

# Astrophysique des Hautes Energies dans l'Observatoire Virtuel

## *Virtual Observatory and High Energy Astrophysics*



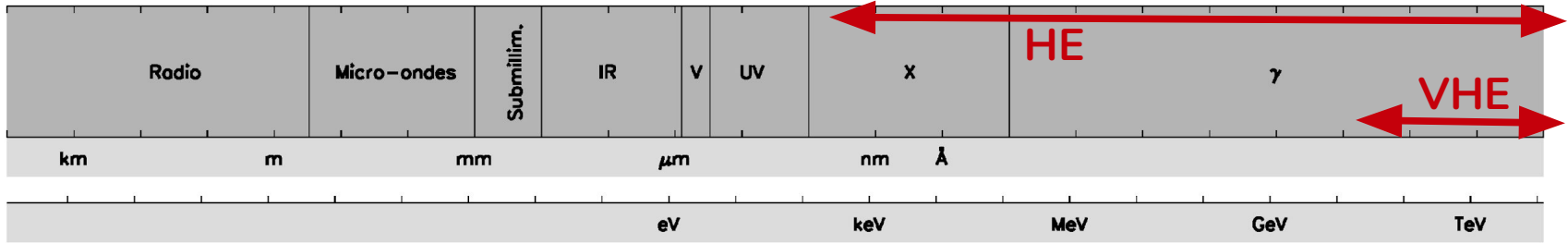
*ASOV Strasbourg meeting  
2022-10-11*

**Mathieu Servillat** (LUTH - Observatoire de Paris / CNRS)



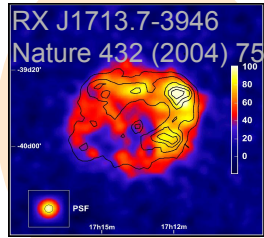
- ◆ Violent, transient, non-thermal phenomena
- ◆ Matter under extreme conditions
- ◆ Particle Acceleration
- ◆ Fundamental Physics
- ◆ Role of Black Holes in the structuration of the Universe

# High Energy Astrophysics

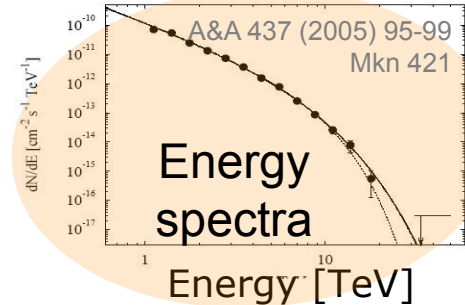
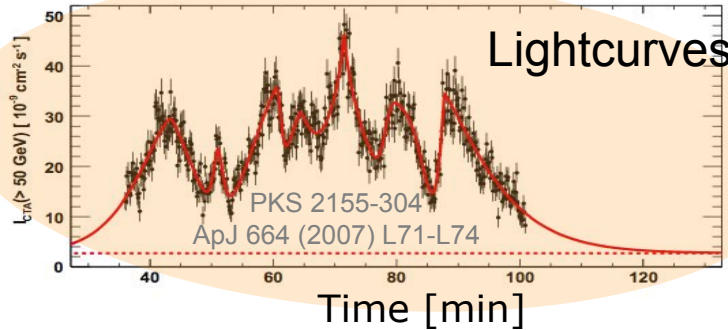


Several orders of magnitude - Photon **counting** - Low count **statistics** - High background

→ **Event lists** (coordinates, time, energy)



Images



+ **multi-messenger data** (photons, cosmic rays, neutrinos, gravitational waves...)

# Virtual Observatory and IVOA

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



The **Virtual Observatory (VO)** is the vision that astronomical datasets and other resources should work as a seamless whole.

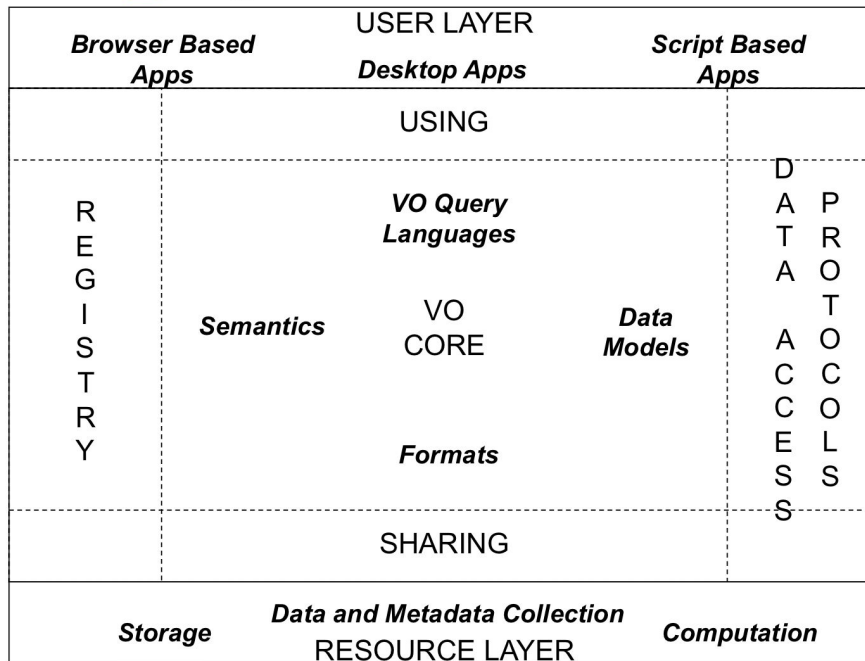
The **International Virtual Observatory Alliance (IVOA)** is an organisation that debates and agrees the technical standards that are needed to make the VO possible.

