Web-based approach to Data Management

My first 25 years of data handling via RDBMS and web-oriented systems

Luciano Nicastro, INAF – OAS Bologna

Why web-based tools?

Kwds: Collaborative data-driven science – Bringing analysis to data – Accessibility – Science for all

- A browser can be seen as a VM
- No (user) code maintenance required
- Client-server architecture (e.g. Node.JS)
- Advanced graphics (e.g. WebGL)
- Can handle large data sets residing on the server / cloud
- Co-existence of relational and No-SQL DBs (e.g. MongoDB)
- Code implemented in various languages (not just Python)
- User friendly $\Longrightarrow$ not just for professional astronomers
- Real-time or batch processing

CARTA as a reference case

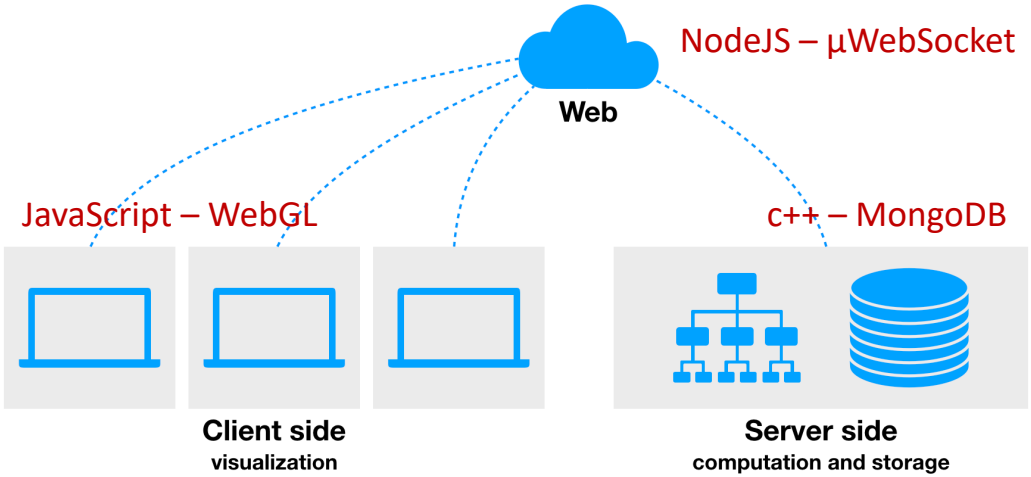
Cube Analysis and Rendering Tool for Astronomy

https://carta.readthedocs.io/en/4.1/

- Image visualization and analysis tool designed for ALMA, VLA, MeerKAT and ASKAP as well as SKA
- CARTA uses a client-server architecture suitable for visualizing images with large file sizes (GB to TB)
- CARTA is mainly built in C++, TypeScript, and JavaScript and various third-party libraries

Note:
CARTA is fundamentally a **web application** with three main components:

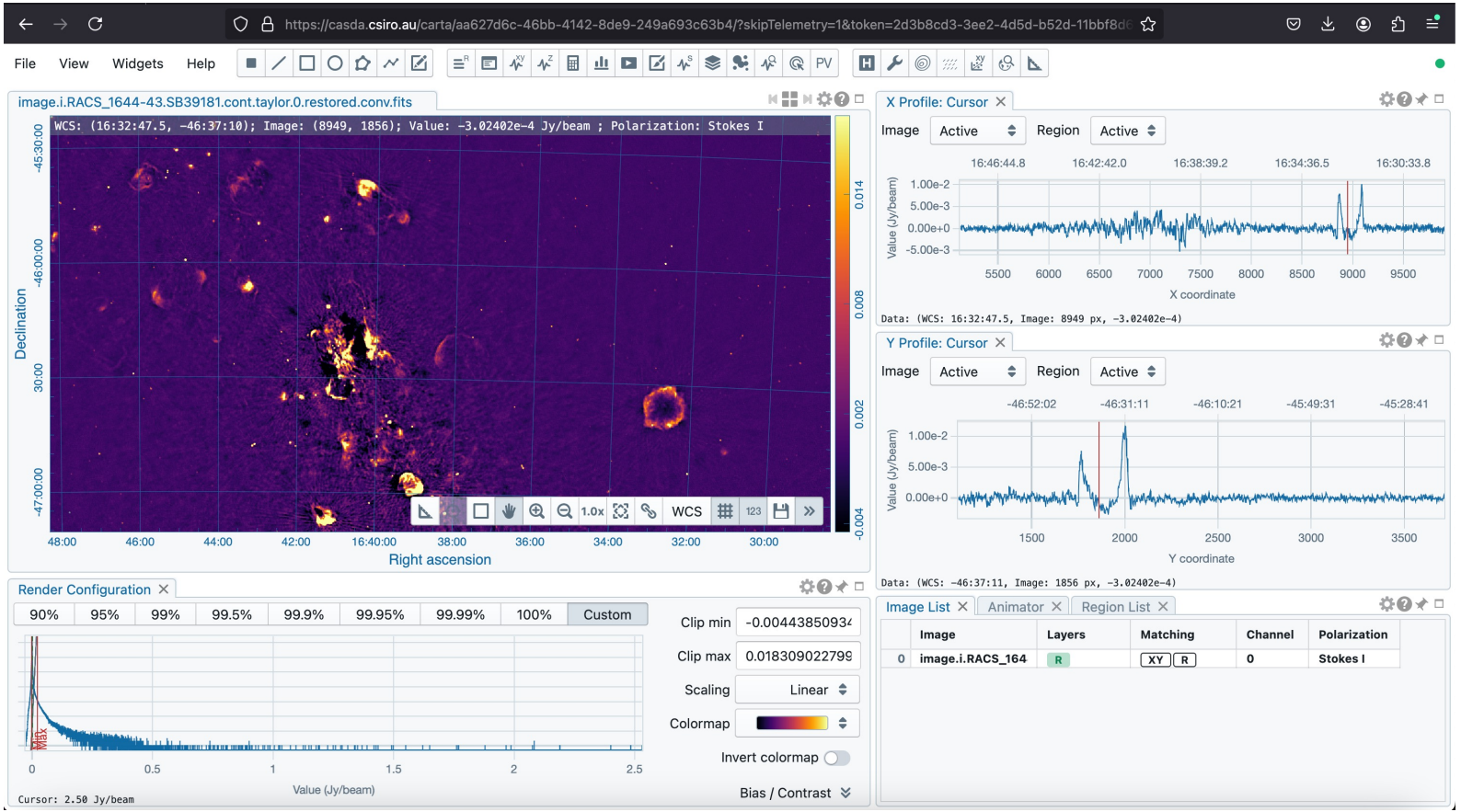
- carta_backend
- carta_frontend
- carta-controller



The source code of CARTA is available on GitHub.

CARTA web

https://casda.csiro.au/carta/...your_image_code...



Web-based data management tools

Project / Observatory specific tools (e.g.)

- CASDA: data.csiro.au/domain/casda (CARTA viewer)
- XMM-Newton: nxsa.esac.esa.int/nxsa-web/#search
- SciServer: www.sciserver.org
- SDSS: skyserver.sdss.org/dr18

Multi-mission / generic tools (e.g.)

- ESASky: sky.esa.int/esasky
- ESO science portal: archive.eso.org/scienceportal
- MAST: archive.stsci.edu
- CADC: www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en
- SIMBAD/VizieR: simbad.cds.unistra.fr – vizier.cds.unistra.fr

REMIR/ROS2 images handling flow

ROS2 PC
$ Camera_ROSS *Img.fits*

REMIR PC
$ Camera_REMIR *Img.fits*

La Silla

MCS Producer/DB PC
1. Read *Img.fits* header
2. Log subset kwds to DB
3. Compress – bzip2
4. Notify Mainsrv (BO)
5. Receive OK
6. Send to BO (status flag)

Bologna

MCS Mainsrv/DB
1. Insert/update transfer log
2. Receive *Img.fits.bz2* – bunzip2
3. Define ownership via PI kwd
4. Log Images to DB (pub. + priv.)
5. Astrometry.net
6. Log astrometry footprint
7. Produce image preview
8. Copy *Img.fits.gz* etc. to repo folder based on DATE-OBS

`$ myimgslog [options] Img.fits`

Defaults from **conf. file**

REMIR/ROS2 images handling Main ingredients

rem.oas.inaf.it/DB

Custom

- MCS: github.com/gcalderone/MCS
- DIF / SID: github.com/lnicastro/DIF – github.com/lnicastro/SID
- MyRO: github.com/lnicastro/MyRO
- Catalogues: catsweb.oas.inaf.it

+ G. Calderone - OATS

Public

- DB server ⇒ MySQL (MariaDB)
- Web server ⇒ Apache
- Languages ⇒ C++, PHP, HTML5, CSS3, JavaScript

