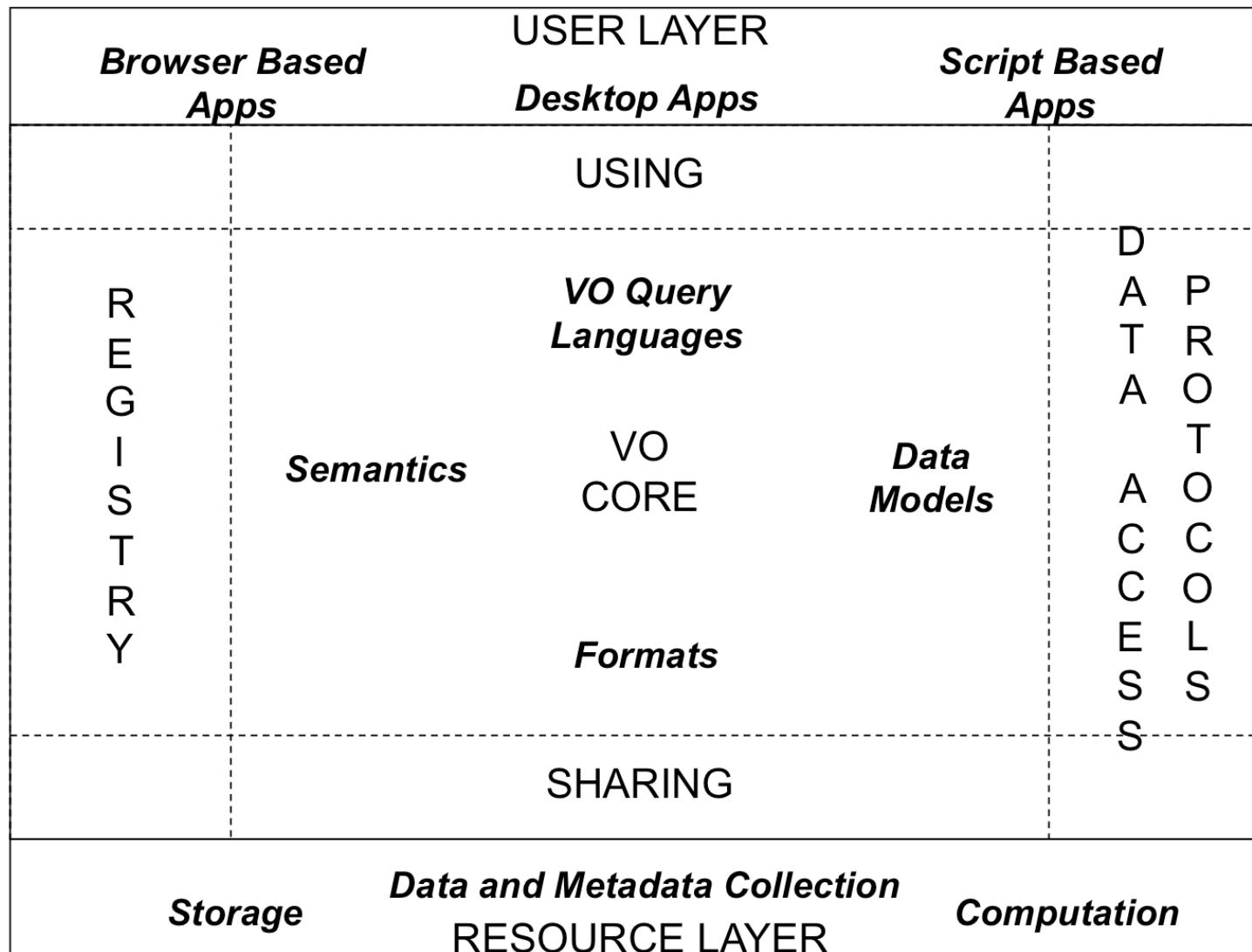


LEVEL 1

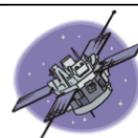
USERS



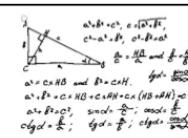
COMPUTERS



20101004  
IVOA Architecture



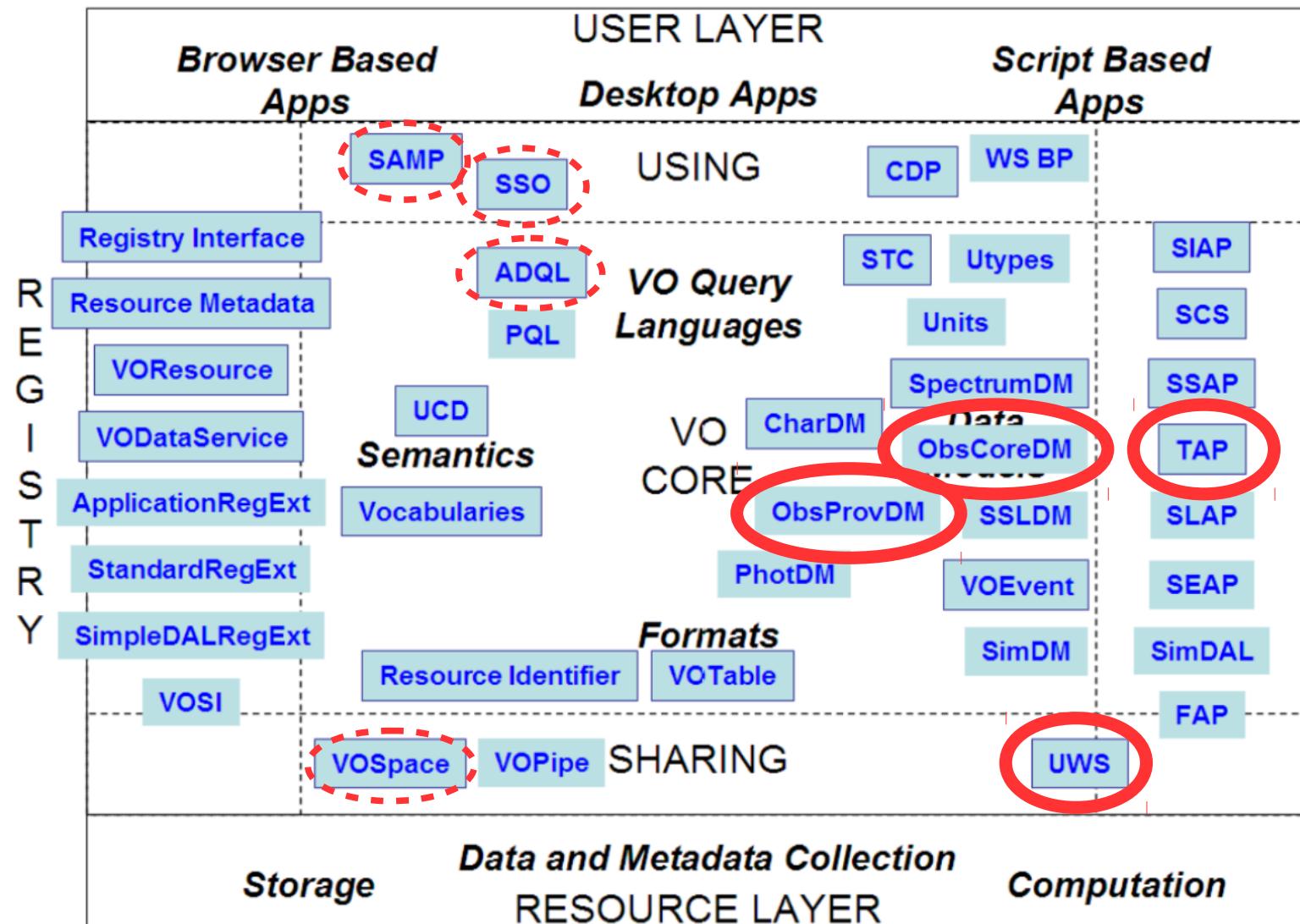
PROVIDERS



**LEVEL 2**  
**All standards**



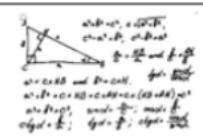
REC  
InProgress



20101004  