LEVEL 1

USERS



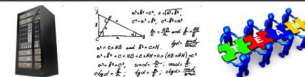
COMPUTERS



20101004  
IVOA Architecture



PROVIDERS



# An event in VO

<https://www.ivoa.net/documents/ObsCore>

**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.




# Cherenkov Astronomy



**MAGIC:** located in La Palma, Spain  
Since 2004: single 17m telescope  
Since 2009: system of two 17m telescopes



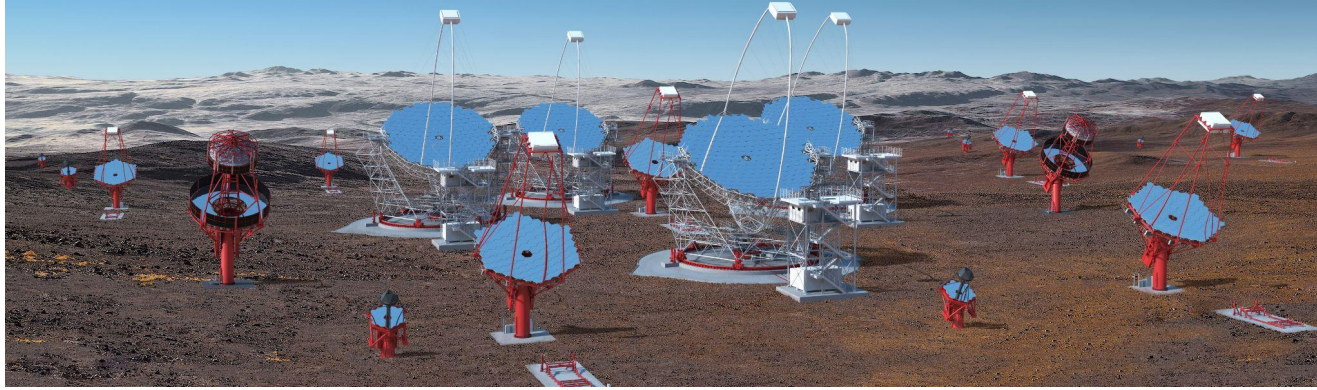
**VERITAS:** located in Mt Hopkins, Arizona  
Since 2007: four 12m telescopes  
Since 2012: upgraded PMTs



**H.E.S.S.:** located in Khomas Highlands, Namibia  
Since 2002: four 12m telescopes  
Since 2012: added 32m by 24m telescope  
Since 2015: camera upgrades on 12m telescopes

@ Jeff Grube

Atacama, Chile



La Palma, Spain



Alpha configuration

CTAO Southern array  
51 telescopes over a  $\sim 3$  km<sup>2</sup> area

14 Medium-Sized Telescopes (MST)  
37 Small-Sized Telescopes (SST)

CTAO Northern array  
13 telescopes distributed over an  
area of about 0.5 km<sup>2</sup>

4 Large-Sized Telescopes (LSTs)  
9 Medium-Sized Telescopes (MSTs)



### Low-energy section

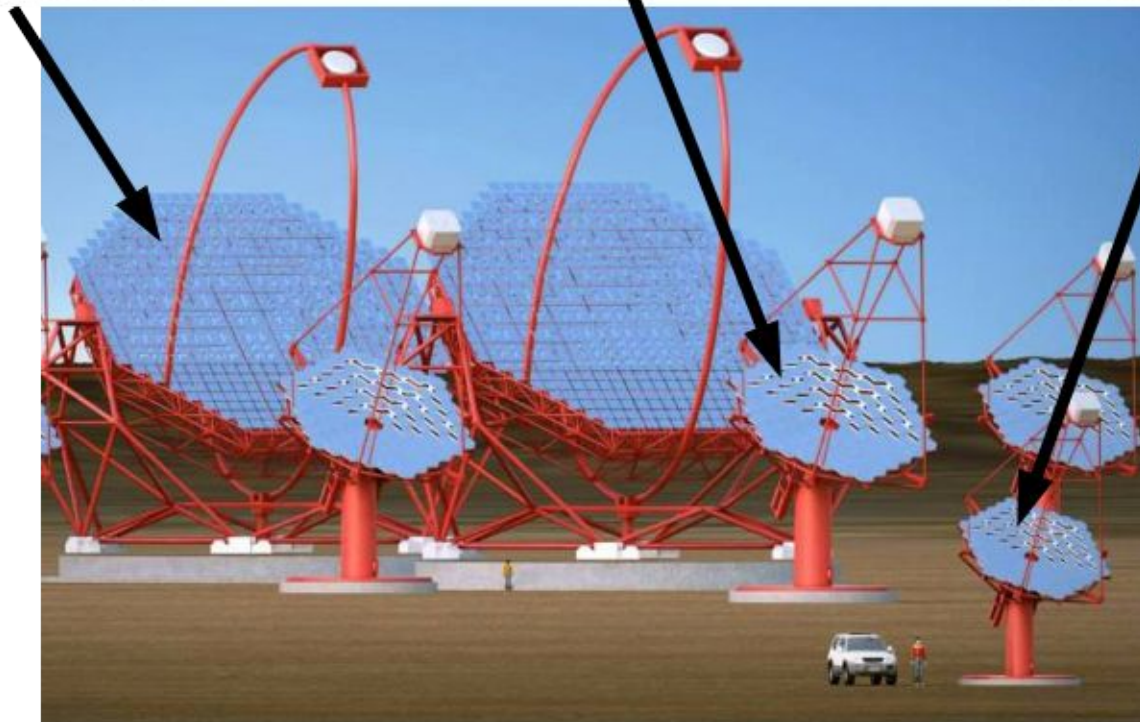
4 x 23m (**LST**)  
FoV: 4-5 degrees  
 $E > 10\text{s of GeV}$

