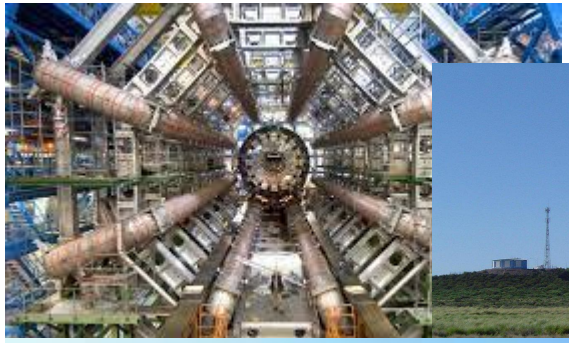
The Virtual Observatory in the age of multimessenger Astroparticle Physics

Claudio Galelli - LUTh, Observatoire de Paris
20/06/2023, 13th CRIS-MAC, Trapani

The “closed” approach

ATLAS

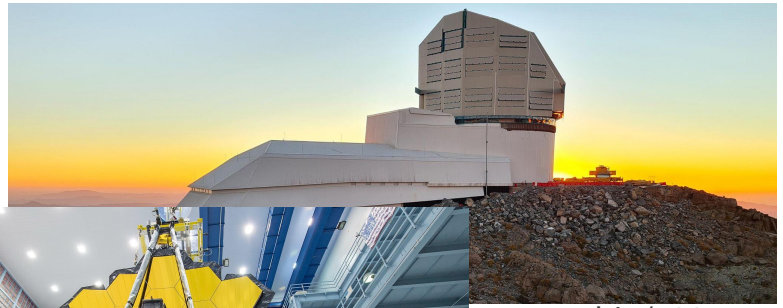


Auger

HESS

Operation decided by the collaborations
Data released as collaboration decision
Data released fractionally

The “open” approach



Rubin/LSST

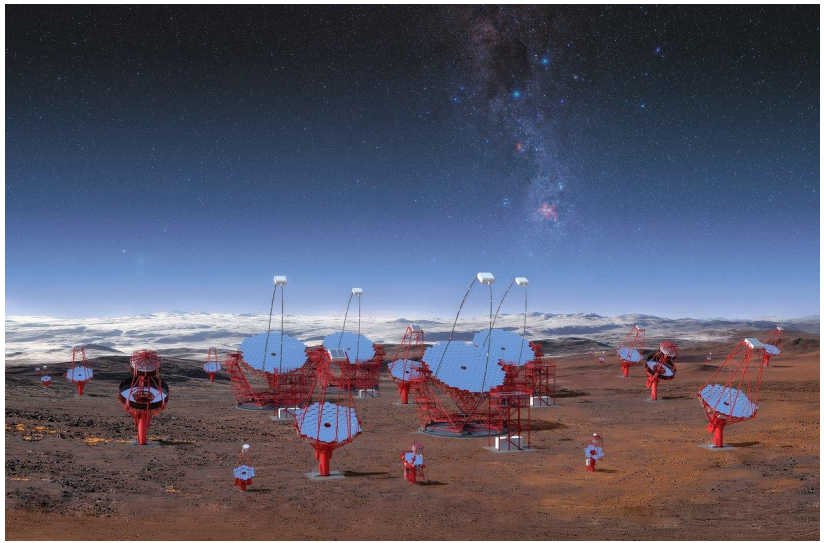
Webb

ALMA

Open calls for observation proposals
Data private for a set time (“proprietary”)
All data open and available to the public

Why is CTAO important for open data?

CTAO will be the first instrument in the VHE environment to operate as an **observatory**



CTA south at Paranal (simulation)

Proposal based - data will be open after a proprietary period

All the **data will be made public** with its **analysis software**

VODF very-high-energy open data format

gammapy A Python package for gamma-ray astronomy

Not as easy as it looks!

Can’t we just format our data as .txt files with incomprehensible keywords for column names and put it in an obscure archive page in a website?

Well yes, but it would be very nice if you could be

Data and metadata should be **easy to find for humans and computers**

Once the data is found, it should be clear how it is **accessed, including authentication**

The data should be easy to **integrate with other data** and analysis and storage workflows

The **reuse of data** should be **optimized** through data and metadata description

M. Wilkinson et al., *The FAIR Guiding Principles for scientific data management and stewardship,* Scientific Data, 2016
M. Servillat et al., *FAIR high level data for Cherenkov astronomy,* Proceedings of the XXXI ADASS conference, Cape Town, 2021

FAIR data in astrophysics

Making open data FAIR can be a nightmare - luckily in astrophysics we have a set of guidelines to help us, and we have to thank the **International Virtual Observatory Alliance**

The IVOA is the body that governs the standards on which the **Virtual Observatory (VO)** is built

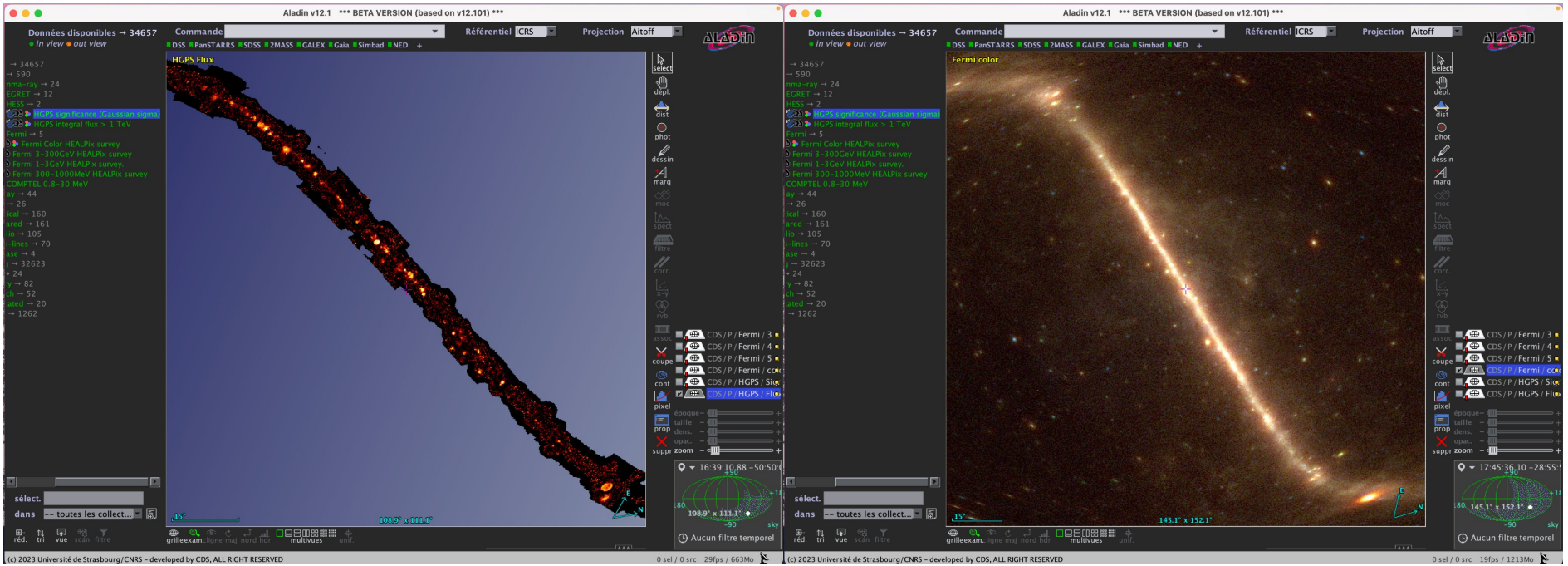
A collection of data centers containing archives and software tools for astronomical research

A good example are the platforms from the Strasbourg Data Center → Data is easy to browse, download and plot together in-app

D.Muna, *NASA SMD Information Policy: Let's All Be FAIR,* Proceedings of the XXXIII ADASS conference, Tucson, 2023

VHE astrophysics and the Virtual Observatory

Some gamma-ray data has already made its way to the VO tools!



