Experience on building an ObsCore service for HE event-lists

Mathieu Servillat (LUTH - Observatoire de Paris / CNRS)
Catherine Boisson, Paula Kornecki

IVOA Sydney
2024-05-21

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

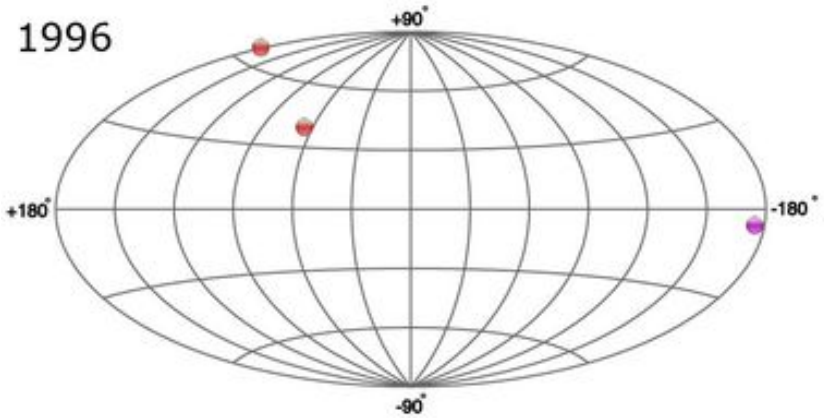
2012-2014

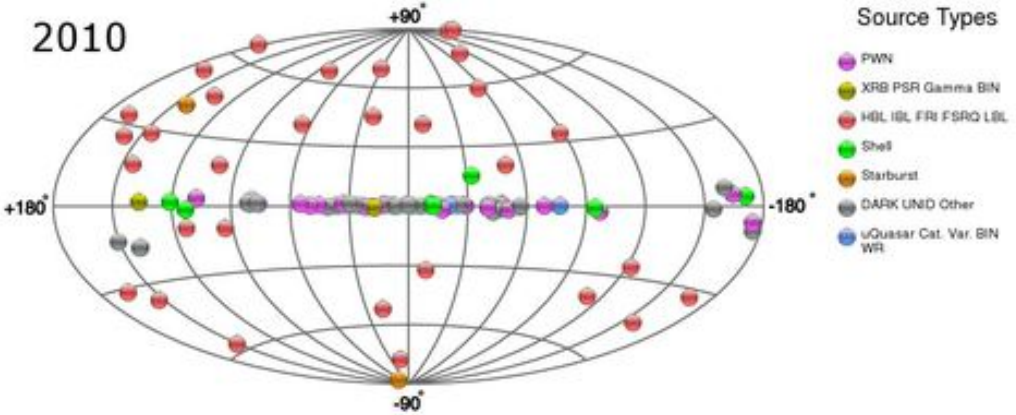
Asterics Astronomy ESFRI & Research Infrastructure Cluster