Main JS packages: JQuery, *Aladin Lite, JS9, amCharts*

DB course: *astrordbms.oas.inaf.it*

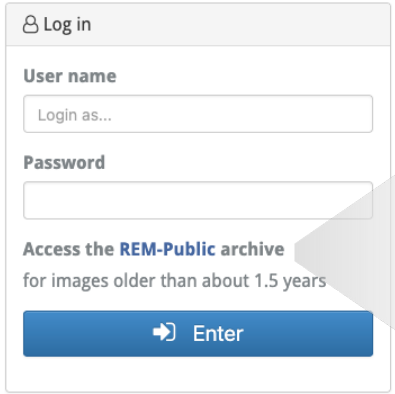
Images browsing

INAF ISTITUTO NAZIONALE DI ASTROFISICA NATIONAL INSTITUTE FOR ASTROPHYSICS

REM Data Archive

rem.oas.inaf.it/DB

Live images counting: ROSS/2: **1824837** + **125620** dark/bias + **79131** flat frames / REMIR: **3894483** + **173634** flat frames



Subscribe

Browse the ROS2 and REMIR observation log

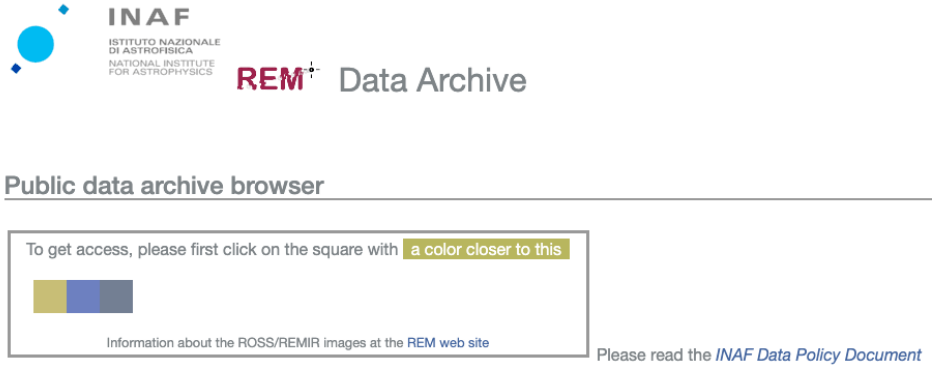
Get the **ROS2** master BIAS and FLAT

Please read the *INAF Data Policy Document* | « **REM** live with the t-REM-o-meter and WEBcam »

REM images browser:
Read the usage permission rules.

Credits

Written by L. Nicastro
Version: 0.9b, 12-Oct-2017



Images browsing

INAF
ISTITUTO NAZIONALE DI ASTROFISICA
NATIONAL INSTITUTE FOR ASTROPHYSICS

REM Data Archive

rem.oas.inaf.it/DB

Live images counting: ROSS/2: **1824837** + **125620** dark/bias + **79131** flat frames / REMIR: **3894483** + **173634** flat frames

Log in

User name

Login as...

Password

Access the **REM-Public** archive
for images older than about 1.5 years

Enter

Subscribe

Browse the ROS2 and REMIR observation log

Get the **ROS2** master BIAS and FLAT

Please read the *INAF Data Policy Document* | « **REM** live with the t-REM-o-meter and WEBcam »

REM images browser:
Read the usage permission rules.

Credits

Written by L. Nicastro
Version: 0.9b, 12-Oct-2017

REM-ROS2 normalised and bias subtracted monthly Master Flats

→ **Master Bias**

Goto to year: **2013**, **2014**, **2015**, **2016**, **2017**, **2018**, **2019**, **2020**, **2021**, **2022**, **2023**

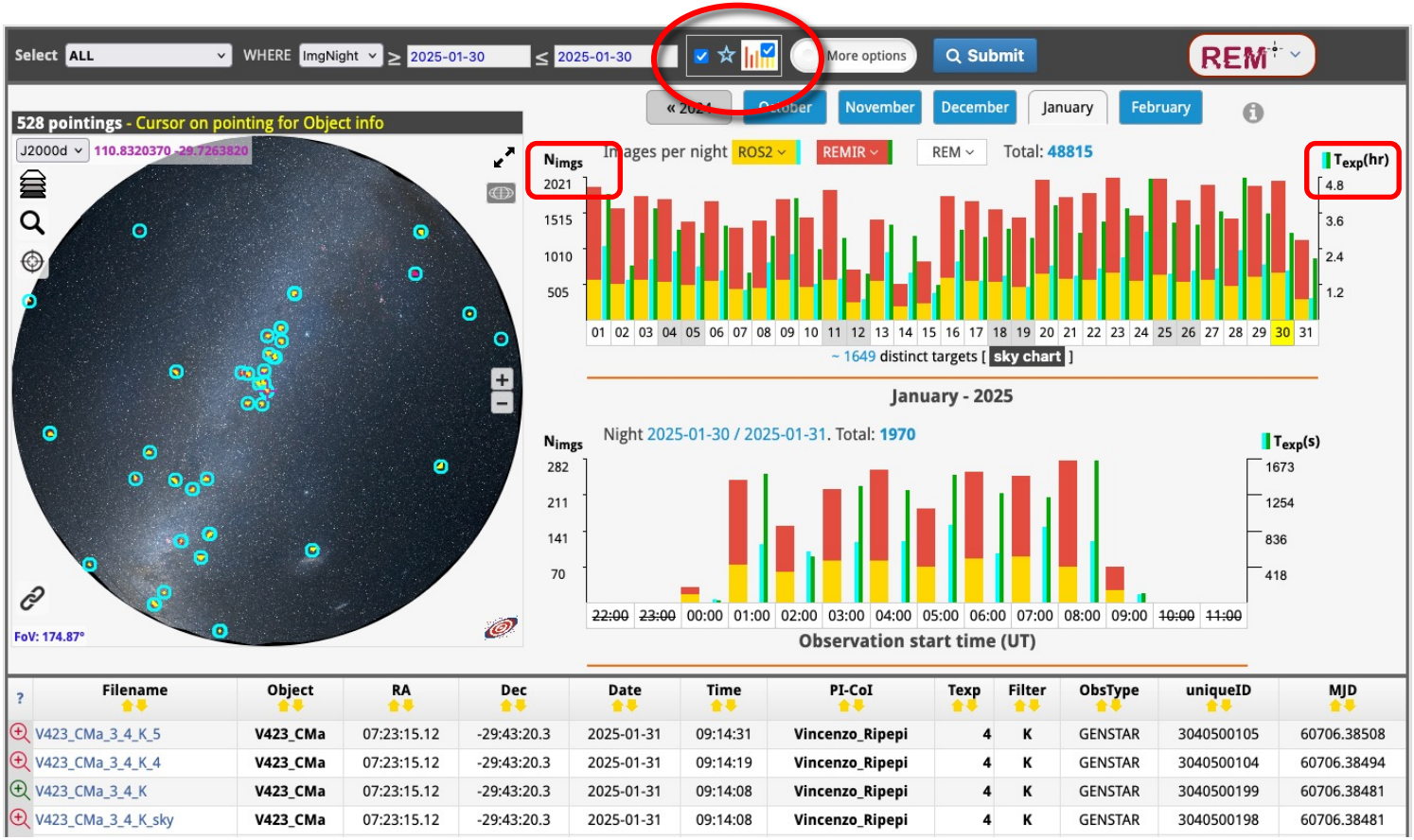
2025 ("s" version → gain = 1.0)

Month	BL (z)	BR (r)	UL (i)	UR (g)
January	JPG FITS	JPG FITS	JPG FITS	JPG FITS

2024 ("s" version → gain = 1.0)