IVOA Architecture

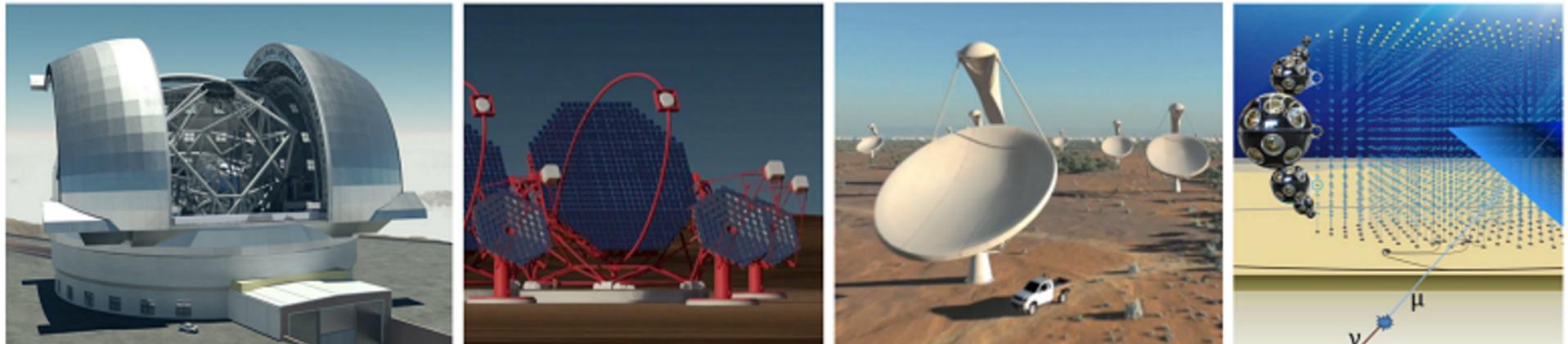


PROVIDERS





- ◆ European Commission's **Horizon 2020** framework
- ◆ ASTERICS aims to address the cross-cutting **synergies** and **common challenges** shared by the various Astronomy ESFRI facilities (E-ELT, CTA, SKA & KM3NeT)
- ◆ DADI: Data Access, Discovery and Interoperability