These are point contributions! Showcase example: HESS public releases on VO

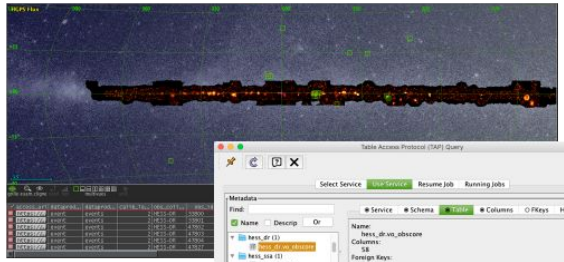
F = Findable

International Virtual Observatory Alliance (IVOA) standards tailored to make data findable:

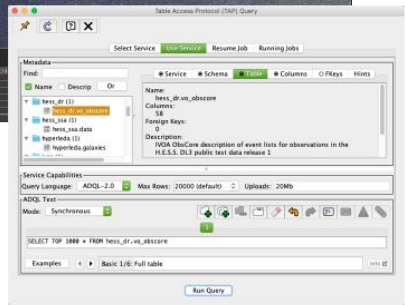
- **IVOA** Observation Data Model Core Components (**ObsCore**)
  [link to IVOA REC] → adapted to Cherenkov data
- **IVOA** Table Access Protocol (**TAP**)
  [link to IVOA REC]
- **Deployed service** at ObsParis
  https://hess-dr.obspm.fr
- Registered to the VO **Registry** via PADC
  (Paris Astronomical Data Centre)

→ Data widely findable
e.g. Aladin, TOPCAT, TapHandle, PyVO…
+ dedicated web pages