### Core-energy section

23 x 12m (**MST**)  
FoV: 7-8 degrees  
 $100 \text{ GeV} < E < 10 \text{ TeV}$

### High-energy section

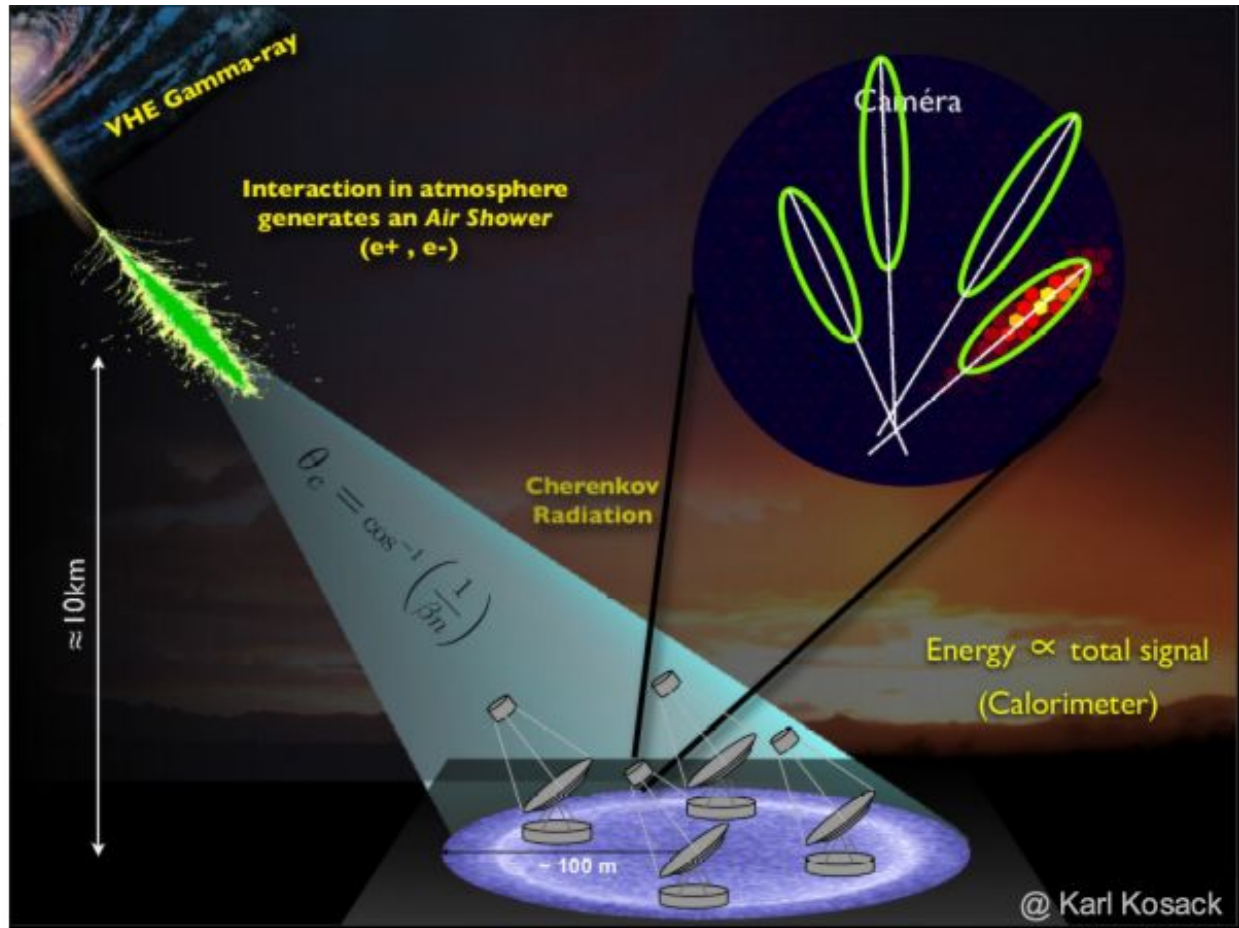
30-70 x 4-6m (**SST**)  
FoV: 10 degrees  
Multi-TeV energies