# VO data access prototype

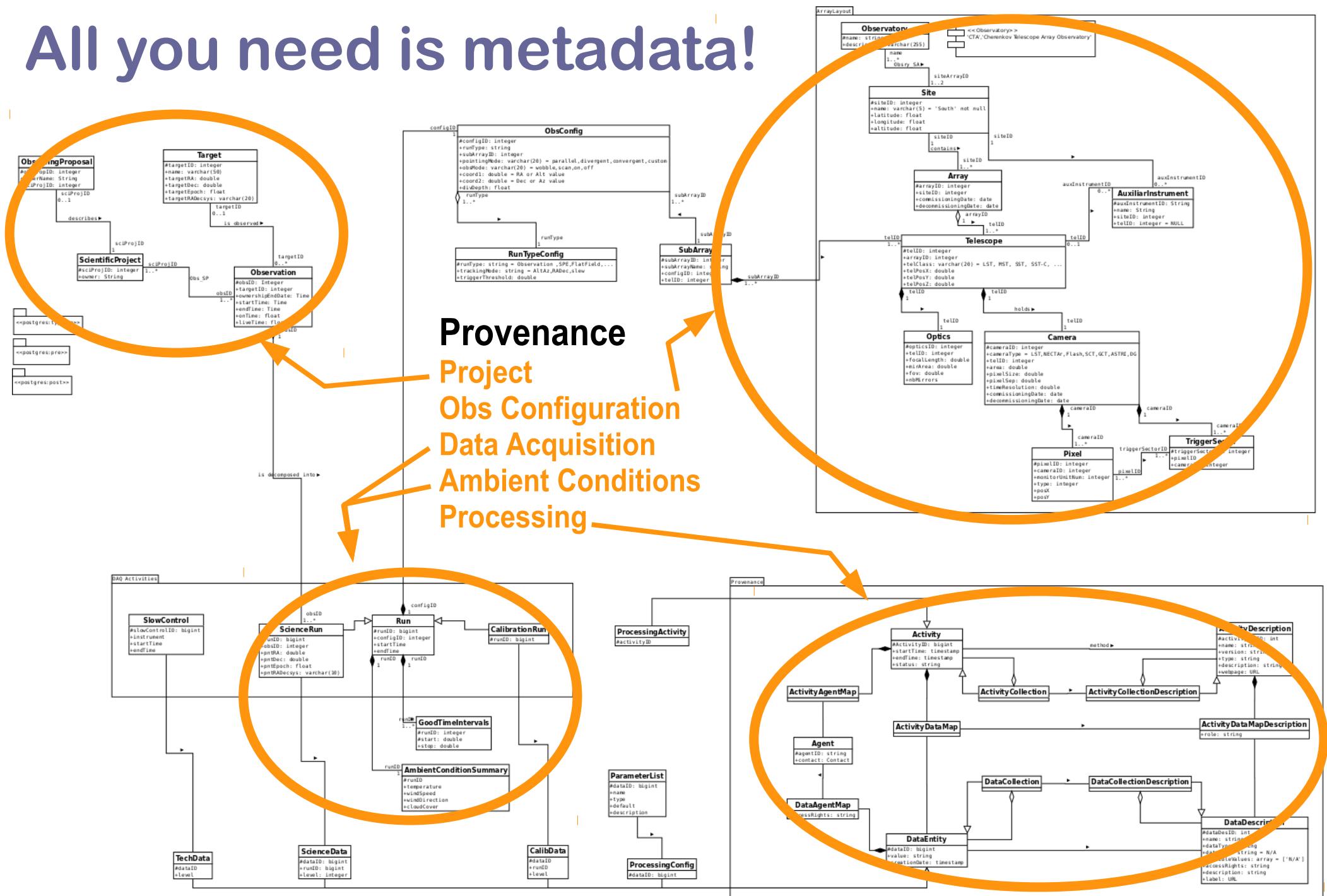
- ◆ **CTA Data Model** (not complete, still evolving)
  - ◆ [https://forge.in2p3.fr/projects/model/wiki/UML\\_models](https://forge.in2p3.fr/projects/model/wiki/UML_models)
  - ◆ Automatic Conversion UML to SQL
  - ◆ Relational database implemented (PostgreSQL)
- ◆ **Data Ingestion:** CTA First Data Challenge (1DC)
- ◆ **VO Compliance**
  - ◆ IVOA ObsCore Data Model
  - ◆ GAVO DaCHS server: [TAP](#), [ADQL](#)
- ◆ **Web Client** (Django, jQuery, BootStrap)
- ◆ **Online Analysis:** [UWS](#), [SAMP](#)
- ◆ **Single Sign On** solutions