TapHandle



Aladin



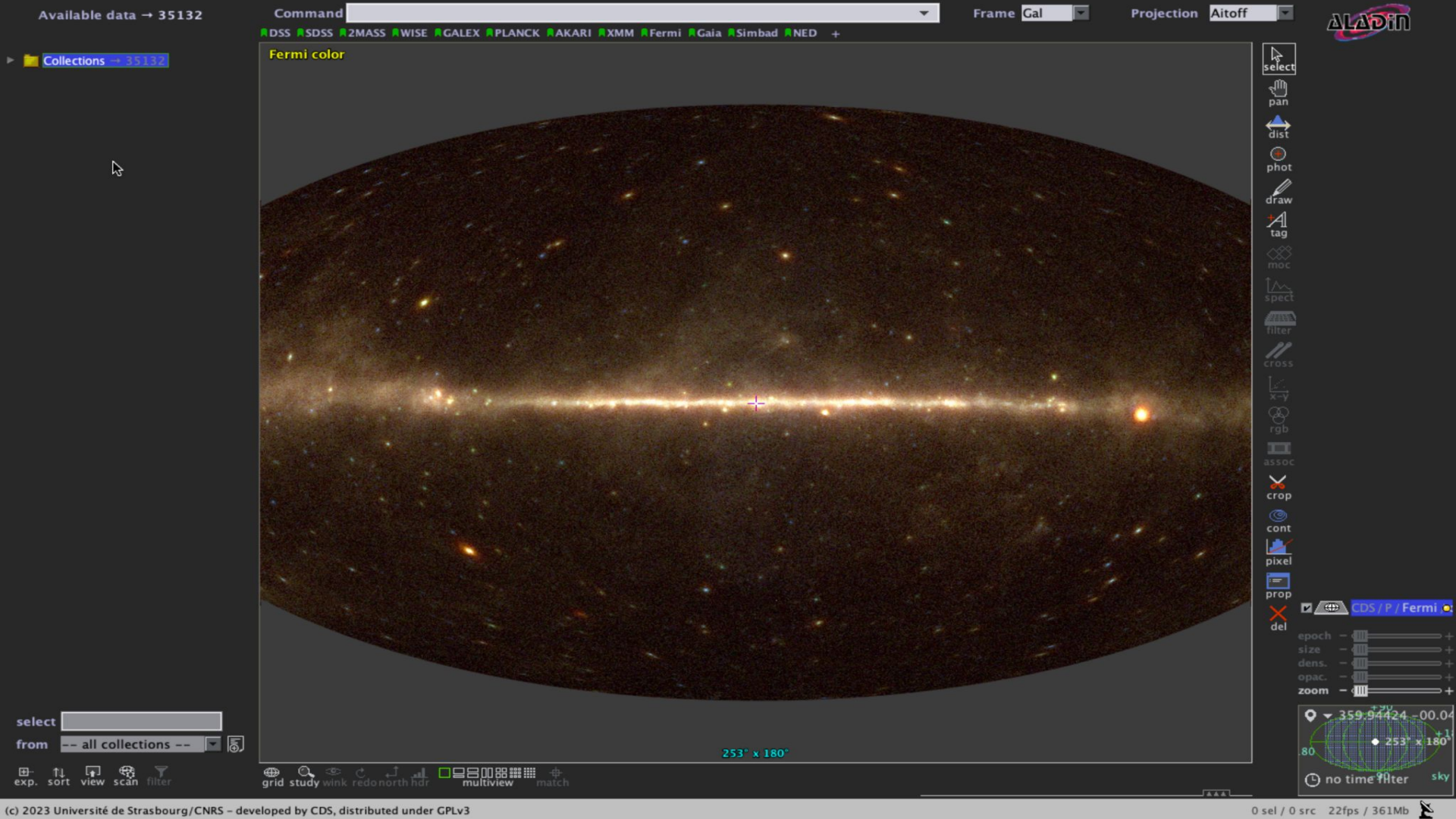
TOPCAT

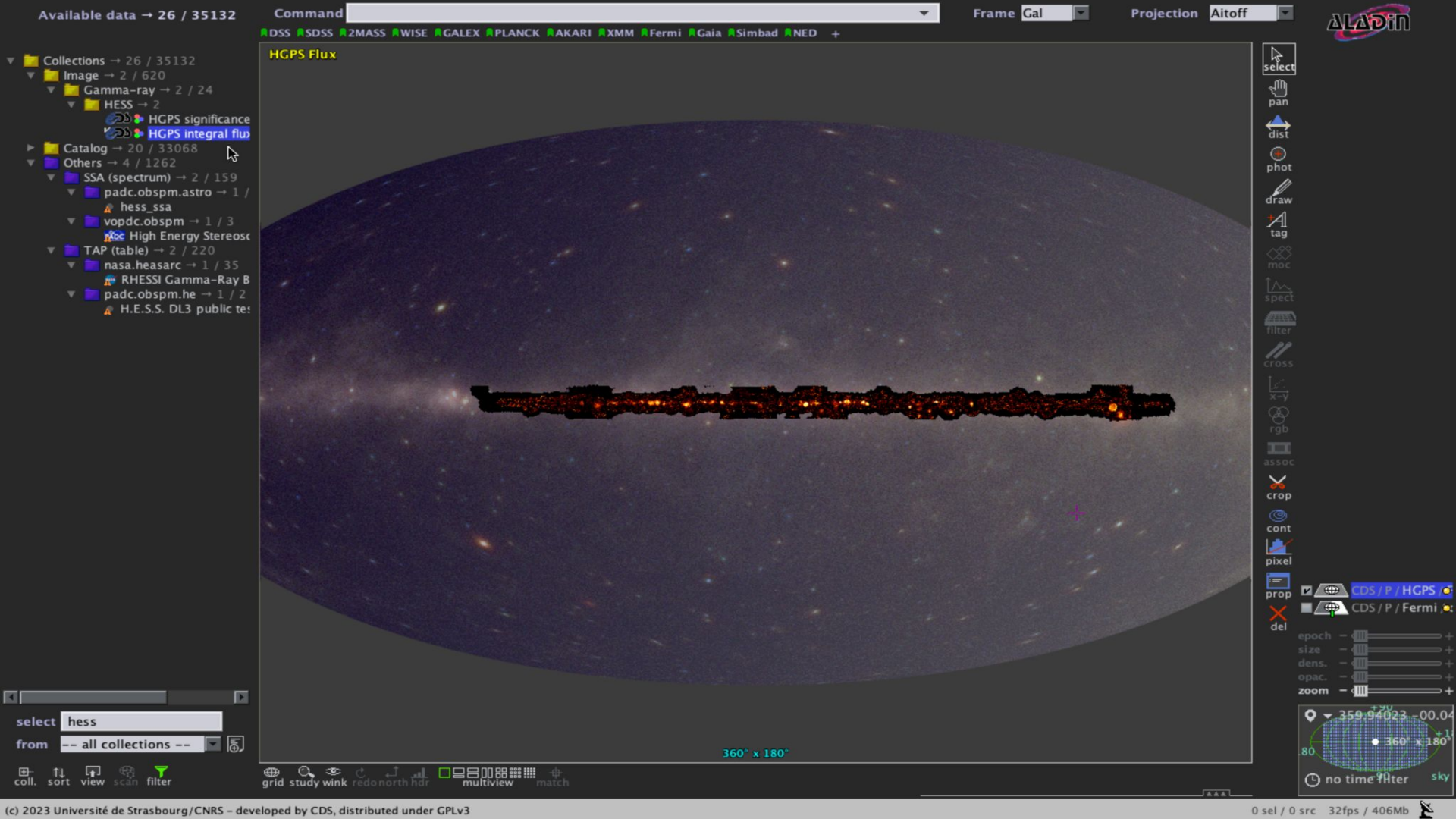
F = Findable : generation of an ObsCore Table

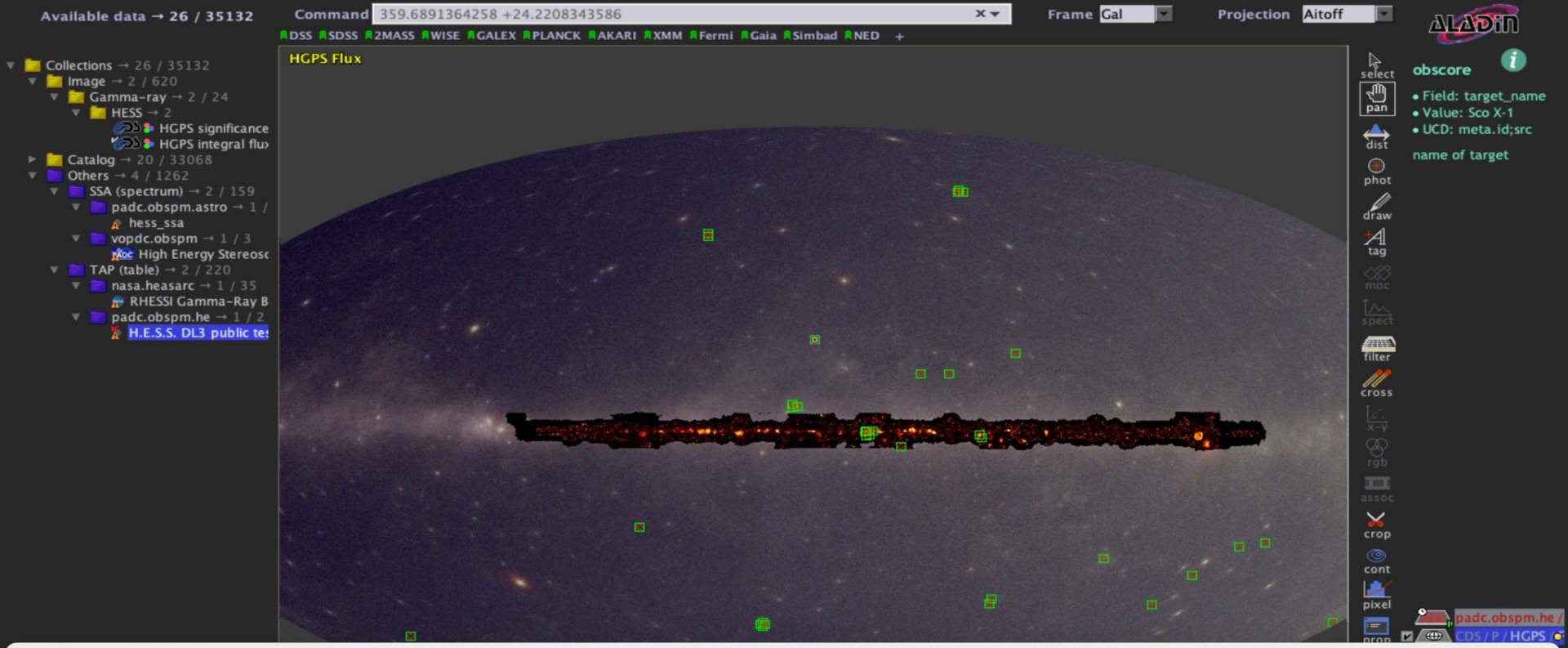
Module **ivoa.py** now included in **Gammapy**

Export of the DataStore into an IVOA ObsCore table with proper metadata to build an IVOA TAP service

dataproduct_type	calib_level	target_name	obs_id	obs_collection	obs_publisher_did	access_url	access_format	access_estsize	s_ra	s_dec	s_fov
								kbyte	deg	deg	deg
str10	int32	str25	str10	str10	str30	str30	str30	int32	float64	float64	float64
EVENTS	2	AGN monitoring	513837	DL3	ivo://ctao#<internal_id>	URL<internal_id>	application/fits	1797	327.5722	-14.7231	10.0
EVENTS	2	AGN monitoring	513839	DL3	ivo://ctao#<internal_id>	URL<internal_id>	application/fits	1785	356.2607	-16.4372	10.0
EVENTS	2	AGN monitoring	513833	DL3	ivo://ctao#<internal_id>	URL<internal_id>	application/fits	1664	262.7	-0.2026	10.0


Available data → 35132
Command
Frame Gal
Projection Aitoff
DSS SDSS 2MASS WISE GALEX PLANCK AKARI XMM Fermi Gaia Simbad NED +
Collections → 35132
Fermi color
select
pan
dist
phot
draw
tag
moc
spect
filter
cross
x-y
rgb
assoc
crop
cont
pixel
prop
del
CDS / P / Fermi
epoch
size
dens.
opac.
zoom
359.94424 -00.04
253° x 180°
80
no time filter
sky
select
from -- all collections --
exp. sort view scan filter
253° x 180°
grid study wink redo north hdr multiview match
(c) 2023 Université de Strasbourg/CNRS - developed by CDS, distributed under GPLv3
0 sel / 0 src 22fps / 361Mb