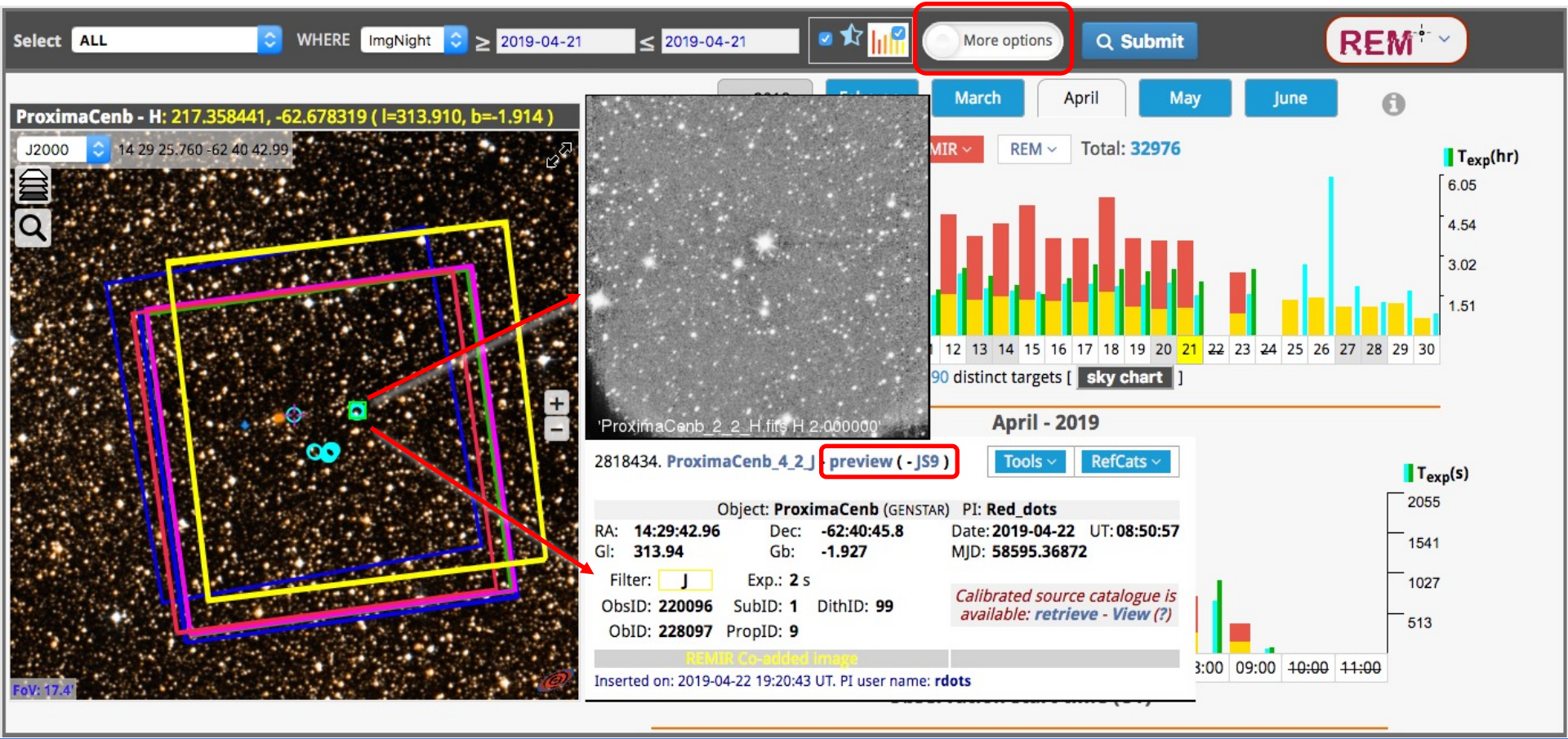
Month	BL (z)	BR (r)	UL (i)	UR (g)
December	JPG FITS	JPG FITS	JPG FITS	JPG FITS
November	JPG FITS	JPG FITS	JPG FITS	JPG FITS
October	JPG FITS	JPG FITS	JPG FITS	JPG FITS
September	JPG FITS	JPG FITS	JPG FITS	JPG FITS
August	JPG FITS	JPG FITS	JPG FITS	JPG FITS
July	JPG FITS	JPG FITS	JPG FITS	JPG FITS
June	JPG FITS	JPG FITS	JPG FITS	JPG FITS
May	JPG FITS	JPG FITS	JPG FITS	JPG FITS

Images browsing – *month/day + Aladin Lite*

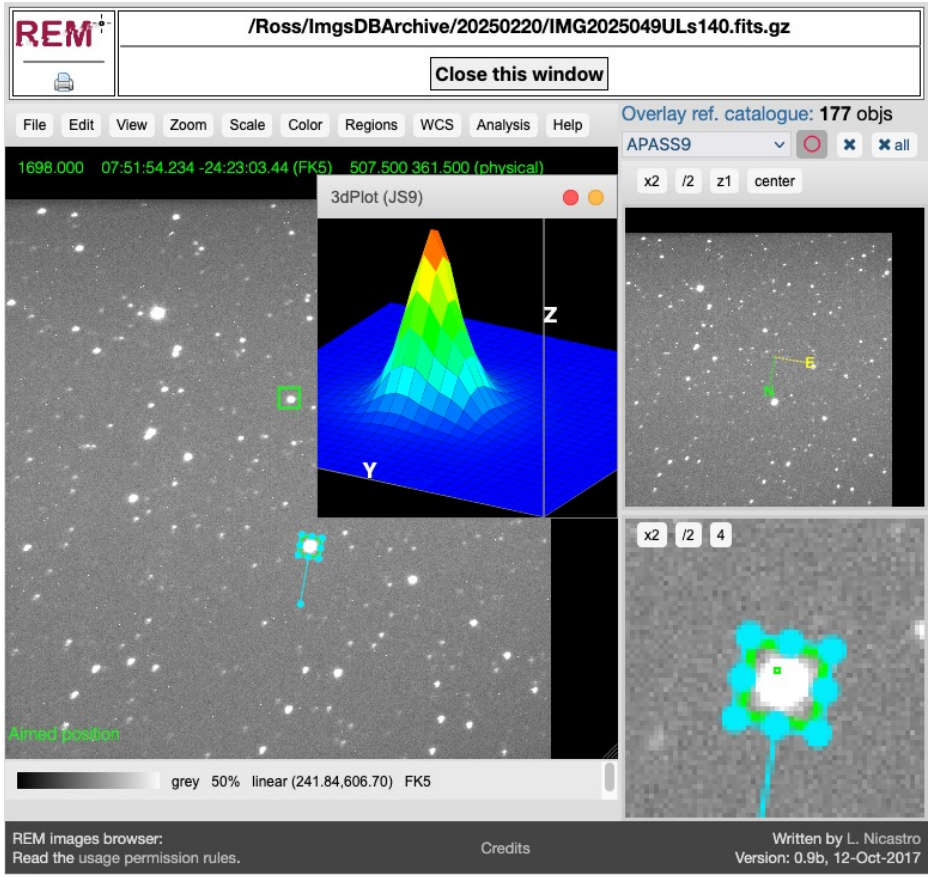


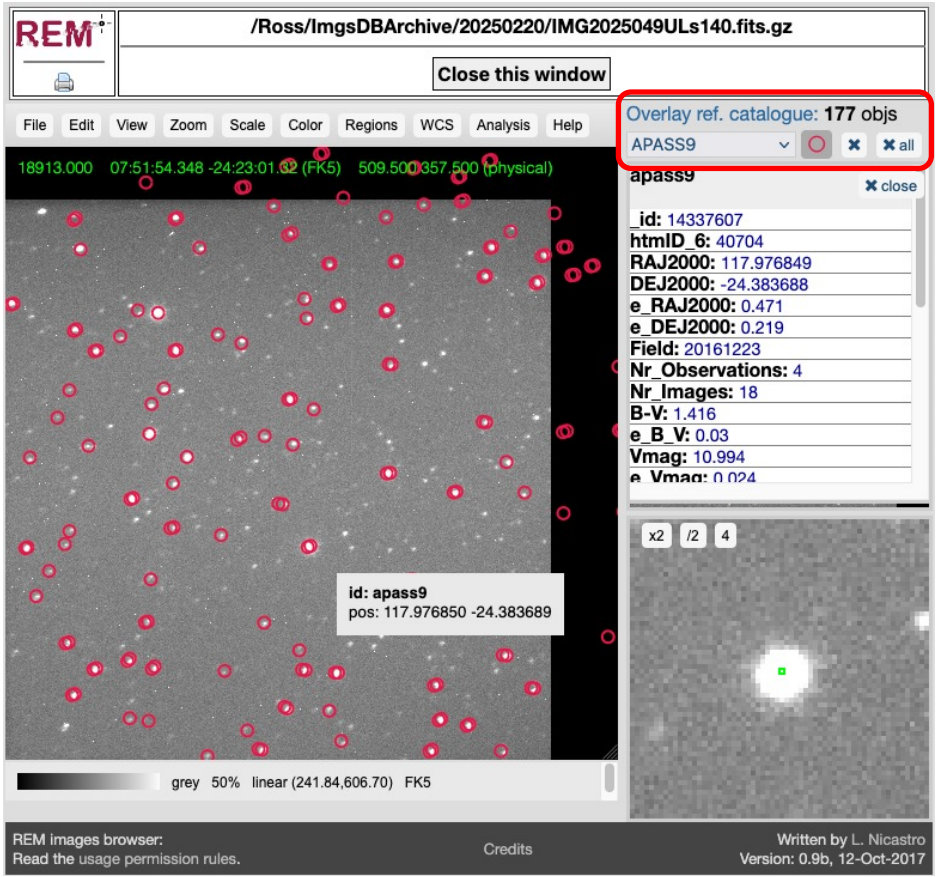
?	Filename	Object	RA	Dec	Date	Time	PI-CoI	Texp	Filter	ObsType	uniqueID	MJD
	V423_CMa_3_4_K_5	V423_CMa	07:23:15.12	-29:43:20.3	2025-01-31	09:14:31	Vincenzo_Ripepi	4	K	GENSTAR	3040500105	60706.38508
	V423_CMa_3_4_K_4	V423_CMa	07:23:15.12	-29:43:20.3	2025-01-31	09:14:19	Vincenzo_Ripepi	4	K	GENSTAR	3040500104	60706.38494
	V423_CMa_3_4_K	V423_CMa	07:23:15.12	-29:43:20.3	2025-01-31	09:14:08	Vincenzo_Ripepi	4	K	GENSTAR	3040500199	60706.38481
	V423_CMa_3_4_K_sky	V423_CMa	07:23:15.12	-29:43:20.3	2025-01-31	09:14:08	Vincenzo_Ripepi	4	K	GENSTAR	3040500198	60706.38481