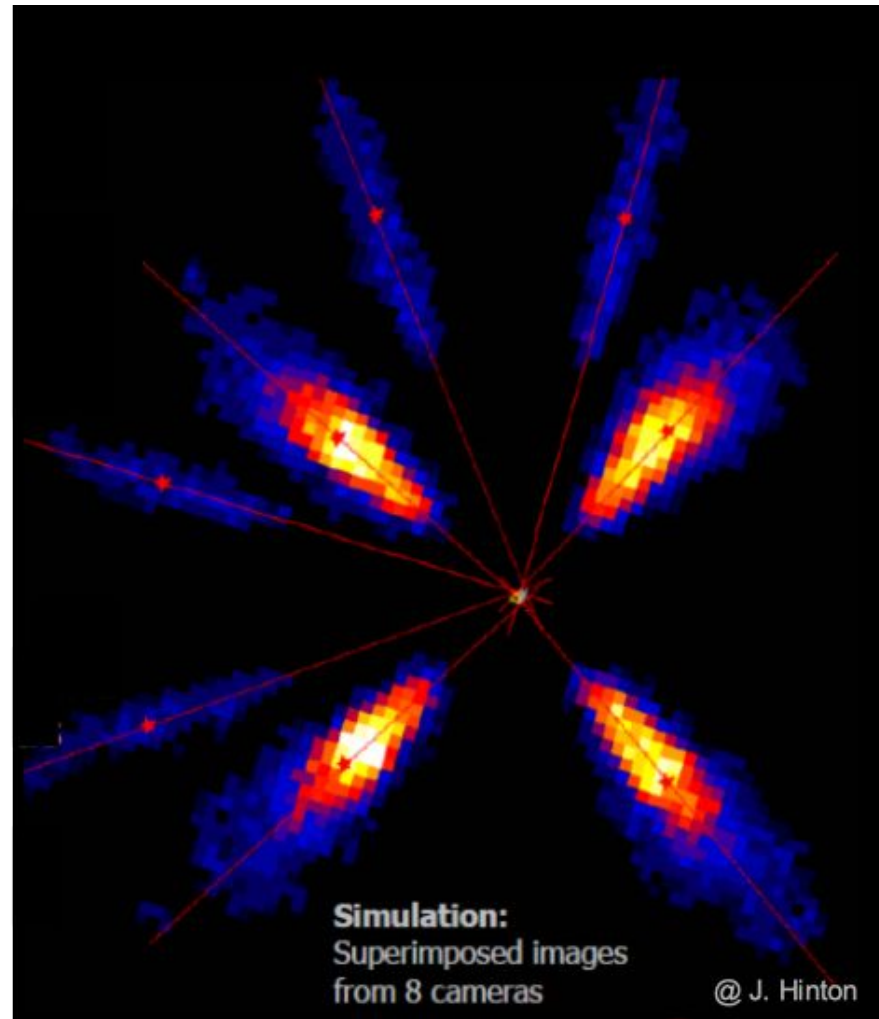
# Cherenkov Astronomy Principles

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

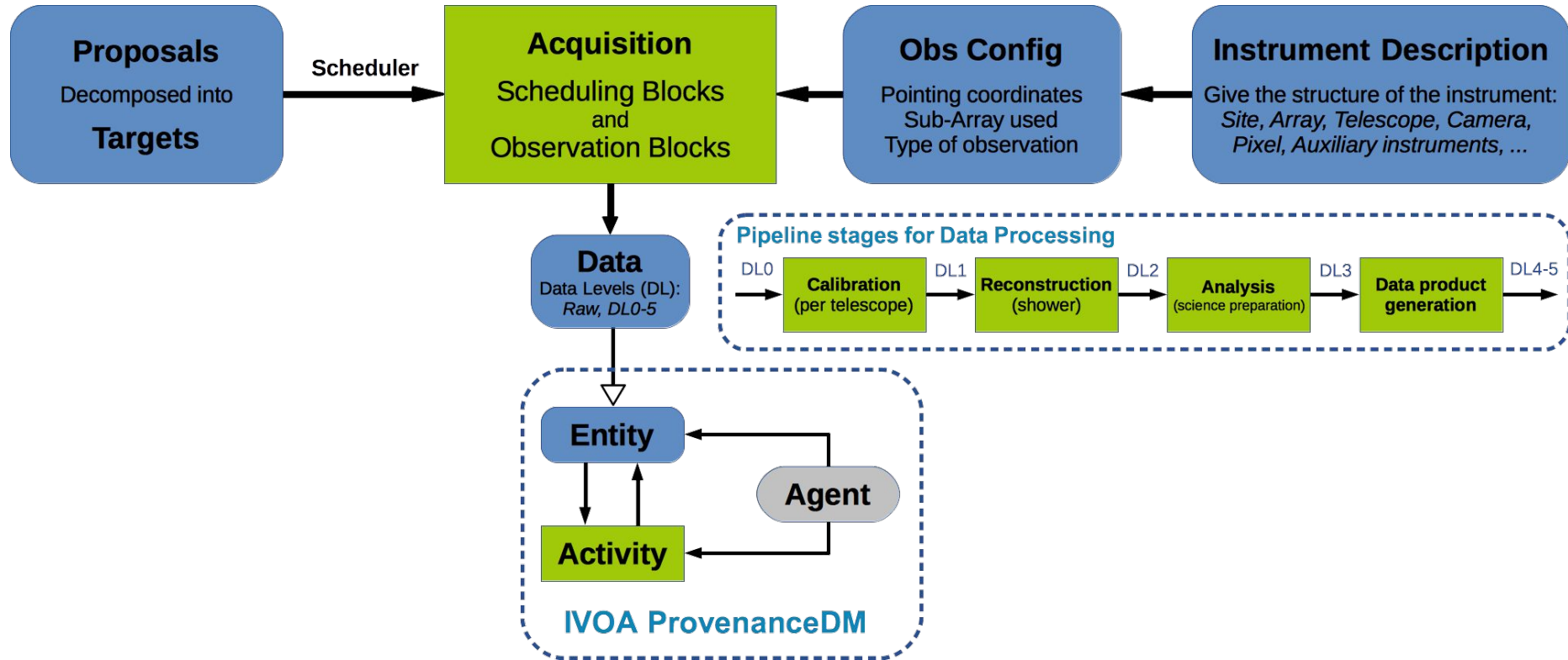




- ◆ **Larger** collection area for gamma-rays
  - ◆ **More events**, more photons
  - ◆ Better spectra, images, fainter sources
  