2015-2019

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

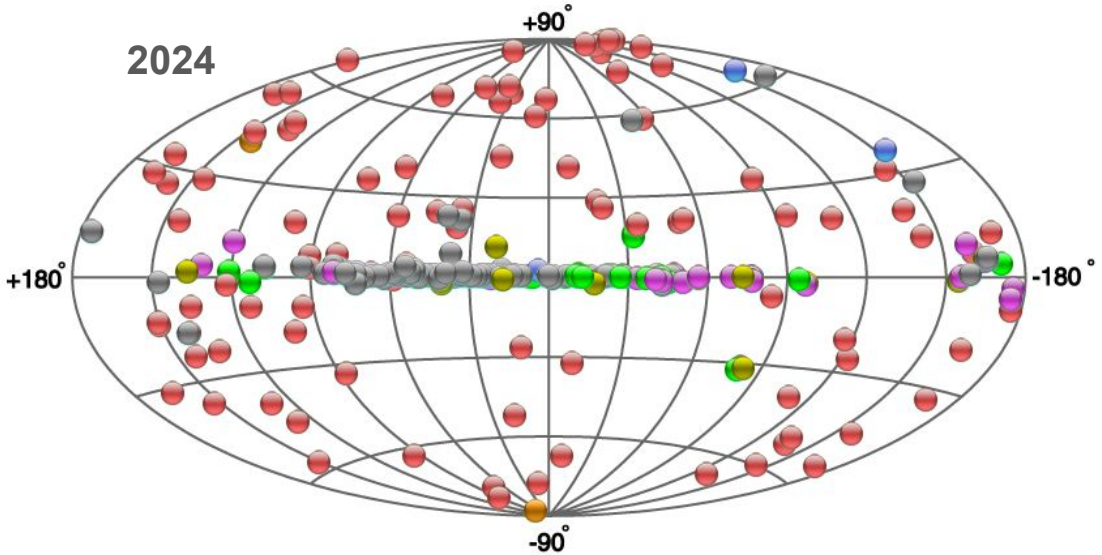
2019-2023





TeVCat2
http://tevcat2.uchicago.edu/

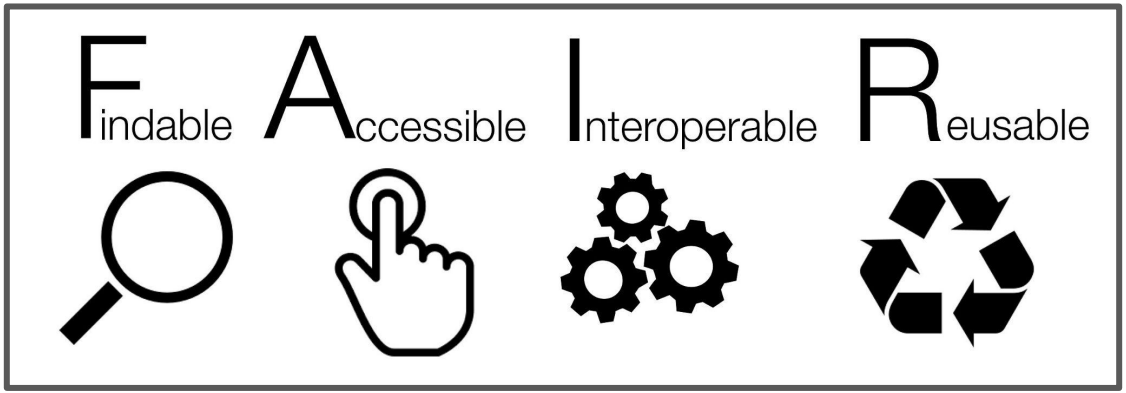
H.E.S.S., MAGIC, VERITAS
> 200 sources
energy > 100 GeV



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

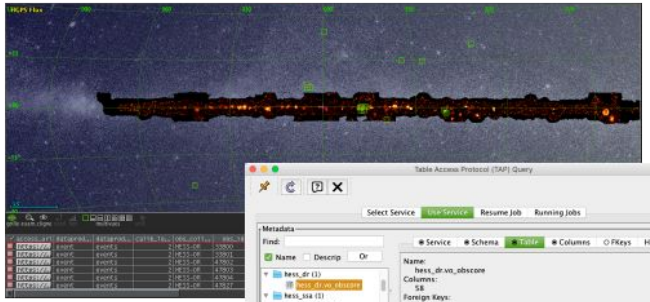
F = Findable

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

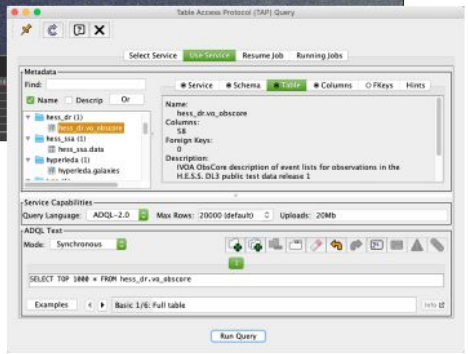
- **IVOA** Observation Data Model Core Components (**ObsCore**)
  [link to IVOA REC] → adapted to Cherenkov data
- **IVOA** Table Access Protocole (**TAP**)
  [link to IVOA REC]
- **Deployed service** at ObsParis
  https://hess-dr.obspm.fr
- Registered to the VO **Registry** via PADC
  (Paris Astronomical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