} ObsTAP



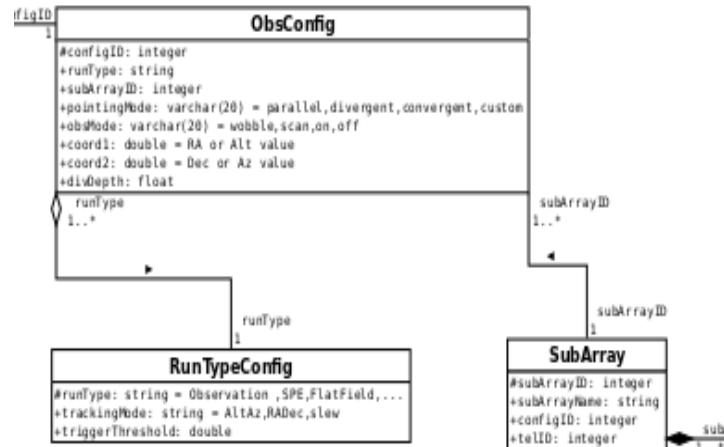
► Complete solution based on VO standards/protocols

# All you need is metadata!



# CTA data model

## ObsConfig



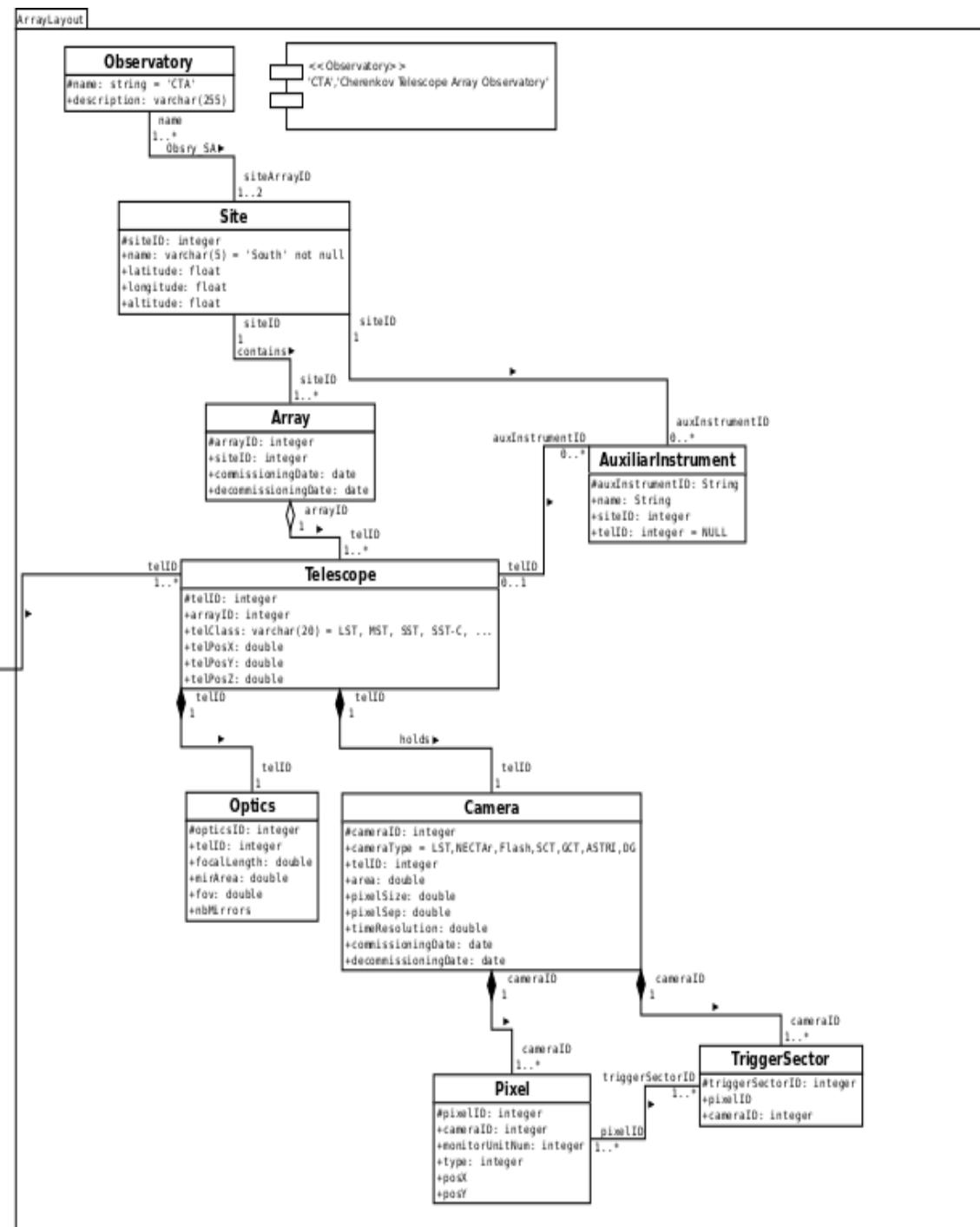
Examples :