- ◆ **Better** events
  - ◆ More precise measurements of atmospheric cascades and hence primary gammas
  - ◆ Improved **angular** resolution
  - ◆ Improved **energy** resolution
  - ◆ Improved **background rejection** power

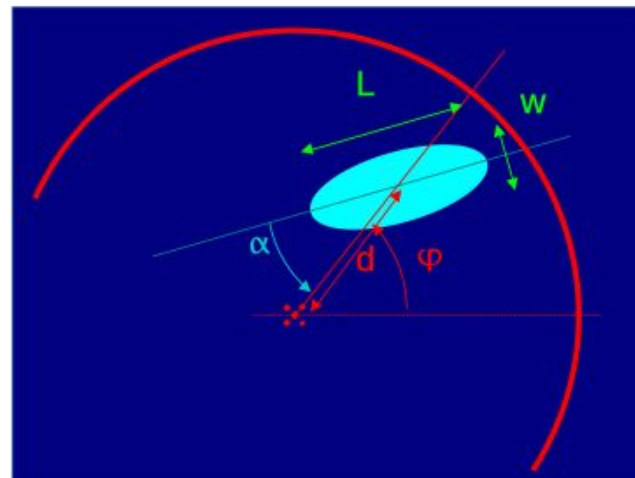
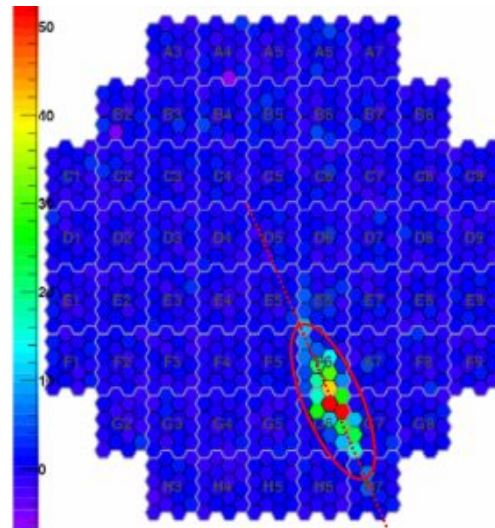


# CTA Master Configuration Data Model



# Reconstruction phase

- ◆ **Hillas Parameters** (1984)
  - ◆ Shower images are **elliptical**
  - ◆ A **few** parameters:
    - ◆ Length ( $L$ ) & Width ( $W$ )
    - ◆ Nominal Distance ( $d$ )
    - ◆ Azimuthal angle ( $\varphi$ )
    - ◆ Orientation angle ( $\alpha$ )
    - ◆ Amplitude (size)
  - ◆ Additional parameters: asymmetry, ...
- ◆ **Advanced methods**
  - ◆ 3D model
  - ◆ Fit to simulated images







# FAIR Principles for data

<https://www.go-fair.org/fair-principles>

## FINDABLE

Unique identifiers and metadata are used to allow data to be located quickly and efficiently



## ACCESSIBLE

Data is open, free and universally available for research discovery efforts



## INTER-OPERABLE

A common programming language is used to allow use in a broad range of applications



## REUSABLE

All data is clearly described and outlines associated data-use standards



# VO access to H.E.S.S. public data

Virtual Observatory (VO) standards tailored to make data findable:

- H.E.S.S. experiment first public data release  
→ <https://zenodo.org/record/1421099>
- **VO ObsTAP Service** implemented at the Observatoire de Paris:
  - **IVOA** Observation Data Model Core Components (**ObsCore**) [\[link to IVOA REC\]](#)  
→ adapted to Cherenkov data
  - **IVOA** Table Access Protocole (**TAP**) [\[link to IVOA REC\]](#)
  - Registered to the **VO Registry** via PADC (Paris Astronomical Data Centre)
- Data widely findable
  - e.g. Aladin, TOPCAT, TapHandle, PyVO...
  - + dedicated web pages



<https://hess-dr.obspm.fr>

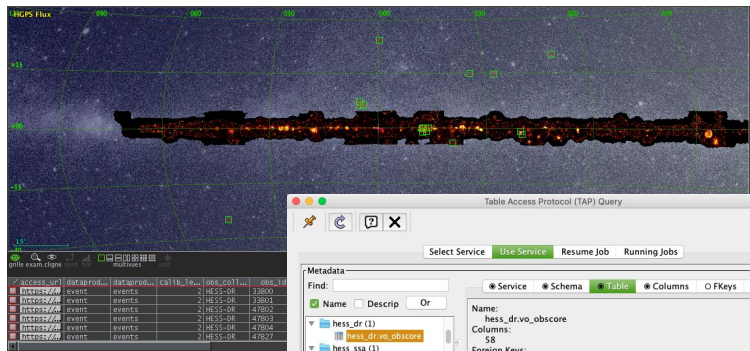
TapHandle

hess

voparis-hess-dr [H.E.S.S. DL3 public test data release 1]  
http://voparis-tap-astro.obspm.fr/\_system\_/tap/run/tap  
ivo://vopdc.obspm/luth/hess-dr