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

TapHandle



Aladin



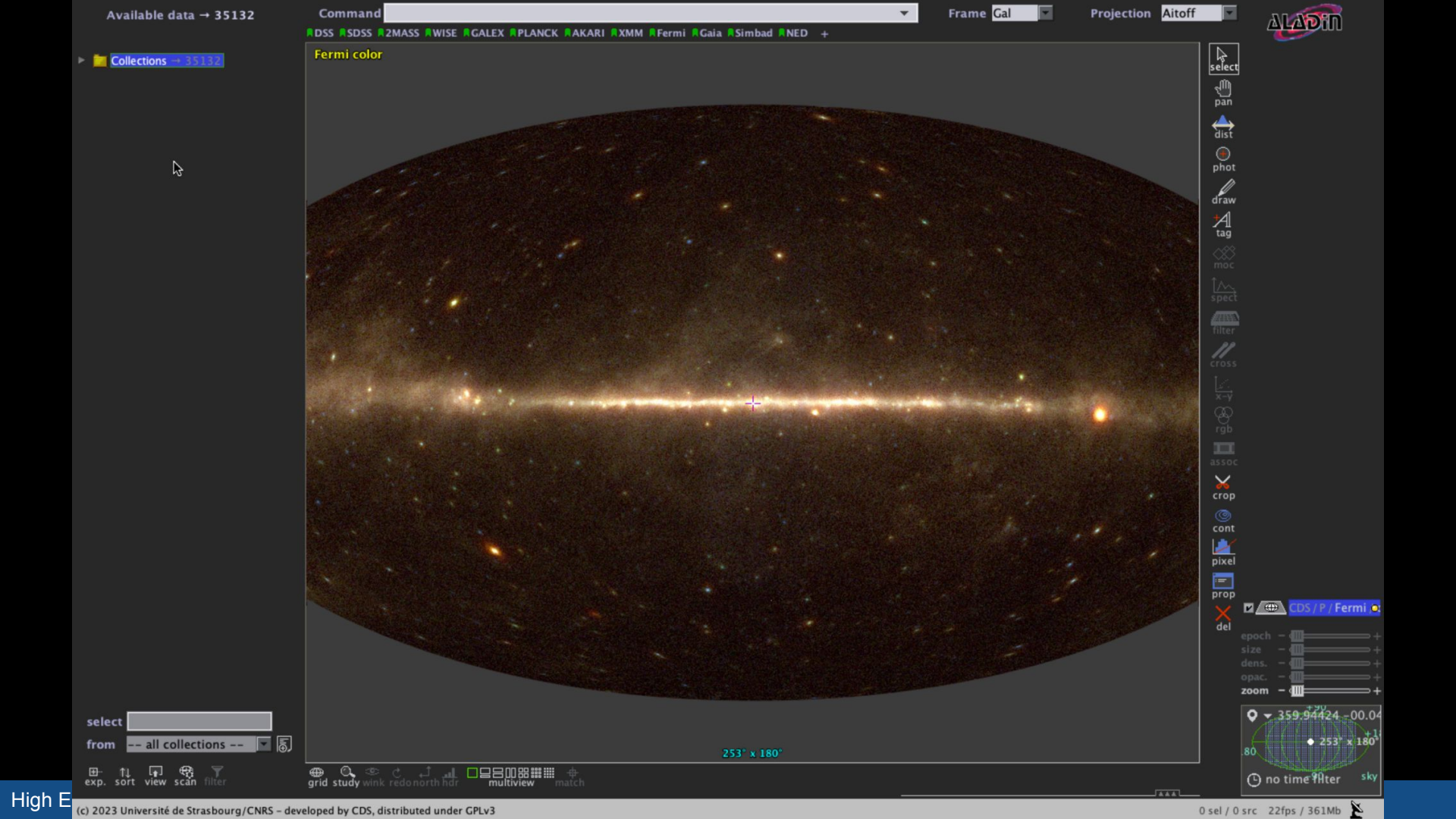
TOPCAT

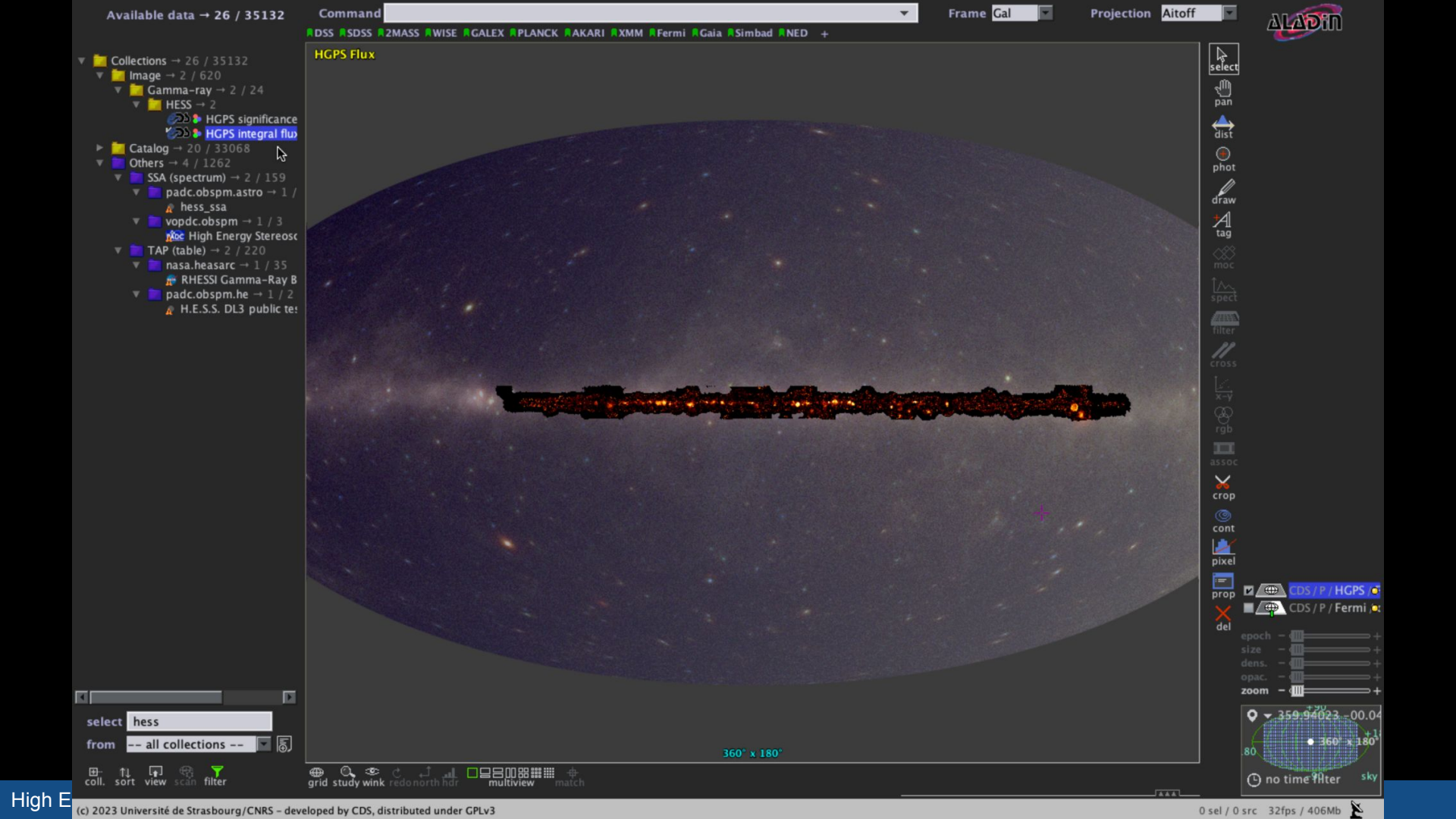
F = Findable : generation of an ObsCore Table