Images browsing – *preview and details*



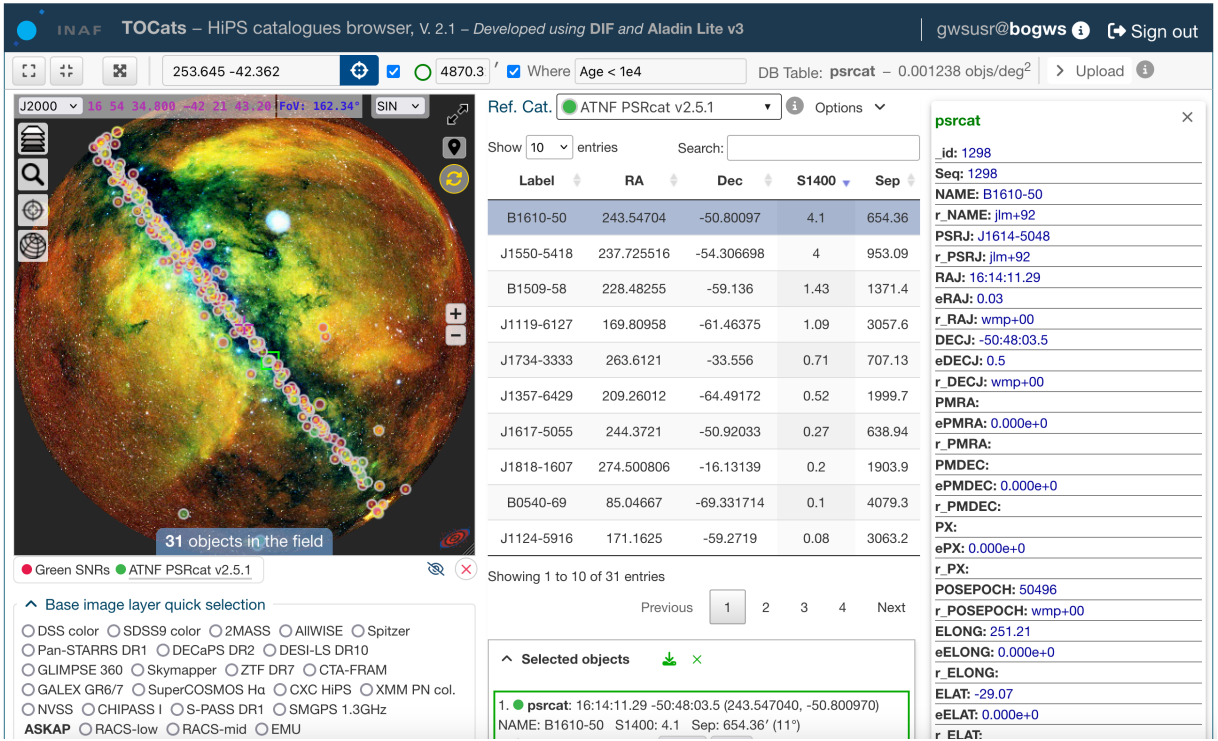
Images browsing – *JS9 + TOCats*





TOCats & TOCatsweb

A tool for fast access to astronomical surveys and object catalogues by using hierarchical scheme techniques to split the celestial sphere, in particular the HEALPix scheme.



catsweb.oas.inaf.it

TOCats three main components

1. A relational DBMS, with hierarchically structured tables containing various types of catalogues and associated metadata
2. A web service that makes images and catalogues easily queryable and accessible from a browser or a script
3. A repository of Hierarchical Progressive Surveys (HiPS) with images and object density maps

TOCats & TOCatsweb – *main components*

Built on

- MySQL: dev.mysql.com
- DIF: github.com/lnicastro/DIF
- Aladin Lite: aladin.cds.unistra.fr/AladinLite/doc
- Hipsgen: aladin.cds.unistra.fr/hips
- DataTables: datatables.net
- amCharts: www.amcharts.com

… and

- Web server ⇒ Apache
- Languages ⇒ C++, PHP, HTML5, CSS3, JavaScript, Python
- + JS packages ⇒ JQuery, JS9, custom code

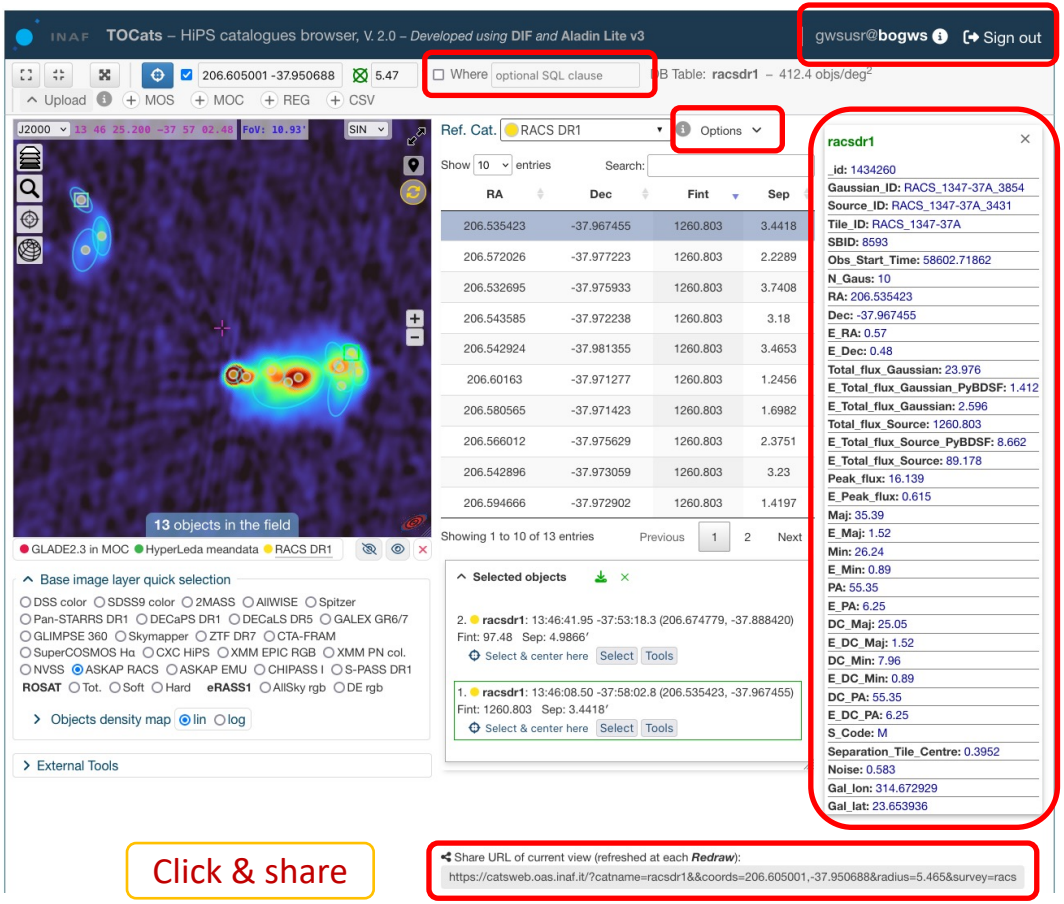
- ~ 110 public catalogues (Opt/IR/NIR/Radio/UV/X-γ)
- ~ 100 private catalogues
- ~ 20 TB (raid 1)
- 10 projects

TOCats & TOCatsweb – *capabilities*

- Combine DIF managed catalogues with HiPS public and private surveys and
- Handle extended sources / footprint
- Retrieve the catalogue's data in multiple formats from the interface or a script (API)
- Accessible from VO tools (e.g. TOPCAT)
- One click external tools access
- Add your project and implement your plug-in tools to handle any type of data
- Take advantage of all the Aladin Lite v.3 capabilities:
  - Multi-surveys view (jpg, png, fits)
  - FITS files (etc.) drop-in
  - HIPS2FITS tool
  - Etc.