Available data → 26 / 35132
Command
Frame Gal
Projection Aitoff
DSS SDSS 2MASS WISE GALEX PLANCK AKARI XMM Fermi Gaia Simbad NED
HGPS Flux
Collections → 26 / 35132
Image → 2 / 620
Gamma-ray → 2 / 24
HESS → 2
HGPS significance
HGPS integral flux
Catalog → 20 / 33068
Others → 4 / 1262
SSA (spectrum) → 2 / 159
padc.obspm.astro → 1 /
hess_ssa
vopdc.obspm → 1 / 3
High Energy Stereosc
TAP (table) → 2 / 220
nasa.heasarc → 1 / 35
RHESSI Gamma-Ray B
padc.obspm.he → 1 / 2
H.E.S.S. DL3 public tes
select hess
from -- all collections --
coll. sort view scan filter
select pan dist phot draw tag moc spect filter cross x-y rgb assoc crop cont pixel prop del
CDS / P / HGPS
CDS / P / Fermi
epoch size dens. opac. zoom
359.94023 -00.04
360° x 180°
no time filter
sky
360° x 180°
grid study wink redo north hdr multiview match
(c) 2023 Université de Strasbourg/CNRS – developed by CDS, distributed under GPLv3
0 sel / 0 src 32fps / 406Mb



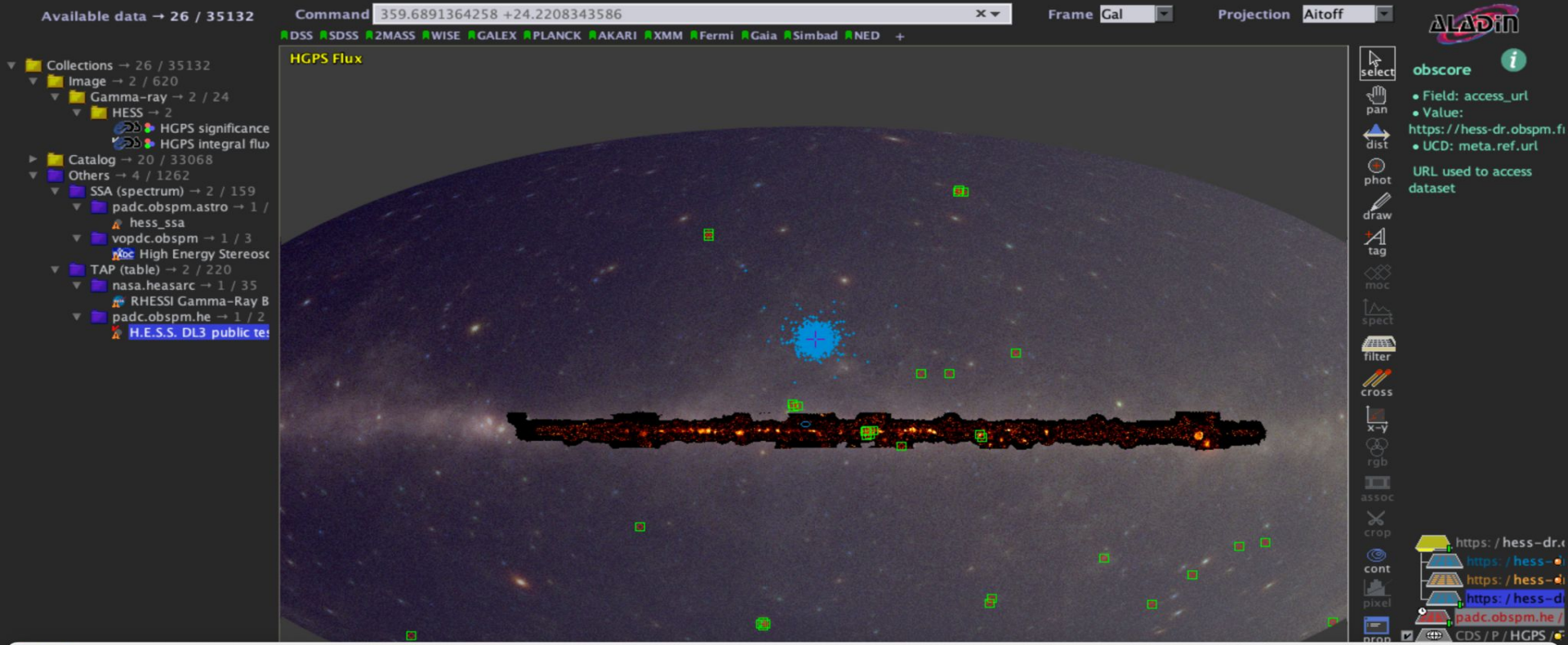
Aladin Java measurements frame

obscore – target_name: name of target

Search

access_url	dataproduct_type	dataproduct_sub...	calib_level	obs_collection	obs_id	obs_publisher_did	access_format	access_estsi...	target_name	s_ra
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	26827	ivo://padc.obspm/hess#...	application/fits	241920	Arp 220	234.5016784667..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	26791	ivo://padc.obspm/hess#...	application/fits	172800	Arp 220	233.7383728027..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	26077	ivo://padc.obspm/hess#...	application/fits	239040	Sco X-1	244.9916687011..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	25511	ivo://padc.obspm/hess#...	application/fits	233280	3C 273	187.7781982421..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	25443	ivo://padc.obspm/hess#...	application/fits	325440	3C 273	187.2778930664..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	25345	ivo://padc.obspm/hess#...	application/fits	334080	3C 273	187.2778930664..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	23736	ivo://padc.obspm/hess#...	application/fits	259200	SN 1987A	83.86775207519..

105 sel / 105 src 872Mb



Aladin Java measurements frame

obscore – access_url: URL used to access dataset

Search

access_url	dataproduct_type	dataproduct_sub..	calib_level	obs_collection	obs_id	obs_publisher_did	access_format	access_estsi..	target_name	s_ra
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	26827	ivo://padc.obspm/hess#..	application/fits	241920	Arp 220	234.5016784667..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	26791	ivo://padc.obspm/hess#..	application/fits	172800	Arp 220	233.7383728027..
https://hess-dr.obspm.fr/retrieve/hess_dl3_dr1_obs_id_026077.fits.gz			2	HESS-DR	26077	ivo://padc.obspm/hess#..	application/fits	239040	Sco X-1	244.9916687011..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	25511	ivo://padc.obspm/hess#..	application/fits	233280	3C 273	187.7781982421..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	25443	ivo://padc.obspm/hess#..	application/fits	325440	3C 273	187.2778930664..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	25345	ivo://padc.obspm/hess#..	application/fits	334080	3C 273	187.2778930664..
https://hess-dr.obspm.fr/	event-list	events	2	HESS-DR	23736	ivo://padc.obspm/hess#..	application/fits	259200	SN 1987A	83.86775207519..

Click on it to load the corresponding archive data in the stack

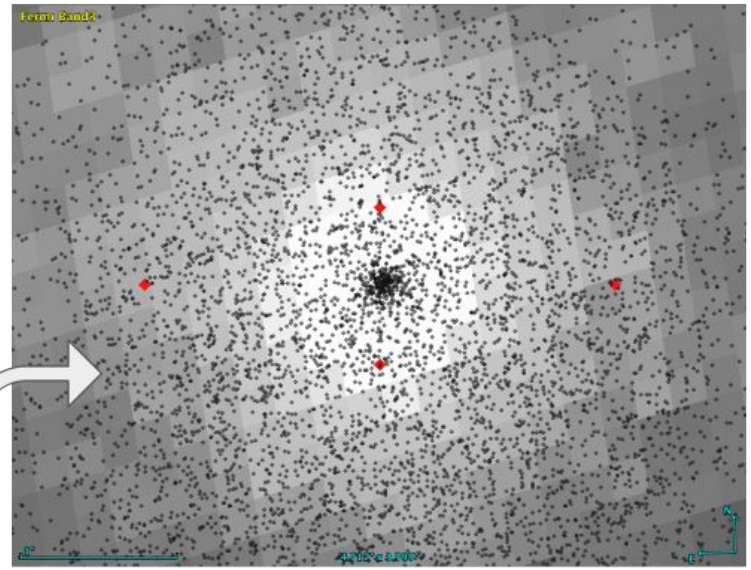
https://hess-dr.obspm.fr/retrieve/hess_dl3_dr1_obs_id_026077.fits.gz