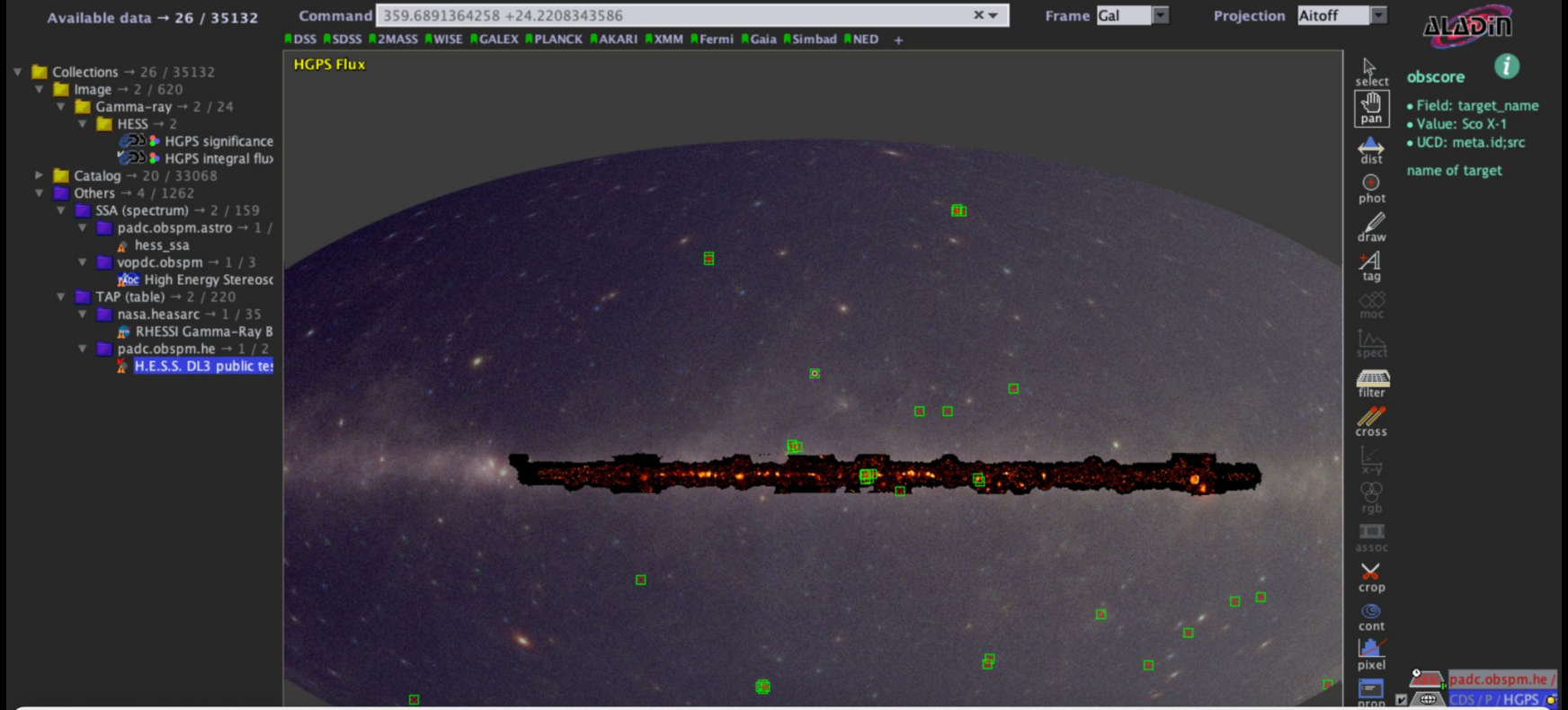
Module **ivoa.py** now included in **GammaPy**