Note:

Aladin Lite is being upgraded continuously



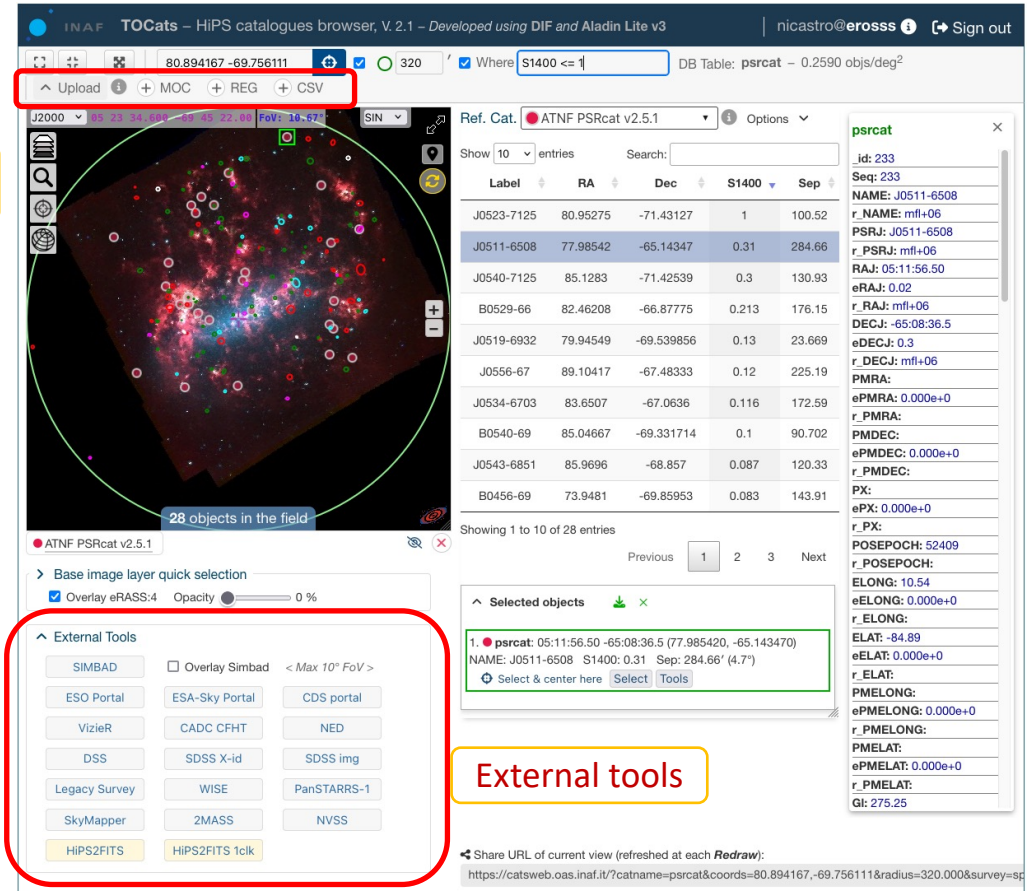
TOCats & TOCatsweb – *use cases*

Import local MOC, REG and CSV files

LMC

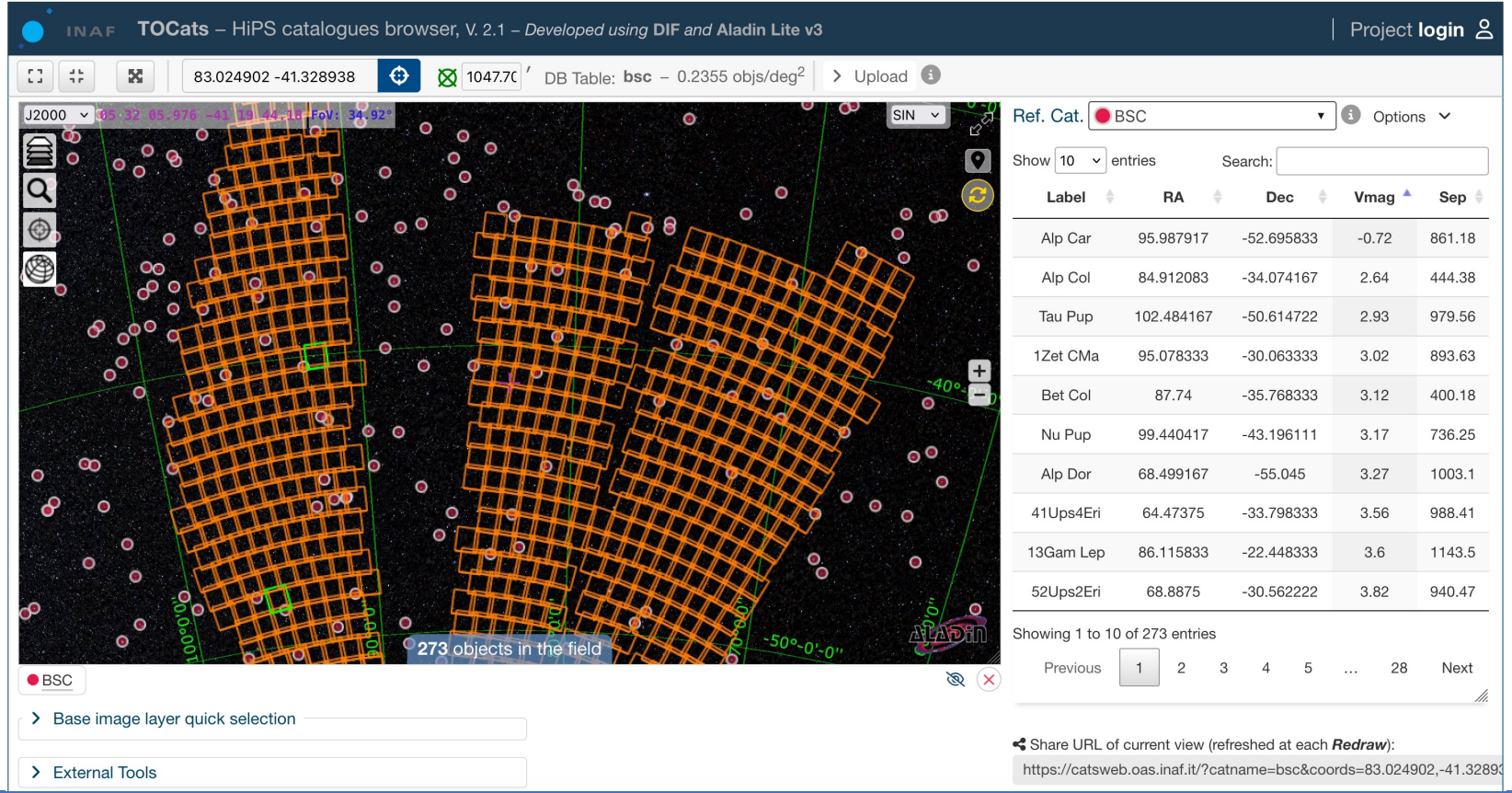
ds9 REGion file

```
ellipse(71.11583,-67.97030, 0.022,0.015,80) #text={J0444-6758}
tag={candidate} color=red

ellipse(71.80072,-69.32124, 0.032,0.05,320) #text={J0447-6918}
tag={MCSNR} color=cyan

circle(72.10489,-67.00686, 0.045)           #text={J0448-6700}
tag={MCSNR} color=green

ellipse(72.391666,-69.05944, 0.031,0.029,0) #text={J0449-6903}
tag={MCSNR} color=cyan

circle (72.330,-69.33889, 0.027)            #text={J0449-6920}
tag={MCSNR} color=green

ellipse(72.55166,-68.30133, 0.126,0.10,105) #text={J0450-6818}
tag={candidate} color=red

circle (72.6125,-70.8375, 0.053)            #text={J0450-7050}
tag={MCSNR} color=green

ellipse(72.94051,-67.29002, 0.081,0.047,60) #text={J0451-6717}
tag={candidate} color=magenta

… … …
```