Tap Nodes

- voparis-hess-dr
- tap\_schema



Aladin

Table Access Protocol (TAP) Query

Select Service Use Service Resume Job Running Jobs

Metadata

Find:

Service:  Schema:  Table:  Columns:  FKKeys:  Hints:

Name: hess\_dr\_vo\_obscore

Columns: 58

Foreign Keys: 0

Description: IVOA ObsCore description of event lists for observations in the H.E.S.S. DL3 public test data release 1

Service Capabilities

Query Language: ADQL-2.0 Max Rows: 20000 (default) Uploads: 20Mb

ADQL Text

Mode: Synchronous

SELECT TOP 1000 \* FROM hess\_dr\_vo\_obscore

Examples: Basic 1/6: Full table

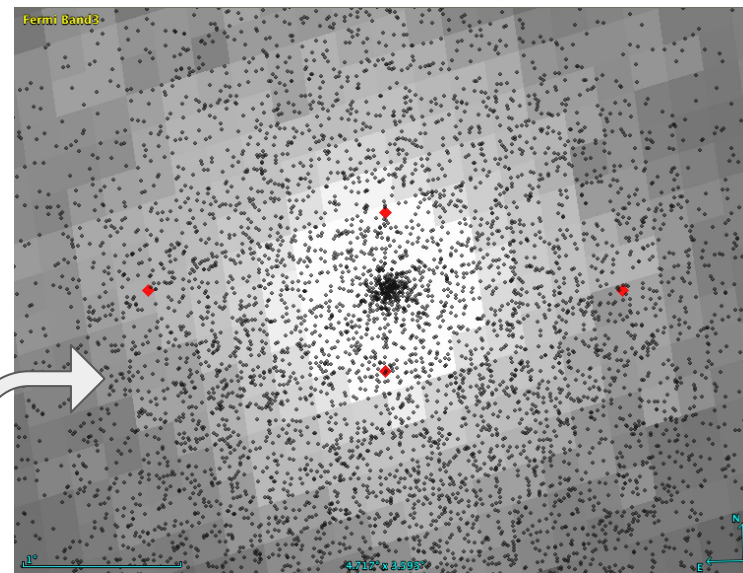
Run Query

TOPCAT

# Obtaining the data

- IVOA ObsCore `access_url`
  - Direct download link to the FITS file
  - IVOA DataLink (to be implemented):
    - Associated calibration data
    - Service that packages the requested **collection** of data + calibration for analysis
- Access rights
  - Public data: no restrictions
  - Anticipating need for permissions:
    - PI proprietary period
    - Federation authentication
    - e.g. IAM ESCAPE service  
→ <https://indigo-iam.github.io/escape-docs>

DL3 event list directly opened in Aladin  
(each black dot is an event)



	access_url	data...	obs_coll...	obs_id	access_format	access...	target_name	s_ra	s_de
	<a href="https://...">https://...</a>	event	HESS-DR	23523	application/fits	285120	Crab Nebula	83.6333...	21.514
	<a href="https://...">https://...</a>	event	HESS-DR	23526	application/fits	282240	Crab Nebula	83.6333...	22.514
	<a href="https://...">https://...</a>	event	HESS-DR	23559	application/fits	285120	Crab Nebula	85.2533...	22.014
	<a href="https://...">https://...</a>	event	HESS-DR	23592	application/fits	273600	Crab Nebula	82.0133...	22.014



# Common data format and FAIR software

<https://gamma-astro-data-formats.readthedocs.io>

A community initiative to define common DL3 data formats for gamma-ray astronomy based on FITS

- Work and discussions in progress !
- Includes formats for: event lists, effective area, energy resolution, point spread function, instrumental background
- A prototype data format for CTA
- Used for the H.E.S.S. public data release
- Partially in use by current instruments: Fermi-LAT, HESS, VERITAS, MAGIC, FACT, ...

[\[A&A 625, A10, 2019\]](#)

[\[A&A 632, A72, 2019\]](#)

[\[A&A 632, A102, 2019\]](#)



A Python package for **gamma-ray** astronomy

→ <https://gammapy.org>

- Open-source Python package (Astropy affiliated package)
- Core library for the Science Tools of CTA
- Used in the analysis of existing gamma-ray instruments, such as H.E.S.S., MAGIC, VERITAS, HAWC...
- **FAIR4RS**: FAIR Principles for Research Software → <https://doi.org/10.15497/RDA00065>
- **ESCAPE OSSR**: open-access repository to share scientific software and services → <https://escape2020.pages.in2p3.fr/wp3/ossr-pages>

# From F-A-I to FAIR

→ **ADASS XXXI** <https://hal.archives-ouvertes.fr/obspm-03516688v1>  
“FAIR high level data for Cherenkov astronomy”

**Findable**  
**Accessible**  
**Interoperable**  
**Reusable?**

- **Findable-Accessible-Interoperable**

- Use the **Virtual Observatory standards**, protocols and services
- Define community **standards** where required
- To be discussed early in projects, but **technical solutions exist**

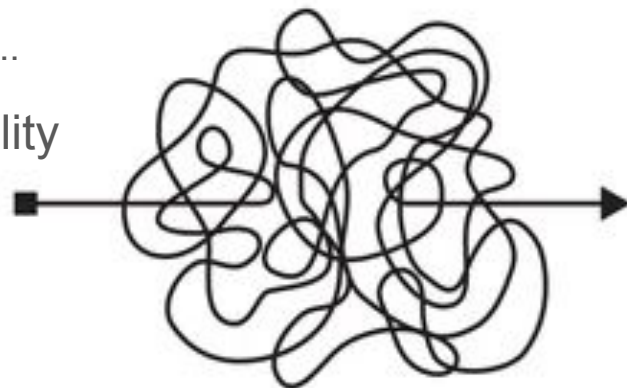
- **Reusability?**

- Based on the **quality / reliability / trustworthiness** of the products
- What calibration was applied? What tools were used and how?  
What assumptions were made during the data preparation?
- **Sustainability**: with time, key information may disappear...