105 sel / 6493 src 35fps / 475Mb

A = Accessible

- ObsCore **access_url**
  - Direct download link to the FITS file
- IVOA **DataLink** (to be implemented):
  - Access to different storage services
  - Access to analysis services, previews
- **Access rights**
  - Public data: no restrictions
  - Anticipating need for **permissions**:
    - PI proprietary period
    - Federation authentication

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
https://...	event	HESS-DR	23523	application/fits	285120	Crab Nebula	83.6333...	21.514
https://...	event	HESS-DR	23526	application/fits	282240	Crab Nebula	83.6333...	22.514
https://...	event	HESS-DR	23559	application/fits	285120	Crab Nebula	85.2533...	22.014
https://...	event	HESS-DR	23592	application/fits	273600	Crab Nebula	82.0133...	22.014

I = Interoperable

VODF
very-high-energy open data format

+

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

A community initiative to define **common data formats for gamma-ray astronomy** based on FITS
https://vodf.readthedocs.io
https://gamma-astro-data-formats.readthedocs.io

- Includes formats for: event lists, effective area, energy resolution, point spread function, instrumental background…
- More an more used by current instruments: Fermi-LAT, HESS, VERITAS, MAGIC, FACT, …

[A&A 625, A10, 2019] [A&A 632, A72, 2019] [A&A 632, A102, 2019]

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

The road to an open format for VHE

The building phase of CTA gave us (quite a lot…) of time to develop its data and analysis tools

GADF

A (short History of GADF)

2011 Prototypes for the CTA data format and science tools

2016
- Establishment of the Gamma-ray Astronomy Data Formats (GADF) initiative
- First preliminary release version (0.1), mainly focused on IACTs

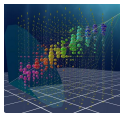
2018
- Version 0.2 released
- Support implemented in the science tools Gammapy and ctools
- H.E.S.S. releases ≈ 50 h of observations of different sources using the format

2019
- FACT, Fermi-LAT, H.E.S.S. MAGIC and VERITAS observations of the Crab Nebula are used to perform the first multi-instrument analysis [doi:10.1051/0004-6361/201834938]
  `https://github.com/open-gamma-ray-astro/joint-crab`
- `ctools` based analysis of the H.E.S.S. data release [doi:10.1051/0004-6361/201936010]
- Comparison of Gammapy and ctools using the H.E.S.S. data release [doi:10.1051/0004-6361/201936452]

from Maximilian Linhoff

Validation of standardized data formats and tools for ground-level particle-based gamma-ray observatories doi:10.1051/0004-6361/202243527

GADF @Gammapy



From B. Khelifi

Prospects for combined analyses of hadronic emission from γ-ray sources in the Milky Way with CTA and KM3NeT arXiv:2309.03007

From GADF to VODF

GADF has been a cornerstone step for interoperability and development, but it has problems...

- Additional parametrization is needed for IRFs (Event types, simulated datasets)
- Standard is vague instead of well described
- Interoperability is not complete or immediate
- There is no governing structure
- Lack of standardized metadata - the **format is not FAIR or VO compliant**

The GADF team decided to create a new more structured initiative with the intent of **linking the VHE community to VO standards**

VODF: a new structure for open science

https://vodf.readthedocs.io

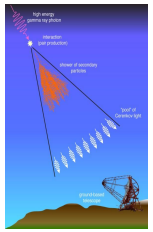
Steering committee (11)
Lead editors (3)
Conveners (2)
Since 2022

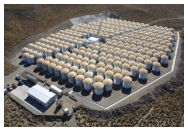
GADF
Data Formats for
Gamma-Ray Astronomy

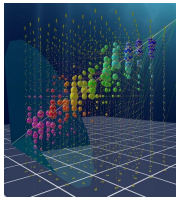
VODF
very-high-energy open data format

A Python package for gamma-ray astronomy

cherenkov telescope array







The VODF committee

very-high-energy open data format

VODF Steering committee

Steering Committee

Facility	Category	Representative
ASTRI	Pointing γ-ray instrument	Fabio Pintore
CTAO	Pointing γ-ray instrument	Roberta Zanin
FACT	Pointing γ-ray instrument	Maximilian Linhoff
Fermi-LAT	Slewing γ-ray instrument	Nicola Omodei
HAWC	Slewing γ-ray instrument	Xiaojie Wang
H.E.S.S.	Pointing γ-ray instrument	Bruno Khélifi
IceCube	Neutrino detector	Marcos Santander
KM3Net	Neutrino detector	Kay Graf
MAGIC	Pointing γ-ray instrument	Cosimo Nigro
SWGO	Slewing γ-ray instrument	Andrew Smith
VERITAS	Pointing γ-ray instrument	Amanda Weinstein

Lead editors:

Karl Kosack (IACTs)
Laura Olivera-Nieto (Slewing instruments)
Jutta Schnabel (Neutrino instruments)

Conveners:

Bruno Khélifi
Roberta Zanin

What's next for VODF ?

very-high-energy open data format

Now that the VODF structure is in place, it's time to take concrete steps towards the creation of the new standards for VHE

- Allow for well-defined usage of multiple IRFs in the same data processing
- Introduction of Event Types
- Define VO keywords with VHE specificities - strong base in X-Ray astronomy
- Set standards for **Metadata**, not just data
- Improve discoverability using VO recommendations - Observatory Core Component Data Model (**ObsCore**)
- Integrate completely into IVOA interest groups (**HEIG**)

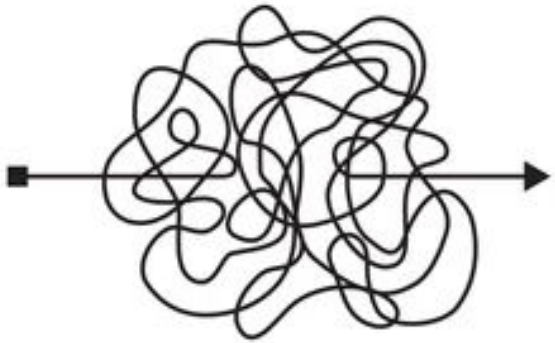
A **Python** package for **gamma-ray** astronomy

From F-A-I to FAIR

F-A-I are very technical points.

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**, **store** it and **link** it to the data

→ **IVOA Provenance standard data model!**

Summary and conclusion

The VHE-multimessenger community strongly values **open, accessible and interoperable data and analysis software** and is taking prioritary steps towards it

VODF
very-high-energy open data format

- Built on the big success of GADF as an interoperable standard
- Will be **VO compliant**
- Has a defined governing structure