Number of telescopes involved

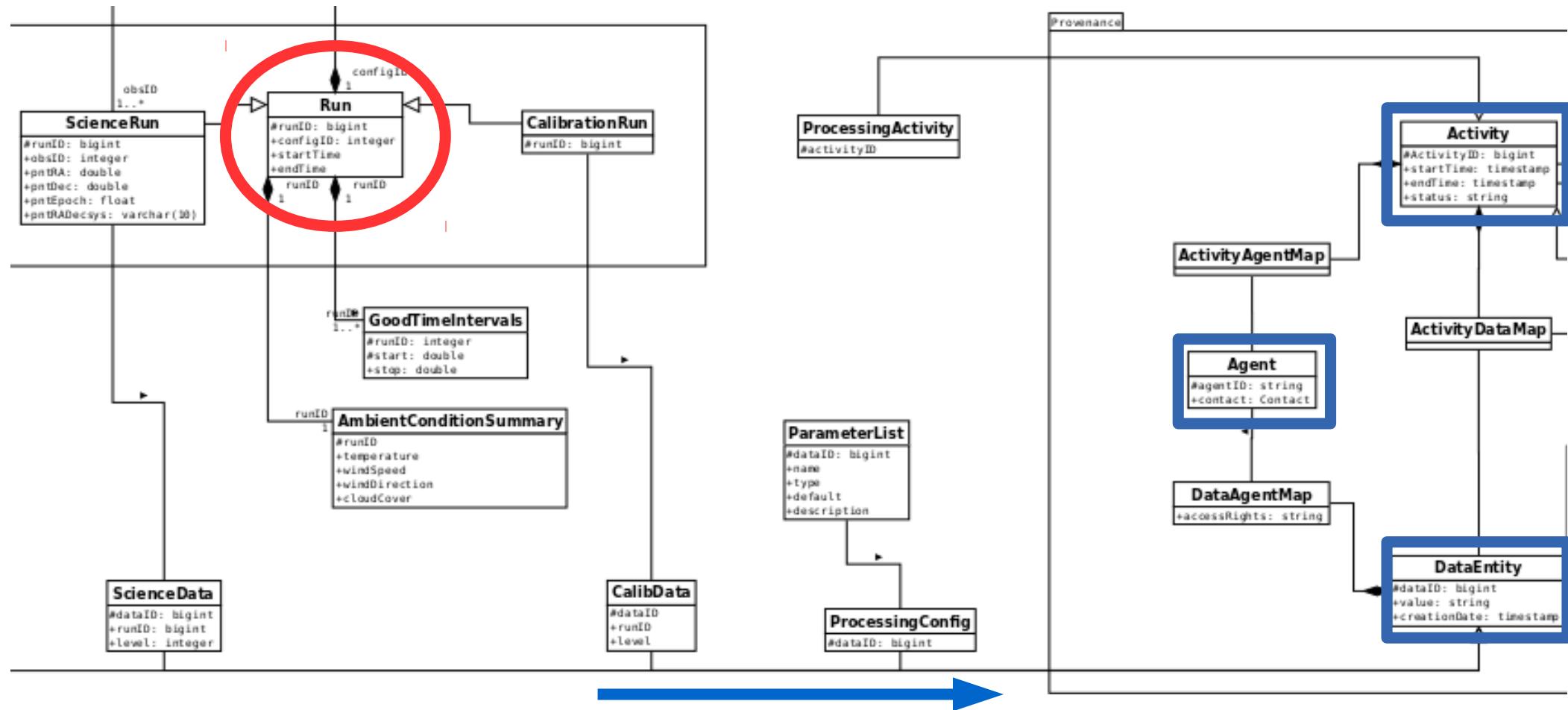
Field of view

Pointing direction

...