- **Provenance** information as an answer to reusability

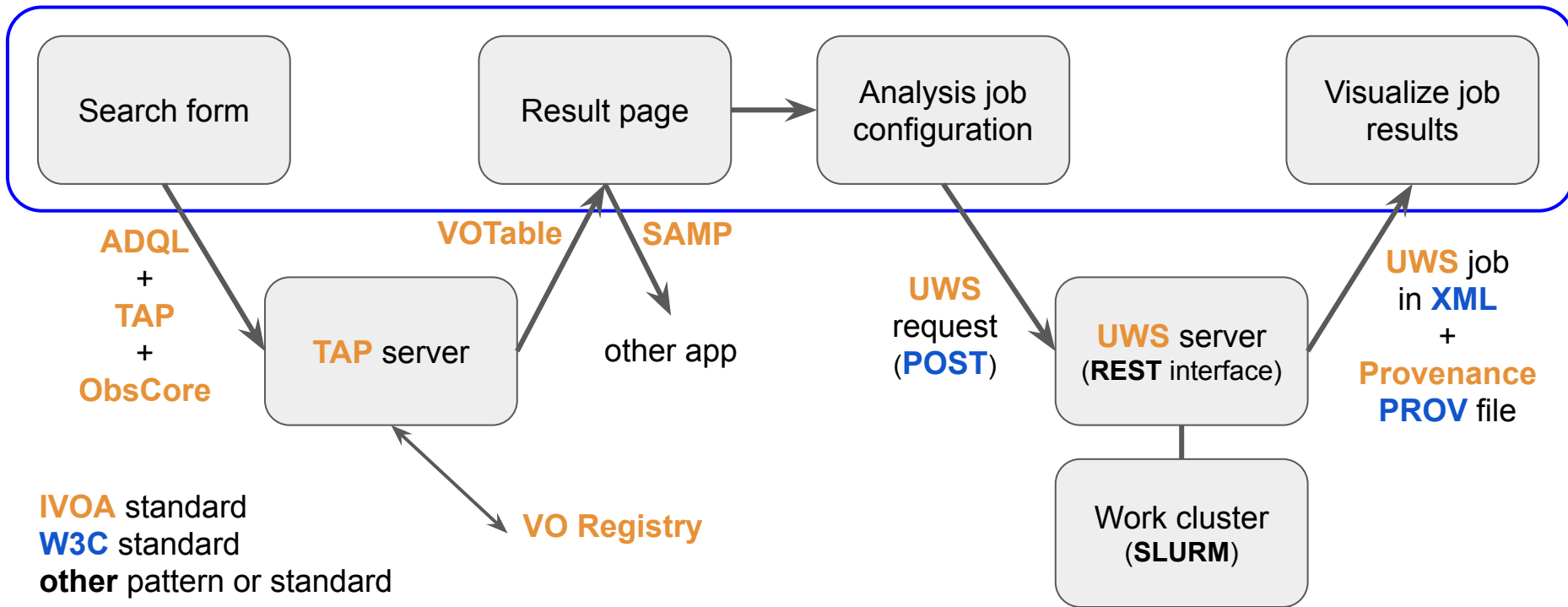
- Need for the **origin, trace**, and detailed manipulations
- Need to **structure** this information
- Need to **keep** it and **link** it to the data
- IVOA Provenance data model!



# Use of standards and modularity to increase development speed

Web application

+ Authn&Authz (SAML, OpenID, SCIM, ...)



TOPCAT

Table List

- 4: hess\_dl3\_dr1\_obs\_id\_020343.fits.gz-3
- 5: hess\_dl3\_dr1\_obs\_id\_020343.fits.gz-4
- 6: hess\_dl3\_dr1\_obs\_id\_020343.fits.gz-5
- 7: hess\_dl3\_dr1\_obs\_id\_023523.fits.gz
- 8: hess\_dl3\_dr1\_obs\_id\_023523.fits.gz-2
- 9: hess\_dl3\_dr1\_obs\_id\_023523.fits.gz-3
- 10: hess\_dl3\_dr1\_obs\_id\_023523.fits.gz-4
- 11: hess\_dl3\_dr1\_obs\_id\_023523.fits.gz-5

Current Table Properties

Label: hess\_dl3\_dr1\_obs\_id\_023523.fits.gz  
 Location: https://hess-dr.obspm.fr/retrieve/hess\_  
 Name: EVENTS-1  
 Rows: 7,613  
 Columns: 5

SAMP  
 Messages:  Clients:

109 / 1821 M

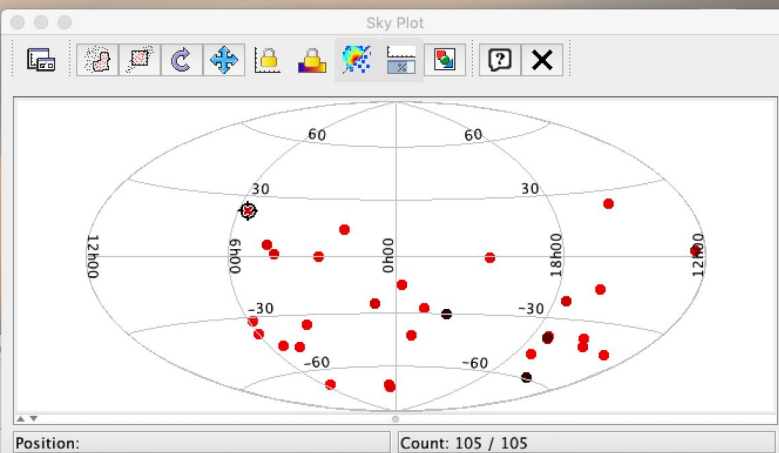


Table Access Protocol

TOPCAT(1): Table Browser for 1: TAP\_2\_hess\_dr.vo\_obscore

datapro...	datapro...	calib_le...	obs_colle...	obs_id	obs_...	
47	event	events	2	HESS-DR	22593	ivo:/
48	event	events	2	HESS-DR	22997	ivo:/
49	event	events	2	HESS-DR	23040	ivo:/
50	event	events	2	HESS-DR	23077	ivo:/
51	event	events	2	HESS-DR	23143	ivo:/
52	event	events	2	HESS-DR	23246	ivo:/
53	event	events	2	HESS-DR	23523	ivo:/
54	event	events	2	HESS-DR	23526	ivo:/
55	event	events	2	HESS-DR	23559	ivo:/
56	event	events	2	HESS-DR	23573	ivo:/
57	event	events	2	HESS-DR	23592	ivo:/
58	event	events	2	HESS-DR	23635	ivo:/
59	event	events	2	HESS-DR	23651	ivo:/
60	event	events	2	HESS-DR	23736	ivo:/
61	event	events	2	HESS-DR	25345	ivo:/

Metadata

Find: hess

Name  Descrip  Or

HESS DL3 DR1 (1/18)

- hess\_dr (1/1)
  - hess\_dr.vo\_obscore

Service Capabilities

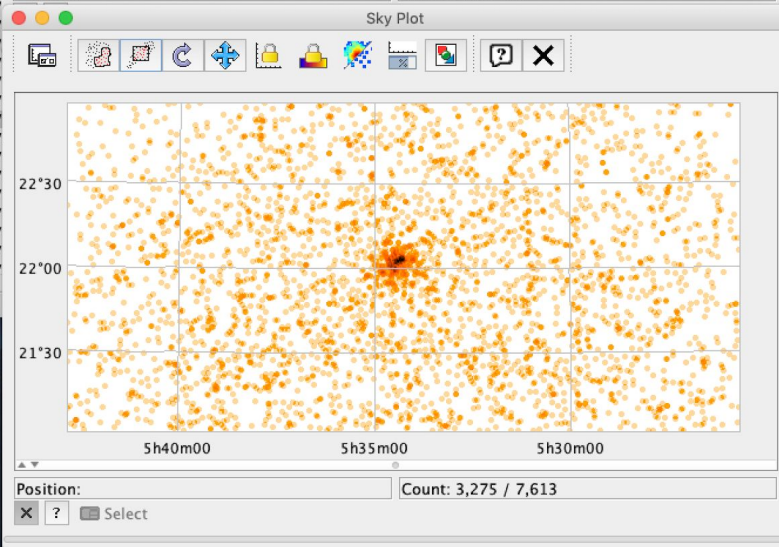
Query Language: ADQL-2.0 Max Rows: 200

ADQL Text

Mode: Synchronous

SELECT TOP 1000 \* FROM hess\_dr.vo\_obscore

Run Query





Use case	Description
<b>Cone Search</b>	Search data available for a given Target
<b>ObsCore search</b>	<p>Search data available corresponding to ObsCore keywords (target_name, time interval, ...), e.g.:</p> <ul style="list-style-type: none"> <li>search data for a given target at a given time</li> <li>search data in a given region of the sky</li> <li>search data that contain events at energy higher than 50 TeV</li> </ul>
<b>ObsCore optional search</b>	<p>Search data available corresponding to ObsCore optional keywords (target_class, data_rights, ...), e.g.:</p> <ul style="list-style-type: none"> <li>search public data for all blazars</li> <li>search data for a given proposal_id</li> </ul>
<b>ObsConfig search</b>	<p>Search data available corresponding to ObsConfig keywords (sub_array_name, pointing_mode, obs_mode ...), e.g.:</p> <ul style="list-style-type: none"> <li>search data that include the Large Size Telescopes (LSTs)</li> <li>search data for a given target, that do not include the divergent pointing mode</li> </ul>
<b>Provenance search</b>	<p>Search data available corresponding to Provenance keywords (calib_version, creation_date ...), e.g.:</p> <ul style="list-style-type: none"> <li>search data produced by a given version of the pipeline and for a given target</li> <li>search data produced using a given reconstruction method</li> <li>search data for a given target produced with loose cuts</li> </ul>

**datapoint\_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 use to expose the CTA SubArray used to acquire the data?

## ◆ **Optional ObsCore fields:**

- ◆ **dataprodct\_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