June 2023 in Paris

IVOA standards for High Energy Astrophysics
IVOA Note to justify a HE Interest Group

https://github.com/ivoa/HighEnergyDataNote

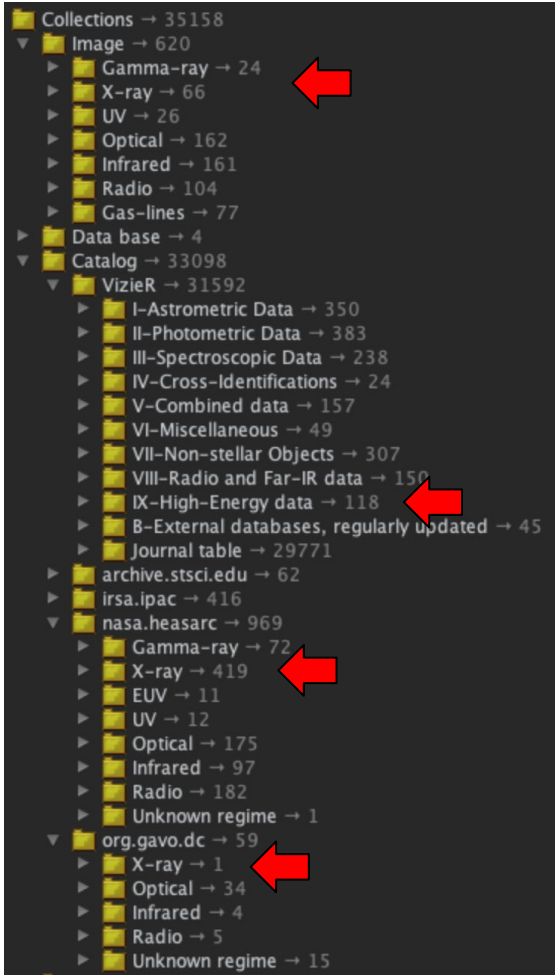
https://indico.obspm.fr/event/1963

BACKUP

Current High Energy data in the VO

- Images
  - Fermi Full Sky, eRosita, XMM-Newton, H.E.S.S. galactic plane
- Catalogs
  - Vizier dedicated section
  - NASA HEASARC (many legacy archives)
  - Generally catalogs of sources
- SIA/TAP services
  - High level data, catalogs, proposals, …

Aladin Search Tree



CHANDRA SOURCE CATALOG
Table Access Protocol

IVOA

XMM-Newton Science Archive
TAP queries to the XSA database

NASA's HEASARC
High Energy Astrophysics Science Archive Research Center

A long way

It's a Long Way... from Private Ground-based Gamma-ray Data to Public Release: Open-data, Open-source Tools, First Real TeV Data Release from H.E.S.S.

[C. Boisson et al. 2020, ADASS XXVII Santiago, ASPC 522 497B]

- Imaging Atmospheric Cherenkov Telescopes
- Gamma **data format** initiative
- **Open tools** for analysis

→ A story of the relations between Cherenkov observatories and the IVOA

H.E.S.S.

cta cherenkov telescope array

EURO VO CoSADIE Collaborative and Sustainable Astronomical Data Infrastructure for Europe

2012-2014

Asterics Astronomy ESFRI & Research Infrastructure Cluster

2015-2019

ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

2019-2023

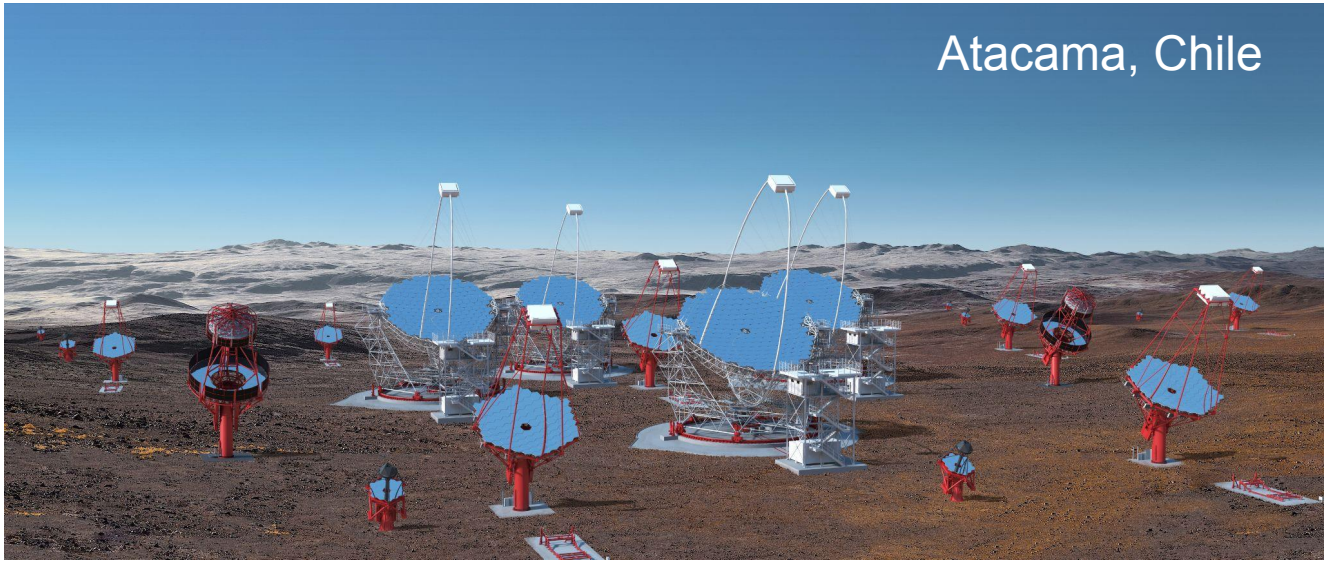
Cherenkov Astronomy

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 Higlands, Namibia
Since 2002: four 12m telescopes
Since 2012: added 32m by 24m telescope
Since 2015: camera upgrades on 12m telescopes

@ Jeff Grube



Alpha configuration

CTAO Southern array

51 telescopes over a ~3 km2 area

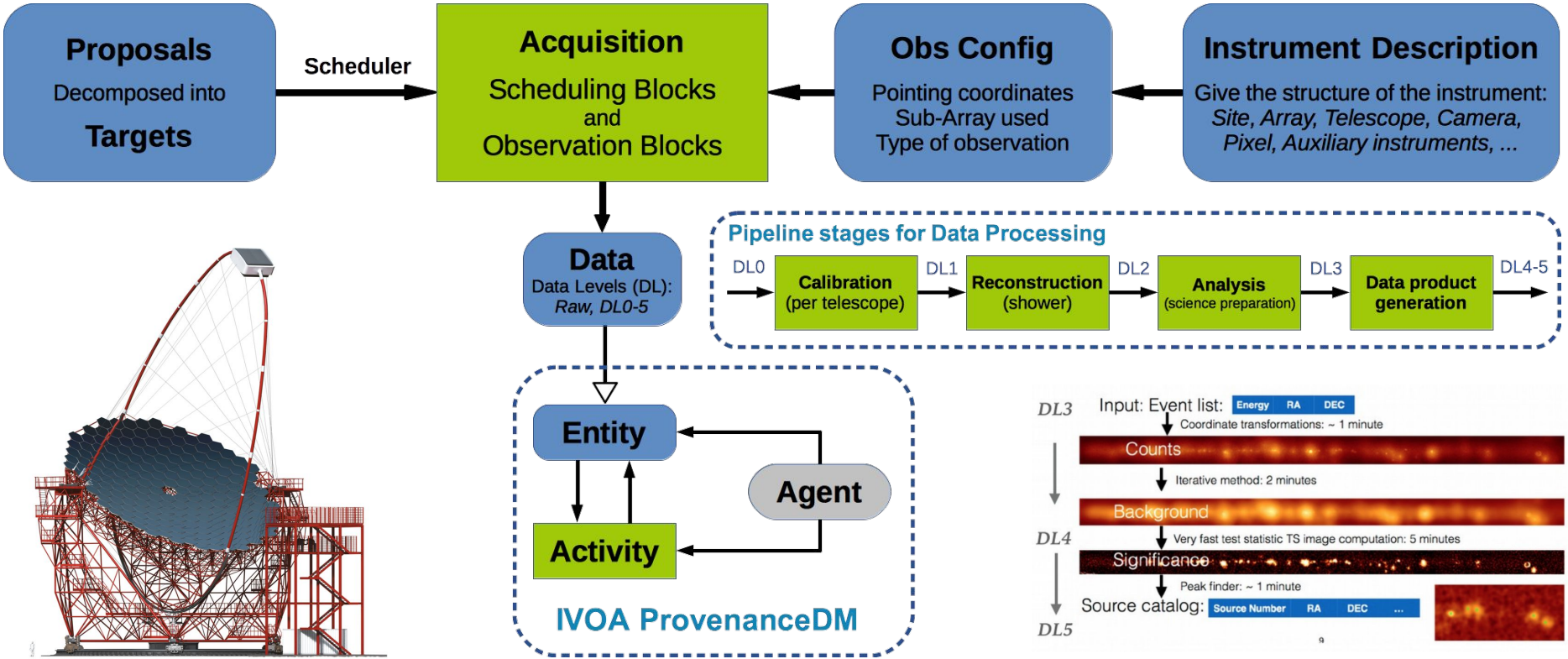
14 Medium-Sized Telescopes (MST)
37 Small-Sized Telescopes (SST)