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	0.0
EVENTS	2	AGN monitoring	513839	DL3	ivo://ctao#<internal_id>	URL<internal_id>	application/fits	1785	356.2607	-16.4372	0.0
EVENTS	2	AGN monitoring	513833	DL3	ivo://ctao#<internal_id>	URL<internal_id>	application/fits	1664	262.7	-0.2026	0.0


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°
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 te
select hess
from -- all collections --
coll. sort view scan filter
grid study wink redo north hdr multiview match
360° x 180°
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
(c) 2023 Université de Strasbourg/CNRS – developed by CDS, distributed under GPLv3
0 sel / 0 src 32fps / 406Mb



Aladin Java measurements frame

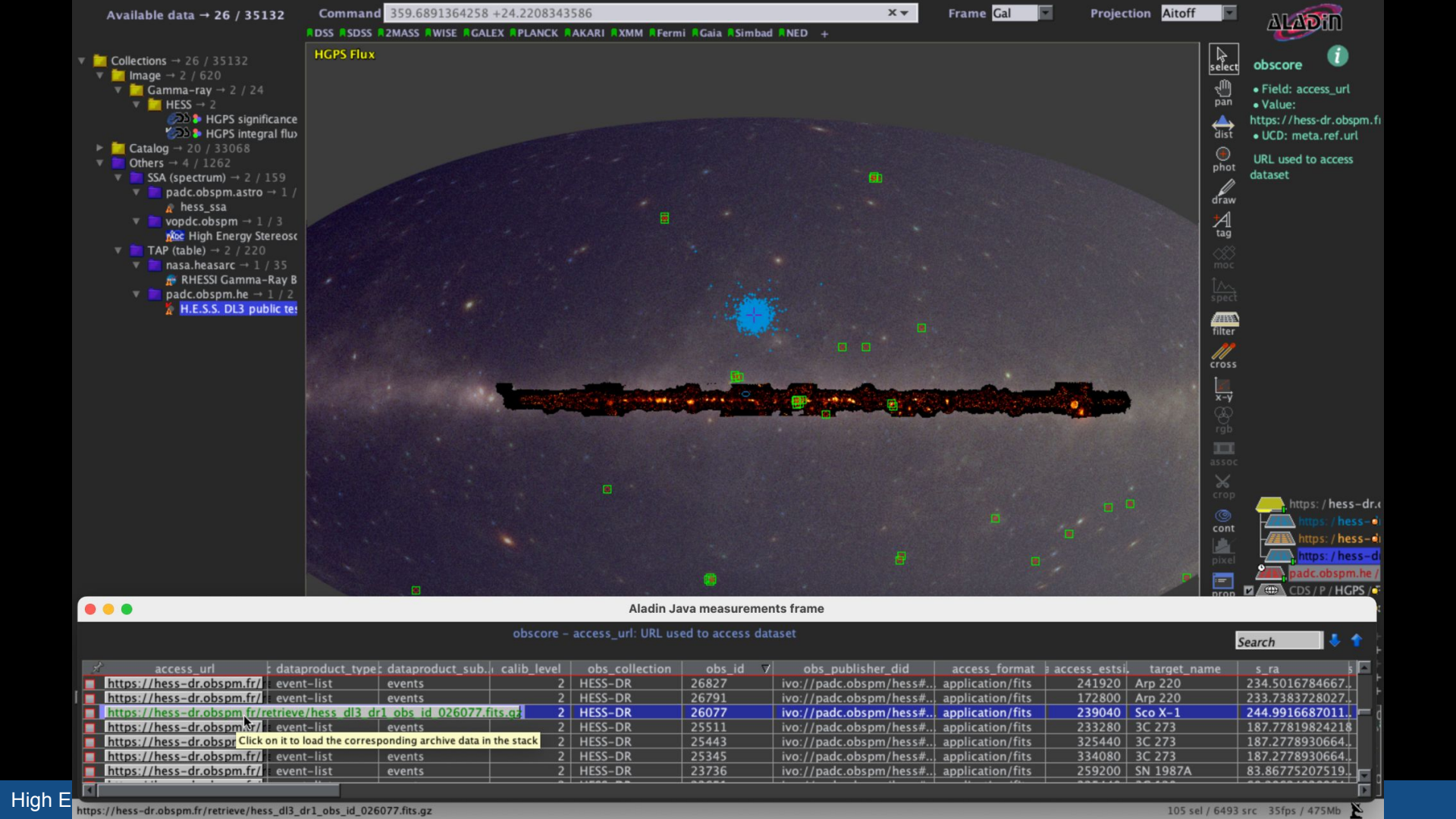
obscore – target_name: name of target

Search

access_url	dataproduct_type	dataproduct_sub...	calib_level	obs_collection	obs_id	obs_publisher_did	access_format	access_estsi...	target_name	s_ra	s
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

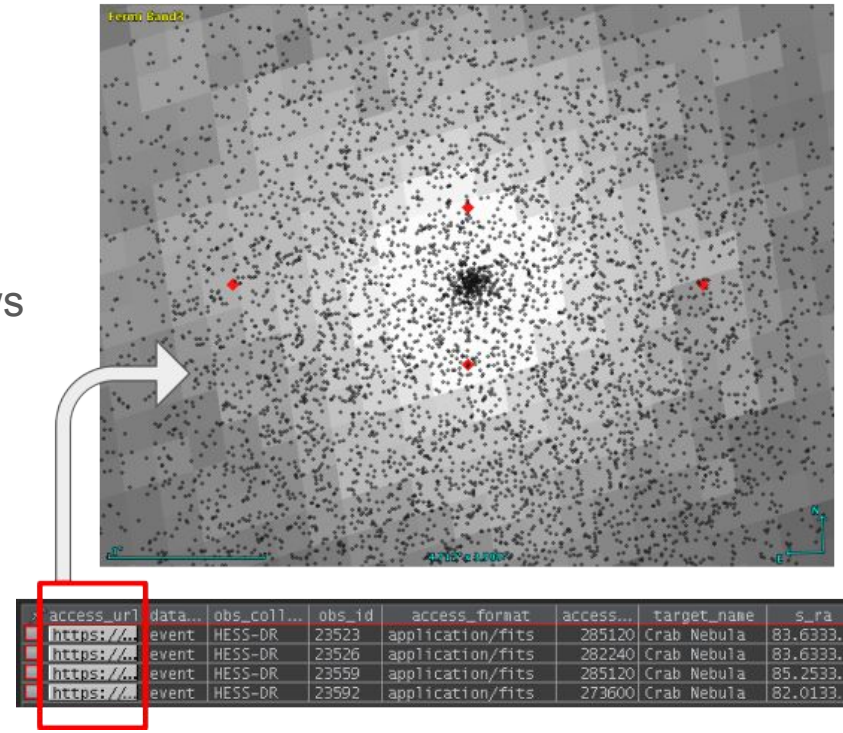
High E


Available data → 26 / 35132