TOCats & TOCatsweb – *use cases*

Euclid-NISP J footprint



TOCats & TOCatsweb – *use cases*

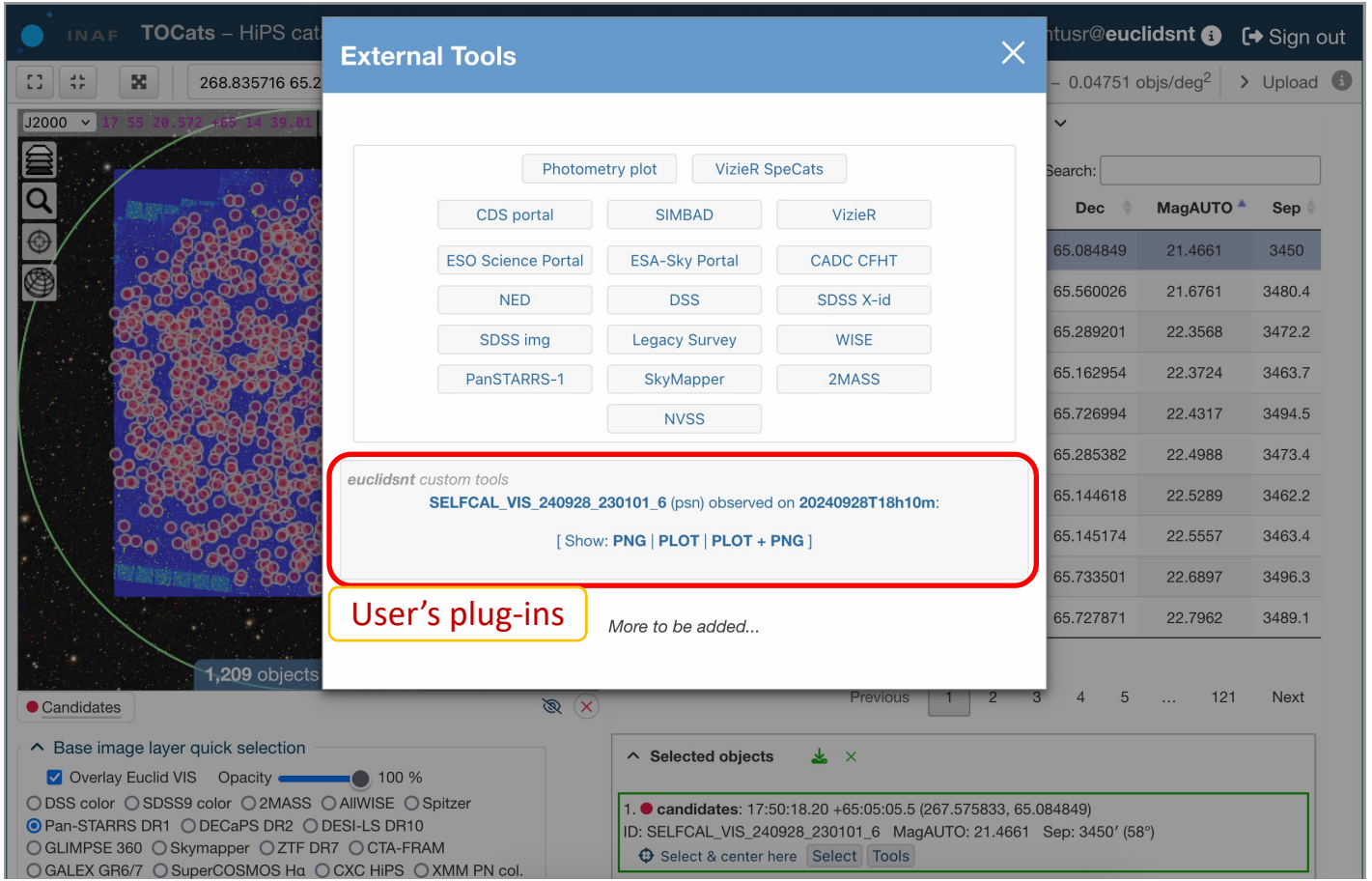
Transient sources in the Euclid SelfCal field