CTAO Northern array

13 telescopes distributed over an area of about 0.5 km2

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

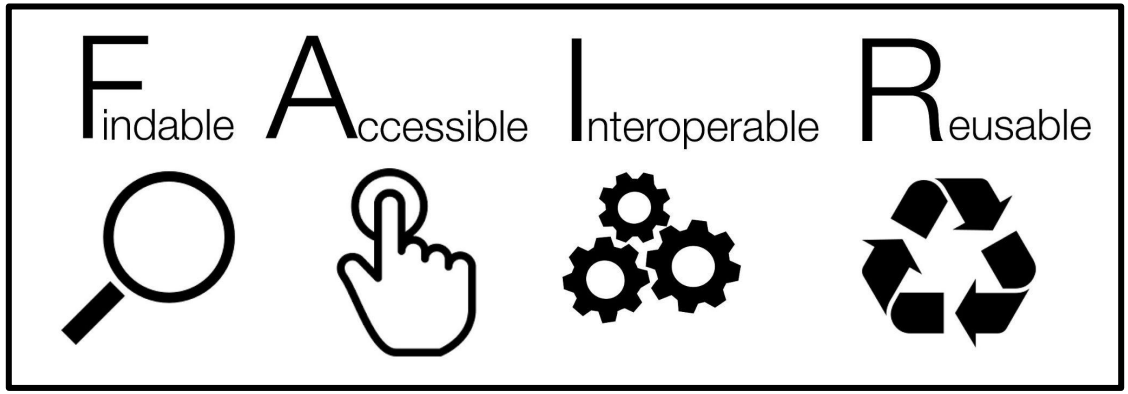
CTA Master Configuration Data Model



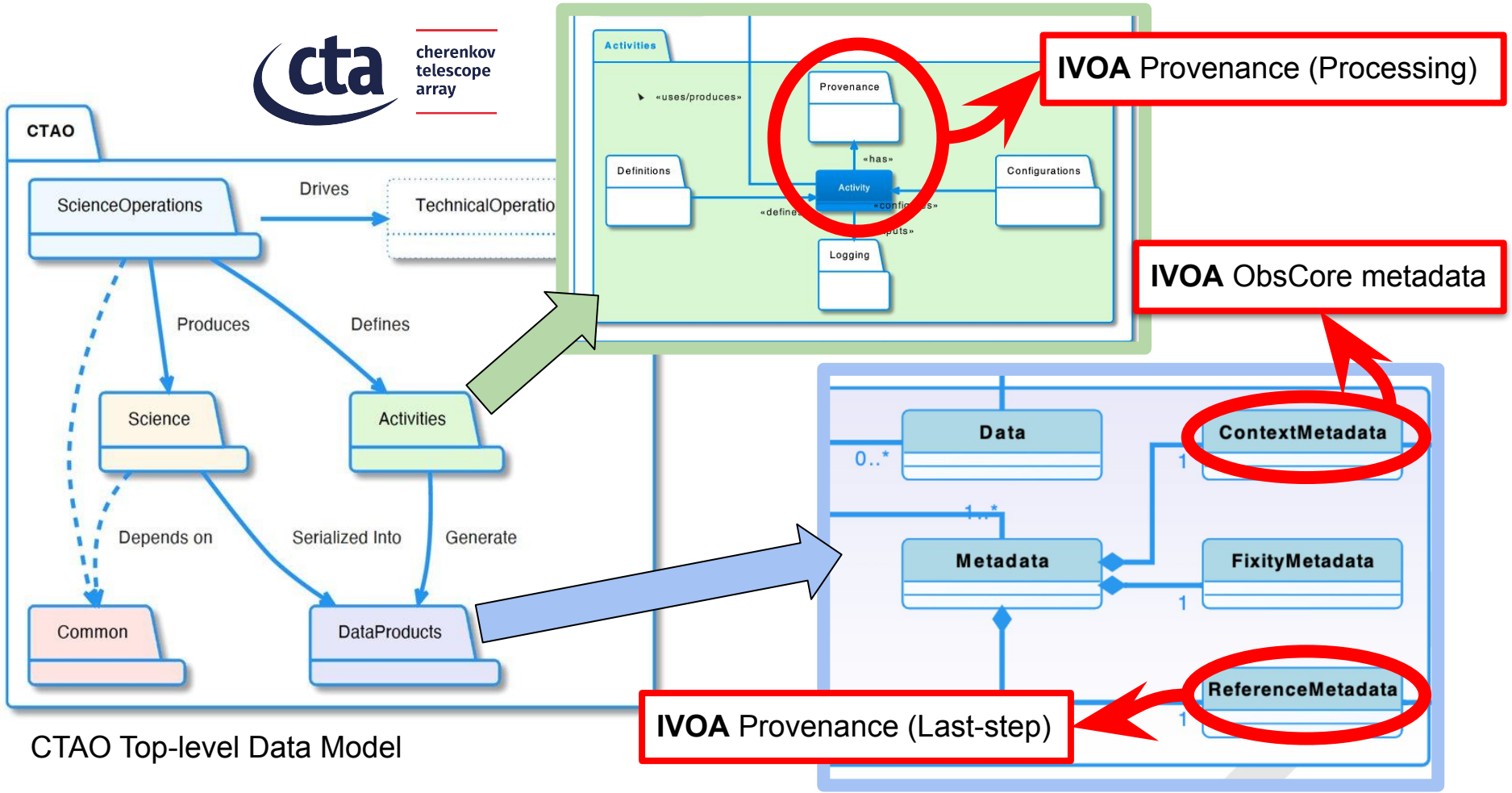
Open Observatory and Open Science

CTA will operate as an **open observatory** and will provide data to the scientific community. In the context of **Open Science**, the data provided by CTA must follow the **FAIR Guiding Principles** for scientific data management:

cherenkov telescope array



See e.g. Servillat et al. 2022, ADASS XXXI, ASP Conference Series
"FAIR high level data for Cherenkov astronomy"
https://hal-obspm.ccsd.cnrs.fr/obspm-03516688



CTAO Top-level Data Model

A long way...

- Interactions between observatories and the IVOA
  - Astronomers and engineers need to be part of IVOA
  - They then need to convince their organisation!
  - And maintain the momentum!