Command 359.6891364258 +24.2208343586
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 te
select
pan
dist
phot
draw
tag
moc
spect
filter
cross
x-y
rgb
assoc
crop
cont
pixel
prop
obscore
• Field: access_url
• Value:
https://hess-dr.obspm.fr
• UCD: meta.ref.url
URL used to access dataset
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 273 187.7781924218
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
High E

A = Accessible

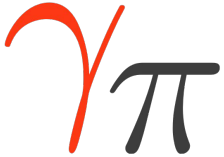
- ObsCore `access_url`
  - Direct download link to the FITS file
- IVOA **DataLink** (to be implemented):
  - Access to different storage services
  - Access to analysis services, previews
- **Access rights**
  - Public data: no restrictions
  - Anticipating need for **permissions**:
    - PI proprietary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

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



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

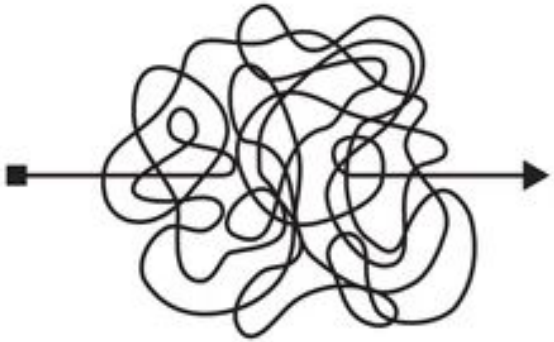
From F-A-I to FAIR

F-A-I

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

Reusability?