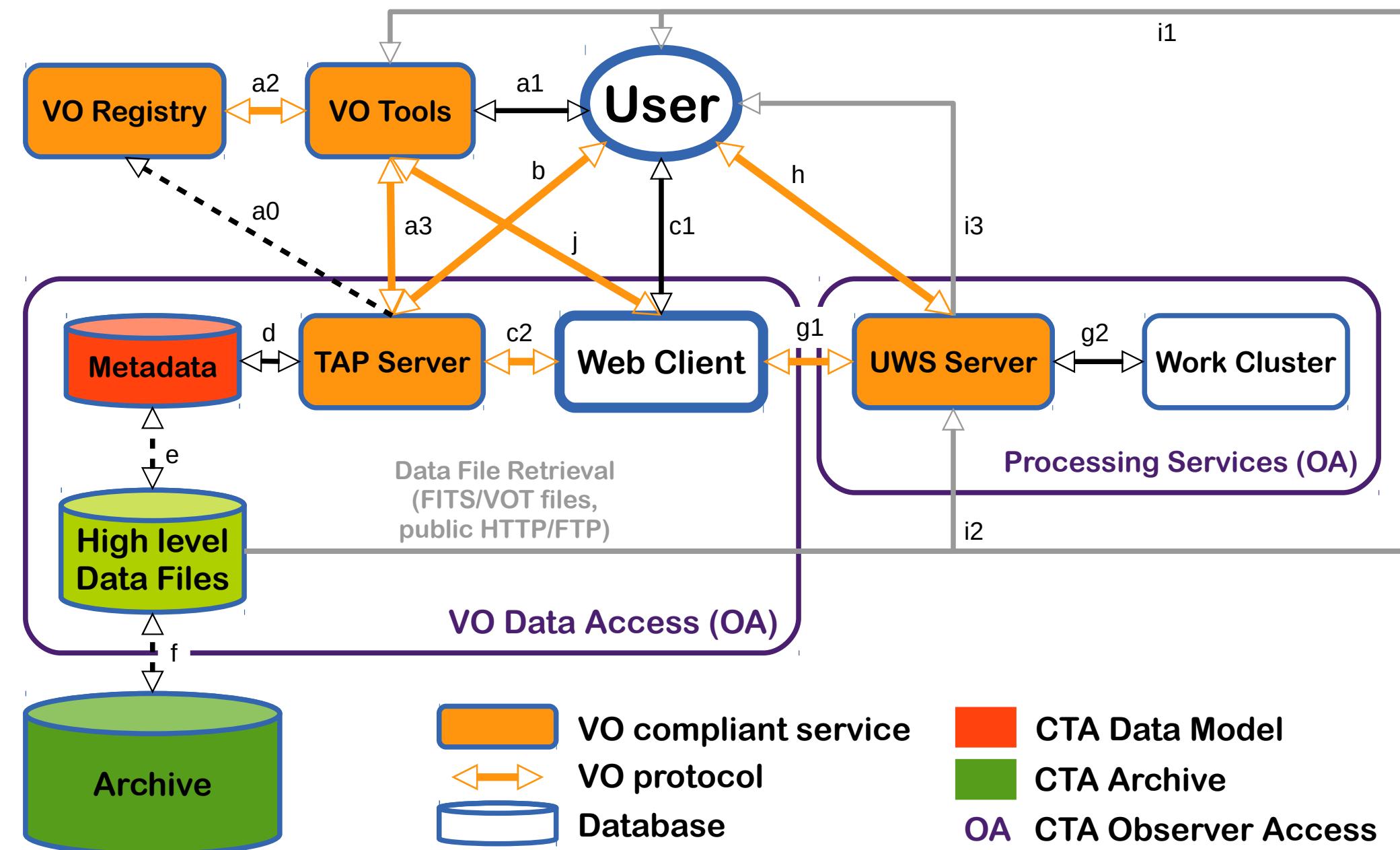


# CTA data model – Data Processing

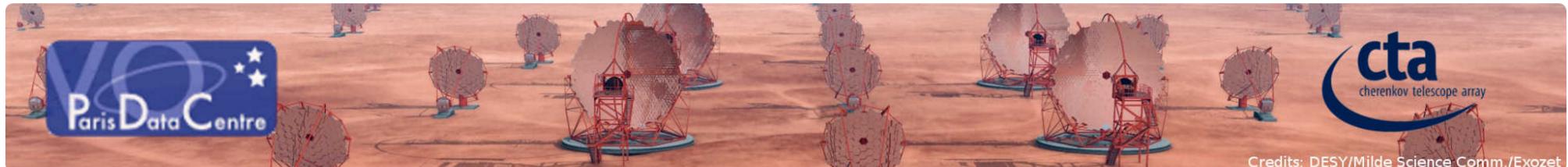


- ◆ Some **CTA specific** Activities and Entities (Run, ObsConfig, ...)
- ◆ Link to W3C/IVOA Provenance data model for Data Processing

# VO data diffusion prototype



# Web Client prototype <http://voparis-cta-client.observatoire-paris.fr>

[Monte Carlo simulations](#)[Data Distiller](#)[Data Reduction](#)[INAF CTA portal](#)[Mathieu.Servillat@obspm.fr](#) logout[CTA Data Distiller](#)[Search Form](#)[Job List](#)[Sign out Mathieu.Servillat@obspm.fr](#)

## Cone Search

**Target Name**

Used to query Simbad with Sesame and set RA/Dec.