VIS SelfCal

TOCats & TOCatsweb – *use cases*

Transient sources in the Euclid SelfCal field

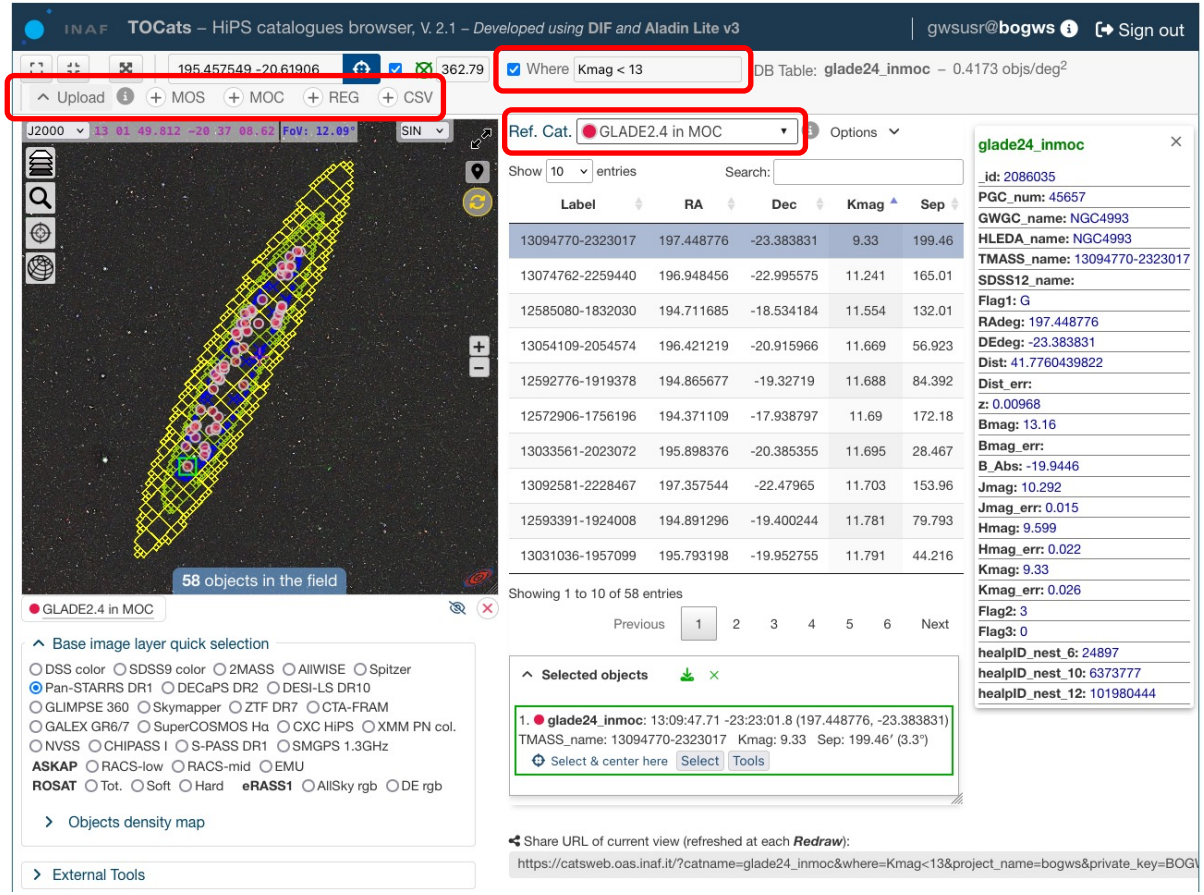
VIS SelfCal



TOCats & TOCatsweb – *use cases*

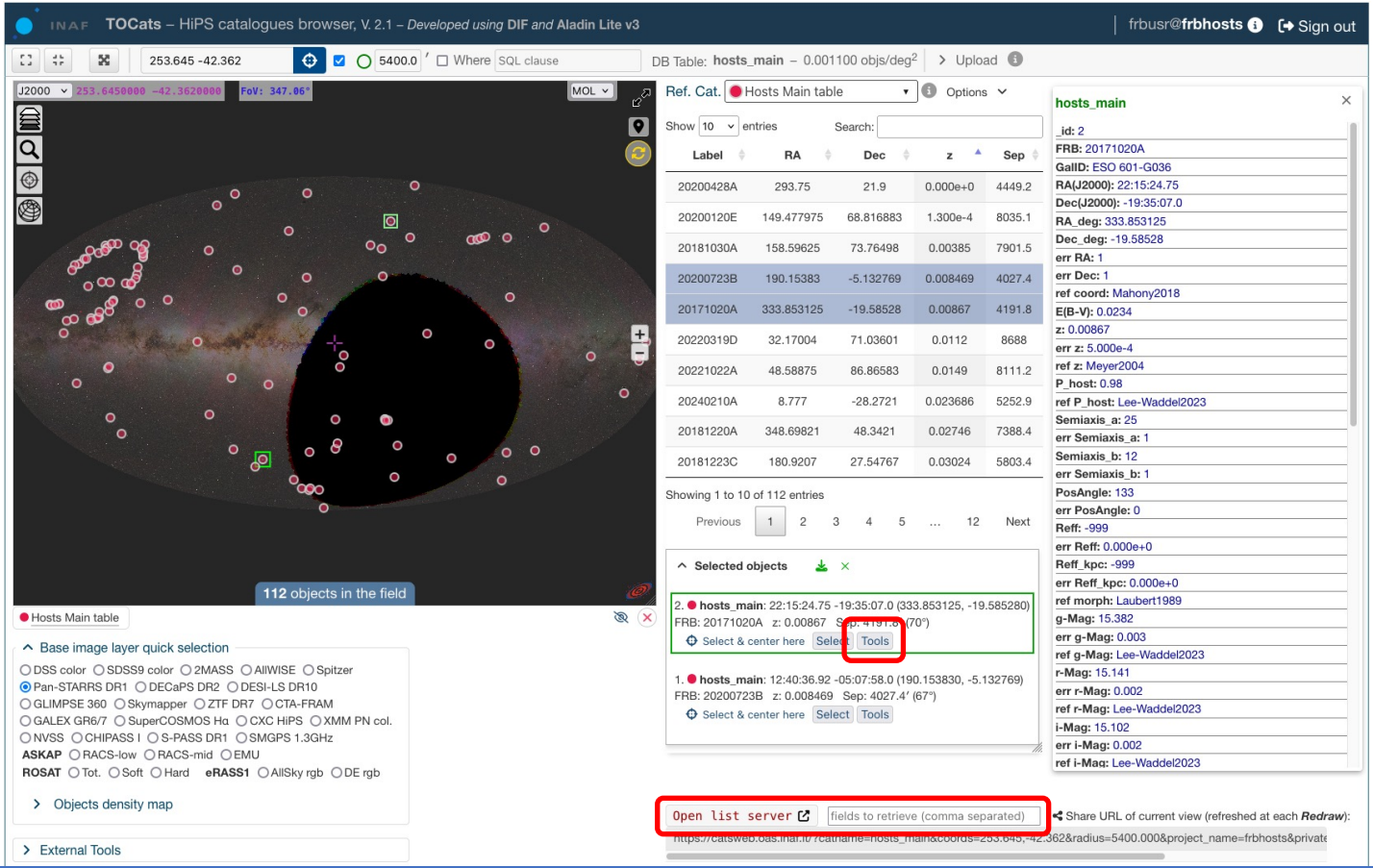
1. Import GW event Multi-Order Skymap (MOS), either local of giving its URL