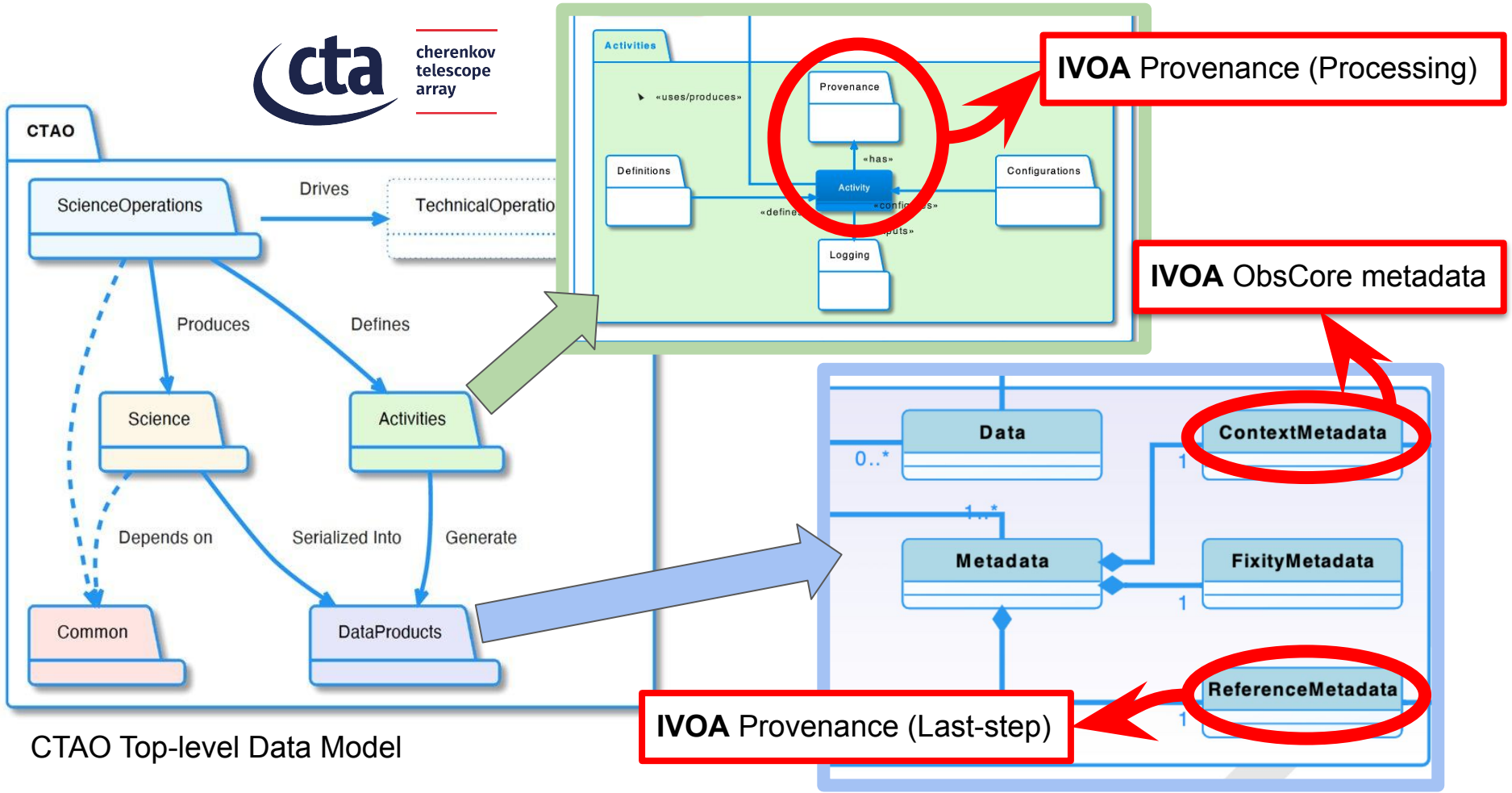
- Based on the **quality** / **reliability** / **trustworthiness** of the products
- What calibration was applied? What tools were used and how?
  What assumptions were made during the data preparation?
- **Sustainability**: with time, key information may disappear...