**Source RA (deg)****Source Dec (deg)****Search radius (deg)**[Submit](#)[Reset](#)

- ◆ Django, jQuery, BootStrap3
- ◆ **Name resolver**  
Simbad through Sesame
- ◆ Builds and Sends the **ADQL query**

## ObsCore Search

**proposal\_id**  
Proposal ID

# Web Client prototype

<http://voparis-cta-client.ospdm.fr>

The screenshot shows the CTA Data Distiller web client prototype. At the top, there's a banner featuring the Paris Data Centre logo, several Cherenkov Telescope Array (CTA)望远镜 models, and the CTA logo. Below the banner, the navigation bar includes 'CTA Data Distiller', 'Search Datasets', 'Results' (with a checkmark icon), 'Job List', 'Selected Job', and 'JS9'. On the right, there's an 'Authentication' section with a 'Sign out user' link. The main area is divided into three sections: 'Search', 'Analyse', and 'Visualisation'. The 'Search' section contains an 'ADQL query' input field with the following content:

```
SELECT * FROM cta.vo_obscore as o WHERE 1 = intersects(o.s_region, circle('ICRS', 83.63308333, 22.0145, 0.001))
```

The 'Analyse' section shows a table of results with the following columns: dataproduct\_type, obs\_collection, obs\_id, target\_name, s\_ra (deg), and s\_dec (deg). The 'Visualisation' section is represented by the JS9 icon. To the right, there's a sidebar with several tool categories:

- SAMP**: Interop (SAMP) - Send Result Table, Send Selected Data.
- Analysis tools**: Create Count Map(s), Extract Spectrum.
- Plotting tools**: TOPCAT, Aladin, VOSpec, SPLAT.

Below the table, it says 'Showing 1 to 5 of 10 rows' and has a 'records per page' dropdown set to 5. At the bottom, there are navigation buttons for page 1 of 2.

dataproduct_type	obs_collection	obs_id	target_name	s_ra (deg)	s_dec (deg)
eventlist	1	23592	Crab Nebula	82.01333618164062	22.01444435119629
eventlist	1	23559	Crab Nebula	85.25333404541016	22.01444435119629
eventlist	1	23526	Crab Nebula	83.63333129882812	22.51444435119629
eventlist	1	23523	Crab Nebula	83.63333129882812	21.51444435119629
eventlist	3	5003499	CrabNebula	83.28087615966797	21.784133911132812

# IVOA ObsCore fields

## ObsCore mandatory fields:

- ◆ dataproduct\_type
- ◆ calib\_level
- ◆ target\_name
- ◆ obs\_collection / obs\_id / obs\_publisher\_did
- ◆ access\_url / access\_format / access\_estsize
- ◆ **Spatial**: s\_ra / s\_dec / s\_fov / s\_region / s\_resolution
- ◆ **Time**: t\_min / t\_max / t\_exptime / t\_resolution
- ◆ **Wavelength**: em\_min / em\_max / em\_res\_power
- ◆ facility\_name / instrument\_name

# ObsCore fields for CTA

**dataproduct\_type:** has to be one of the following: image, cube, spectrum, sed, timeseries, visibility, event.  
Set to "event" in the prototype, has it exposes the 1DC DL3 files.

**calib\_level:** one of the following integer values: 0 (instrumental or raw data in a non-standard/proprietary format), 1 (instrumental data in a standard format, e.g. FITS), 2 (calibrated data in standard format, with instrument signature removed), and 3 (more highly processed data product). CTA defines 5 data level, for example DL3 data are calibrated data in scientific units but still include an instrument signature, hence its calib\_level would be between 1 and 2.

**access\_url:** to be defined by the Archive, however the CTA 1DC data should not be accessible to the public. We thus include simulated data hosted on <http://voplus.obspm.fr/cta/> and always point to this URL in the prototype. In the VO context, the access URL is generally a public link. To handle data rights, this may point to a retrieval system with the ID of the requested data product.

**em\_min, em\_max:** The spectral coordinates are in TeV for us and should be converted to meters to follow the ObsCore standard. This could lead to precision issues in spectral data (though it is not an issue for discovery purposes).