- Early engagement of observatories/facilities
  - Construction and operations = different priorities
  - Seeds in the CTA requirements and data models
  - CTAO involved in ESCAPE open collab., OSTrails…
  - Role of European Projects
  - Role of **OV-France**

HE dedicated workshop at OV-France

- **October 2022 in Strasbourg**
  - https://indico.obspm.fr/event/1489

- Continue activities of the **ESCAPE European project** that focused on High Energy Facilities (**2019-2023** H2020 project).
- Bring together representatives of high energy observatories (VHE, HE, GW, neutrino)
- Presentations of HE observatory operations and data:
  - CTA (Mathieu Servillat)
  - Ligo Virgo Kagra (Pierre Chanial)
  - Neutrino (Damien Dornic)
  - XMM & SVOM (Laurent Michel)
  - GADF/VODF (Bruno Khelifi)

HE meeting at IVOA Interop

- **May 2023 at the IVOA Bologna**
  - Dedicated talk at the DM session by M. Servillat:
    - https://wiki.ivoa.net/internal/IVOA/IntropMay30
  - Fruitful splinter IVOA meeting

- Creation of **HE “Club”**
  - IVOA mailing list and wiki page
    - https://wiki.ivoa.net/twiki/bin/view/IVOA/HEGroup
  - Several online meetings
    - IVOA data models (cube, dataset)
    - Detailed of HE data, with a focus on Instrument Response Functions
  - Next call: **April 3rd 16h30**

Second OV-France workshop enlarged to IVOA

- **June 2023 in Paris**
  - *IVOA standards for High Energy Astrophysics*
  - https://indico.obspm.fr/event/1963

- Review of previous documents since 2021
- Focus on **user scenarios** in HE
  - Access and Analysis of HE data
  - Used IVOA standards
  - What specific developments are needed

- Prepare an IVOA Note to justify a HE Interest Group
  - https://github.com/ivoa/HighEnergyDataNote

Content of the Note

- Common practices and concept of event-list
  - **Lower level** dataset, used to generate images, lightcurves, spectra
  - Generally **reprocessed** from **event lists** for a dedicated analysis
  - Calibrated data, but instrument signature not totally removed
  - **Instrument Response Functions** (**IRFs**) are tightly connected
- Data Discovery
  - **ObsCore** for a HE event list?
  - Possible extension for HE
- An event-ist context data model
  - **Relations** to IRF and Instrument Configuration
- Modelling the content of an event-list
  - Cube and Dataset Data Model

HE “event” in the VO

a HE event is **not** a VOEvent

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.

http://www.ivoa.net/rdf/product-type (Preliminary)

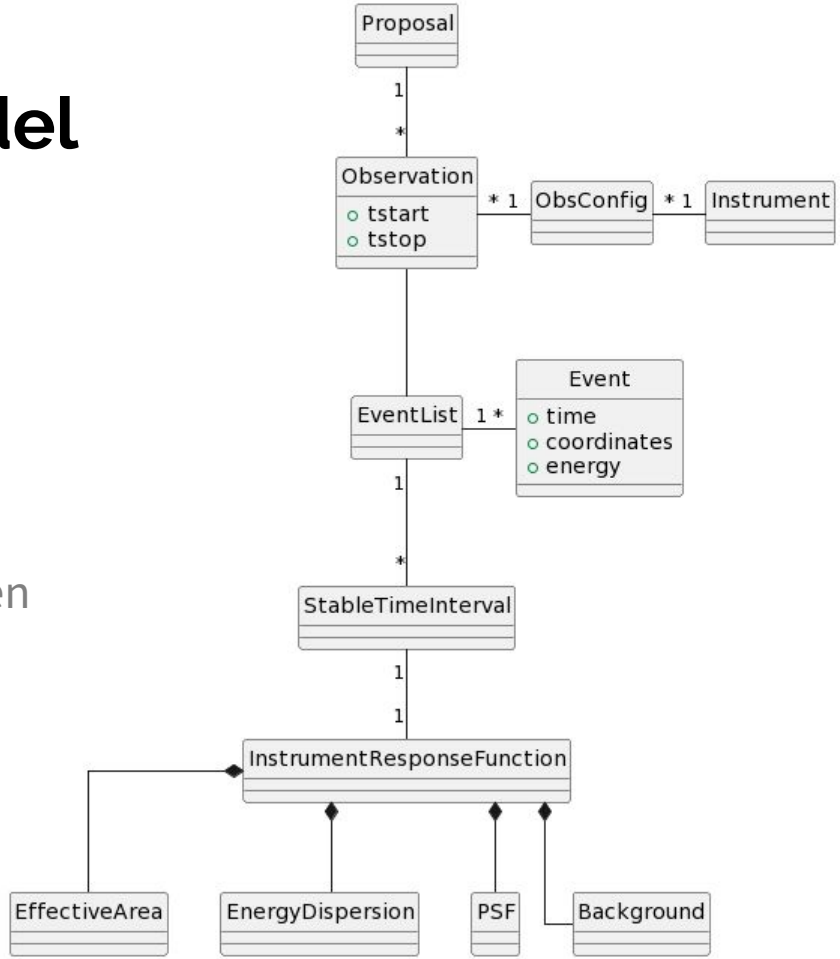
event-list: A collection of observed events, such as incoming high-energy particles. A row in an event list is typically characterised by a spatial position, a time and an energy.

Event-list Context Data Model

- Issue
  - What is really in the event-list dataset?
  - Does it include IRFs? only an event-list?
  - Where can one find the corresponding IRFs?
- Need a way to link an event-list to its IRFs

→ A proper **data model** with relations between those elements would help

→ Possible **ObsCore extension** fields would appear in this data model



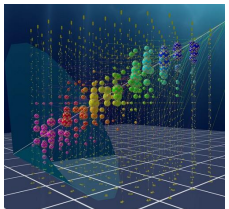
Summary and conclusions

- **HE data available** through the VO
- HE domain have **specificities**
  - In particular the concept of **event-list**
  - and **Instrument Response Functions** (IRFs)
- Very limited number of VO services giving access to event-lists
- On the path to build an **HE Interest Group** at IVOA
  - **Session at IVOA Sydney** in May (plenary or with data model)
  - Next call: April 3rd 16h30

Pointing Gamma-Ray instruments

Slewing Gamma-Ray instruments

Neutrino instruments



Science Analysis: DL3-DL5

Happens at CTA data centers (automatic) + by users on user's laptops or e.g. ESCAPE science platform

Search for data
covering region of interest in space, time, maybe other parameters

VO tools?

DL3
Science-Ready

Retrieve Event Lists and IRFs that cover region of interest

obs_id = 84753

event_id	Time	Energy	RA	Dec
12351	55933.30	0.03	53.6	-29.2

obs_id = 99584

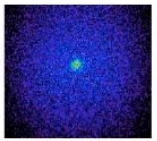
event_id	Time	Energy	RA	Dec
12351	55933.30	0.03	53.6	-29.2
12356	55933.31	0.1	53.2	-29.3
12378	559			
12389	559			

obs_id = 100207

event_id	Time	Energy	RA	Dec
12351	55933.30	0.03	53.6	-29.2
12356	55933.31	0.1	53.2	-29.3
12378	55933.36	12.0	53.8	-29.5
12389	55933.45	0.2	53.7	-28.7

DL4
Science-Binned

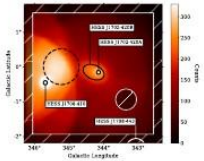
Use *Science Tools* to make intermediate Binned Data Products



Exposure Cube
Counts Cube
Exclusion Map
...

DL5
Science-Quicklook

Use *Science Tools* to Fit models to binned data (forward-folding) and make Flux cubes

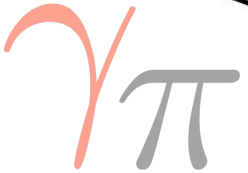


Sky model
Light-Curve
Spectra
Flux Map/Cube

DL6
Science-Observatory

Observatory or CTAC produces some legacy products

Catalogs,
Survey Products



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

From K. Kosack