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 standard data model!**



CTAO Top-level Data Model

A long way…

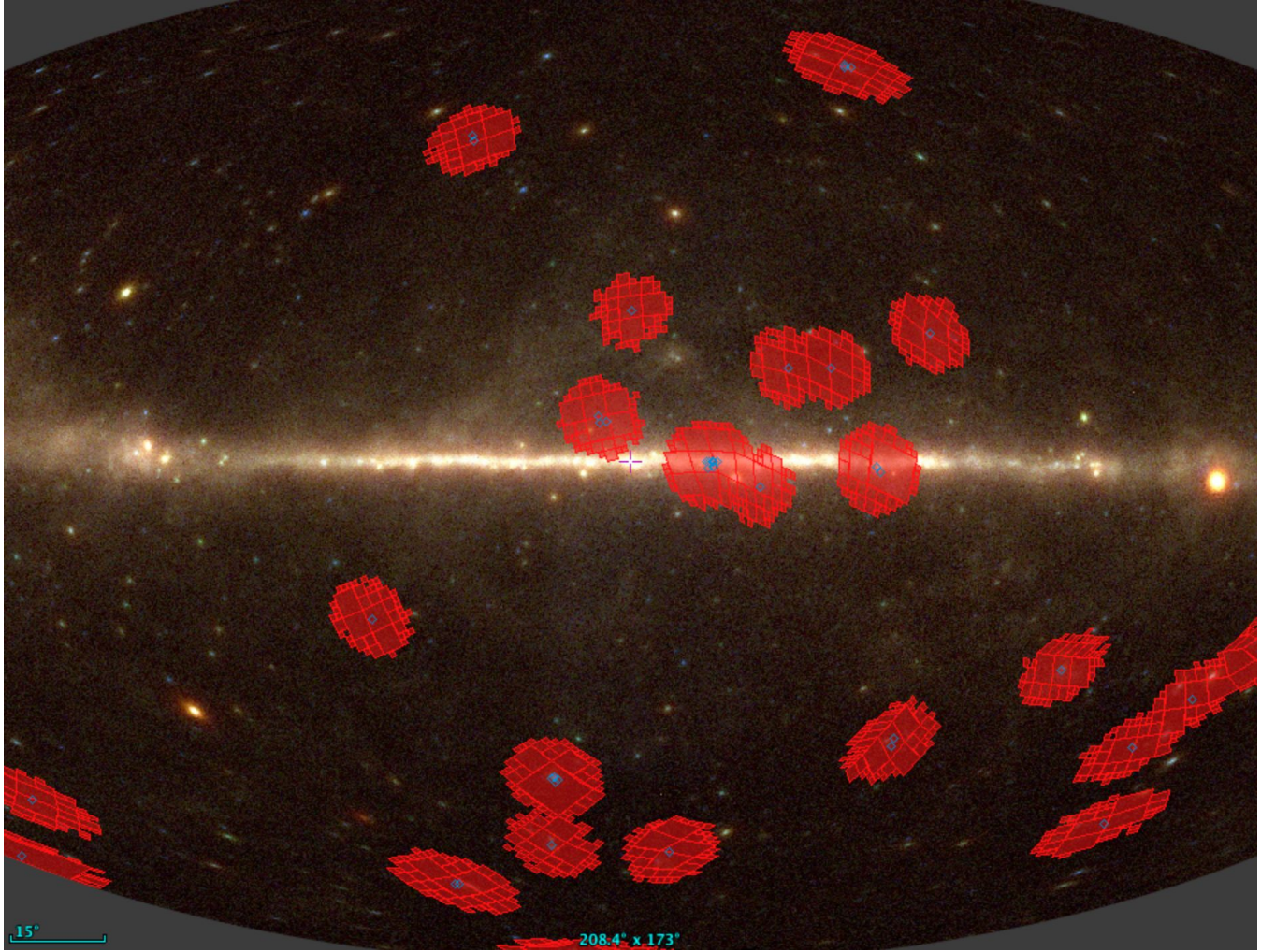
- Interactions between observatories and the IVOA
  - Astronomers and engineers need to be part of IVOA
  - They then need to convince their organisation!
  - And maintain the momentum!
- Early engagement of observatories/facilities
  - Construction and operations = different priorities
  - Seeds in the CTA requirements and data models
  - CTAO involved in ESCAPE open collab., OSTrails…
  - Role of European Projects
  - Role of OV-France

ObsCore fields

- **dataproduct_subtype** = DL3, maybe specific data format (VODF)
- **calib_level** = between 1 and 2
- **obs_collection** could contain many details : obs_type (calib, science, simulation), obs_mode (subarray configuration), pointing_mode, tracking_type, event_type, event_cuts, analysis_type…
- **s_ra, s_dec** = maybe telescope pointing coordinates
- **target_name** : several targets may be in the field of view
- **s_fov, s_region, s_resolution, em_resolution**... all those values are energy dependent, one should specify that the value is at a given energy, or within a range of values.
- **em_min, em_max** : add fields expressed in energy (e.g. eV, keV or TeV)
- **t_exptime** : ontime, livetime, stable time intervals... maybe a T-MOC would help
- **facility_name, instrument_name** : minimalist, would be e.g. CTAO and a subarray

ObsCore possible additions

- **ev_number** : number of rows in the event-list (not possible with _xel)
- **access_format** : Adding MIME-type to table
  - application/x-fits-ogip ...
  - application/x-gadf ...
  - application/x-vodf
- **energy_min, energy_max** : connected to em_min, em_max, but in eV
- **t_gti** : the searching criteria in terms of time coverage require the list of stable/good time intervals to pick appropriate datasets
- **IRF Table**?
- **DataLinks**?


15°
208.4° x 173°