**facility\_name:** we use the observatory name, e.g. "CTA".

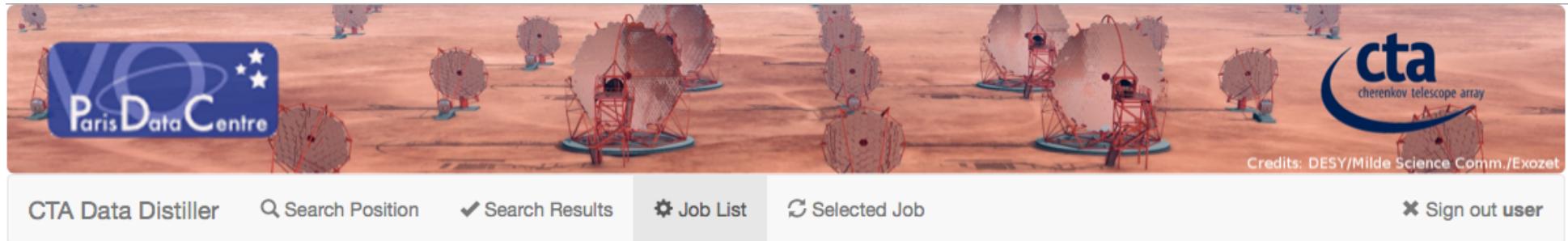
**instrument\_name:** As our test data comes from several experiments, we describe them here: HESS, MAGIC, VERITAS or CTOOLS (for simulated data with the ctools). This could be used to expose the CTA SubArray used to acquire the data?

# Extended ObsCore fields for CTA

- ◆ **Optional ObsCore fields:**
  - ◆ **dataproduct\_subtype**: show DL0-5?
  - ◆ **obs\_release\_date**
  - ◆ **data\_rights** (Public/Secure/Proprietary)
  - ◆ **s\_resolution min, s\_resolution max** (as it is dependent on energy)
  - ◆ **proposal\_id**
- ◆ **ObsConfig:**
  - ◆ **site**: North or South site.
  - ◆ **sub\_array\_name** (or directly in **instrument\_name**)
  - ◆ **pointing\_mode**: parallel, divergent, convergent, custom...
  - ◆ **obs\_mode**: wobble, scan, on, off
  - ◆ **run\_type**: flatfield, science, SPE...
- ◆ **Provenance:**
  - ◆ **data\_quality**: flag giving information on the data quality
  - ◆ **calib\_version**: version of the calibration stage of the Pipeline
  - ◆ **reco\_version**: version of the reconstruction stage of the Pipeline
  - ◆ **reco\_method**: reconstruction method used to obtain DL2 data
  - ◆ **applied\_cuts**: selection criteria used to obtain e.g. a DL3 photon event list
  - ◆ **spectral\_model**: spectral model assumed to obtain spectrum

# Web Client prototype

<http://voparis-cta-client.observatoire-paris.fr>



## Job List

Type	Start Time	Phase	Actions			Control		
ctbin	2014-10-07 21:32:58	ABORTED	<a href="#"><i>ⓘ</i> Details</a>	<a href="#"><i>📁</i> Edit</a>	<a href="#"><i>⬆️</i> Results</a>	<a href="#"><i>▶️</i> Start</a>	<a href="#"><i>🚫</i> Abort</a>	<a href="#"><i>✖️</i> Delete</a>
ctbin	2014-10-06 17:12:03	COMPLETED	<a href="#"><i>ⓘ</i> Details</a>	<a href="#"><i>📁</i> Edit</a>	<a href="#"><i>⬆️</i> Results</a>	<a href="#"><i>▶️</i> Start</a>	<a href="#"><i>🚫</i> Abort</a>	<a href="#"><i>✖️</i> Delete</a>
ctbin	2014-10-04 14:05:12	COMPLETED	<a href="#"><i>ⓘ</i> Details</a>	<a href="#"><i>📁</i> Edit</a>	<a href="#"><i>⬆️</i> Results</a>	<a href="#"><i>▶️</i> Start</a>	<a href="#"><i>🚫</i> Abort</a>	<a href="#"><i>✖️</i> Delete</a>
ctbin	2014-10-03 13:22:46	ABORTED	<a href="#"><i>ⓘ</i> Details</a>	<a href="#"><i>📁</i> Edit</a>	<a href="#"><i>⬆️</i> Results</a>	<a href="#"><i>▶️</i> Start</a>	<a href="#"><i>🚫</i> Abort</a>	<a href="#"><i>✖️</i> Delete</a>
ctbin		PENDING	<a href="#"><i>ⓘ</i> Details</a>	<a href="#"><i>📁</i> Edit</a>	<a href="#"><i>⬆️</i> Results</a>	<a href="#"><i>▶️</i> Start</a>	<a href="#"><i>🚫</i> Abort</a>	<a href="#"><i>✖️</i> Delete</a>

- ◆ **Asynchronous** management of processes
  - ◆ Job sent on a **work cluster**
  - ◆ JavaScript library developped at PADC

# Online processing (server)

## Main features

- ◆ **IVOA standard**
  - ◆ Universal Worker System (UWS)
- ◆ **REST architecture**
  - ◆ Python micro-framework: bottle.py
- ◆ **Collaborative development**
  - ◆ Git server at PADC (gitolite)
  - ◆ GitHub: [https://github.com/mservillat/uws\\_server](https://github.com/mservillat/uws_server)
- ◆ **Tests and quality**
  - ◆ Unit tests with unittest and webtest
  - ◆ Activity history with logging

## Prototype available

<https://voparis-uws-test.obspm.fr>