2. Identify NED-LVS and GLADE 2.4 sources in a given probability region



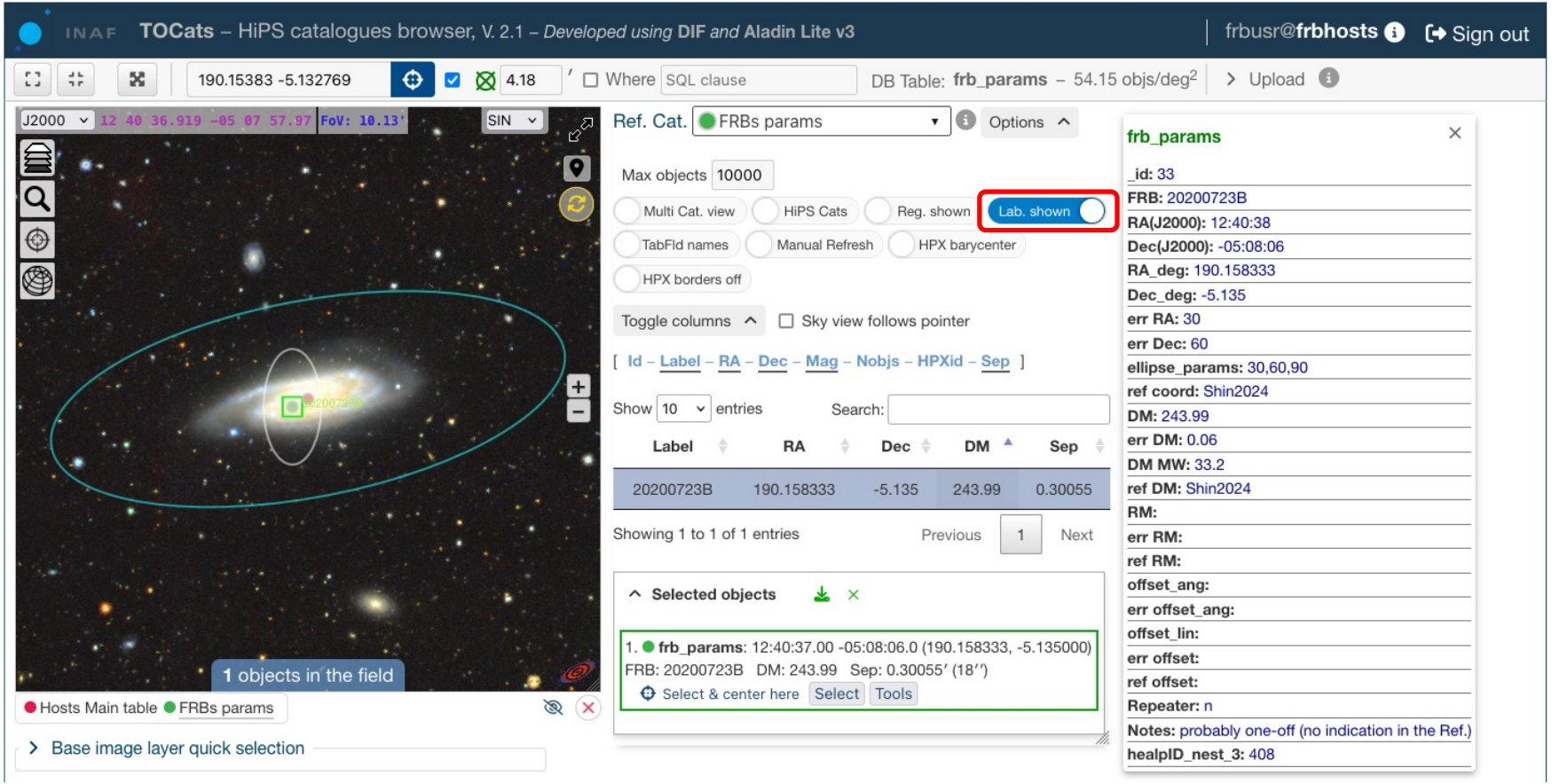
TOCats & TOCatsweb – *use cases*

FRBs and their host galaxies



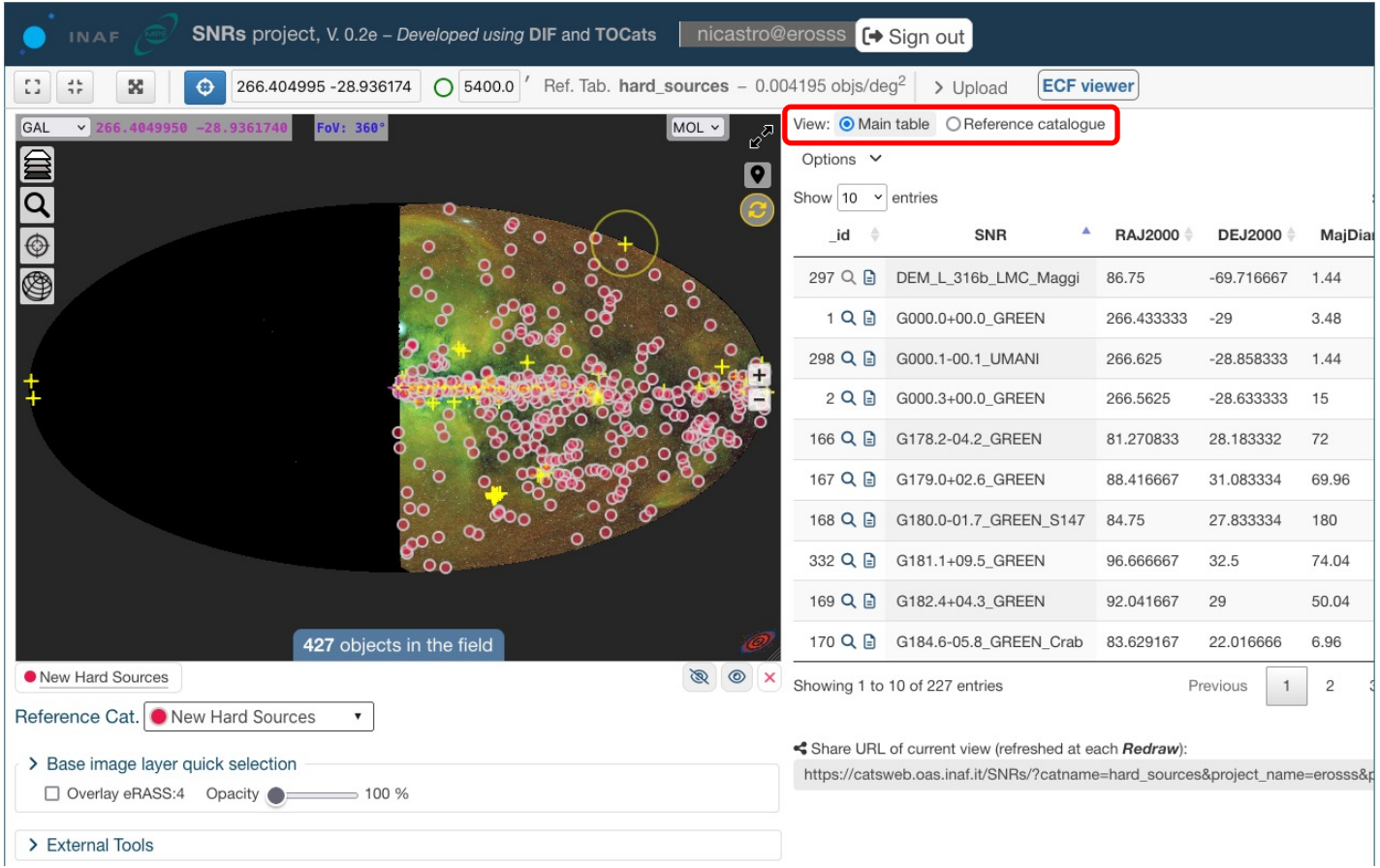
TOCats & TOCatsweb – *use cases*

FRB 20200723B



TOCats & TOCatsweb – *use cases*

eROSITA SNRs



VSTbrowse

grawita.inaf.it/VSTbrowse/

A tool for investigating ESO-VST multi-epoch visits of GW sky regions and search for transients

Preliminary tasks

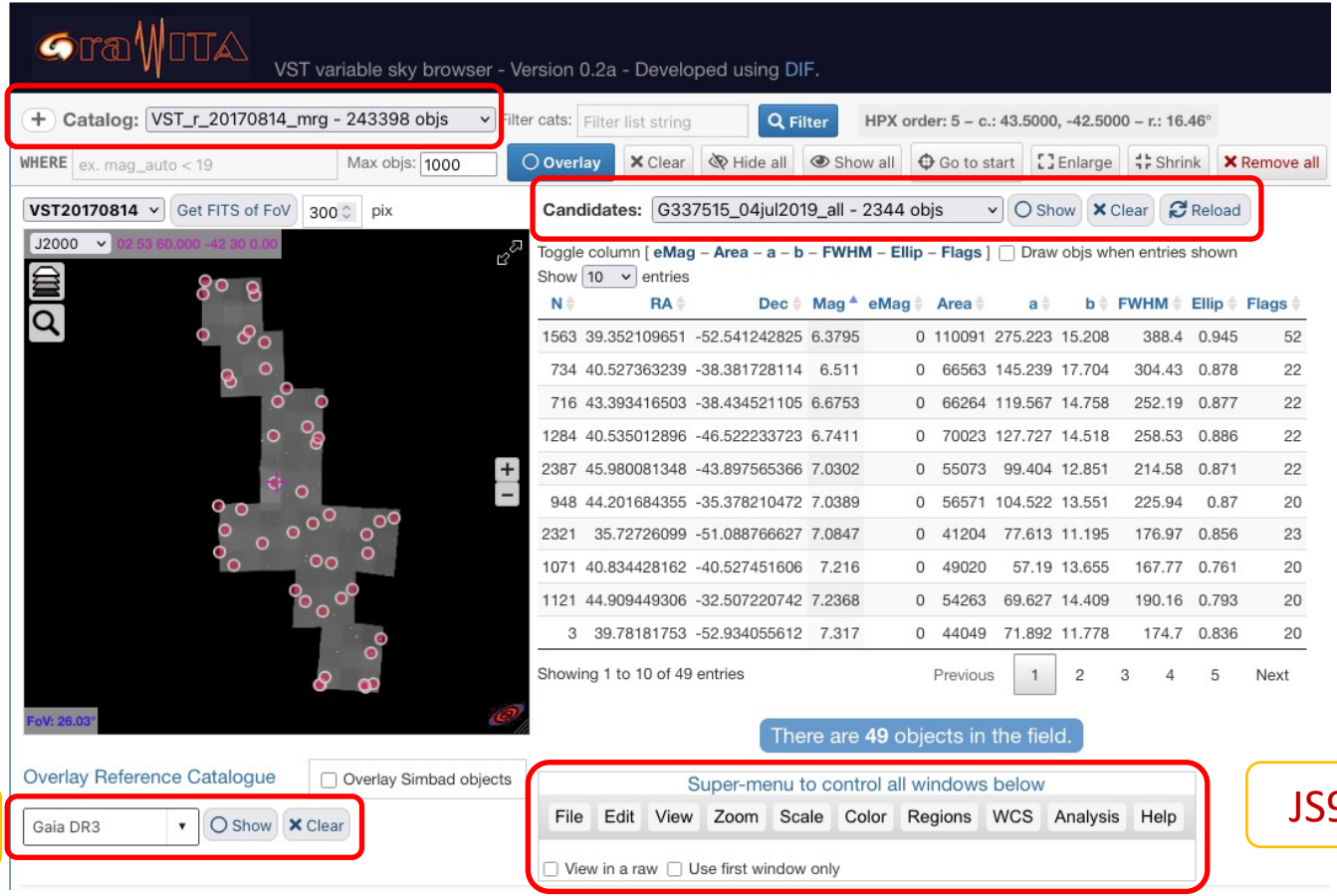
- Hipsgen the collected VST sky survey
- DB insert extracted source catalogues
- DB insert identified variable objects (transient source candidates)

Web tool

- Browse sky: compare with existing surveys and catalogues
- Extract and inspects FITS stamps (JS9)
- Plot transients LC
- Mark and annotate candidates

GRAWITA: VSTbrowse – GW170814

grawita.inaf.it/VSTbrowse/GW170814/

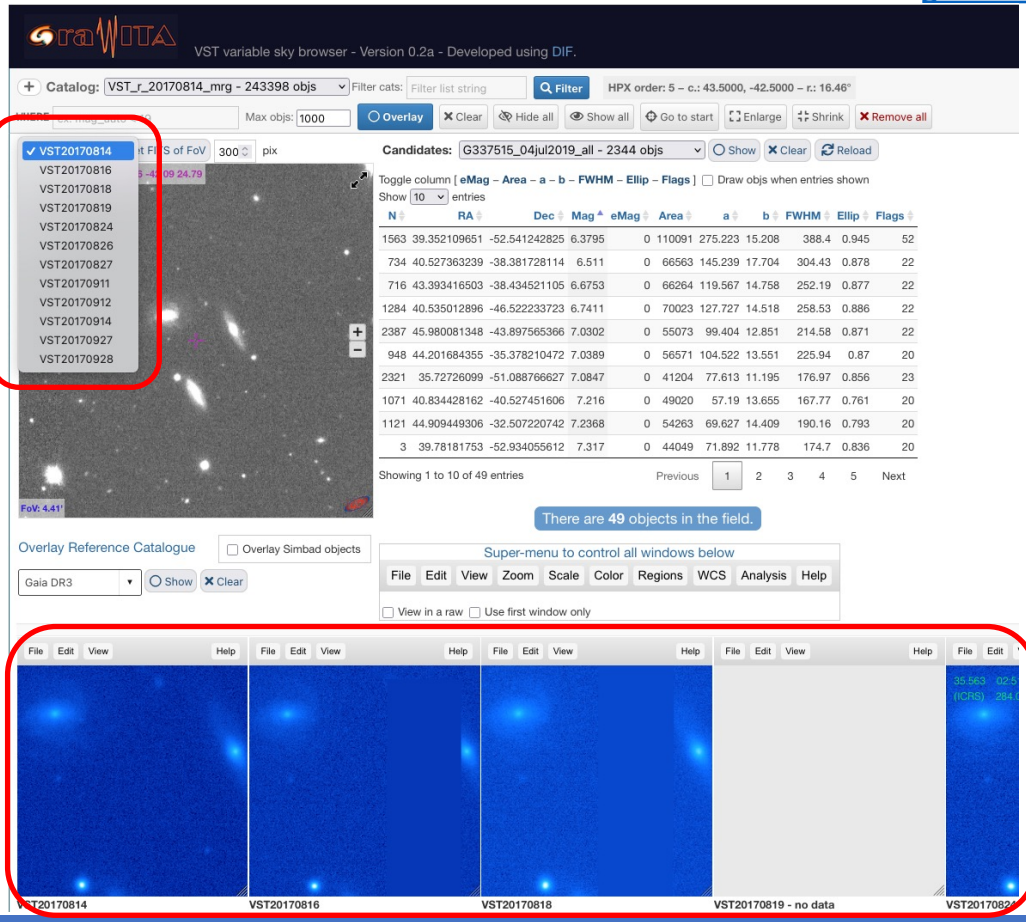


TOCats

JS9

GRAWITA: VSTbrowse – GW170814

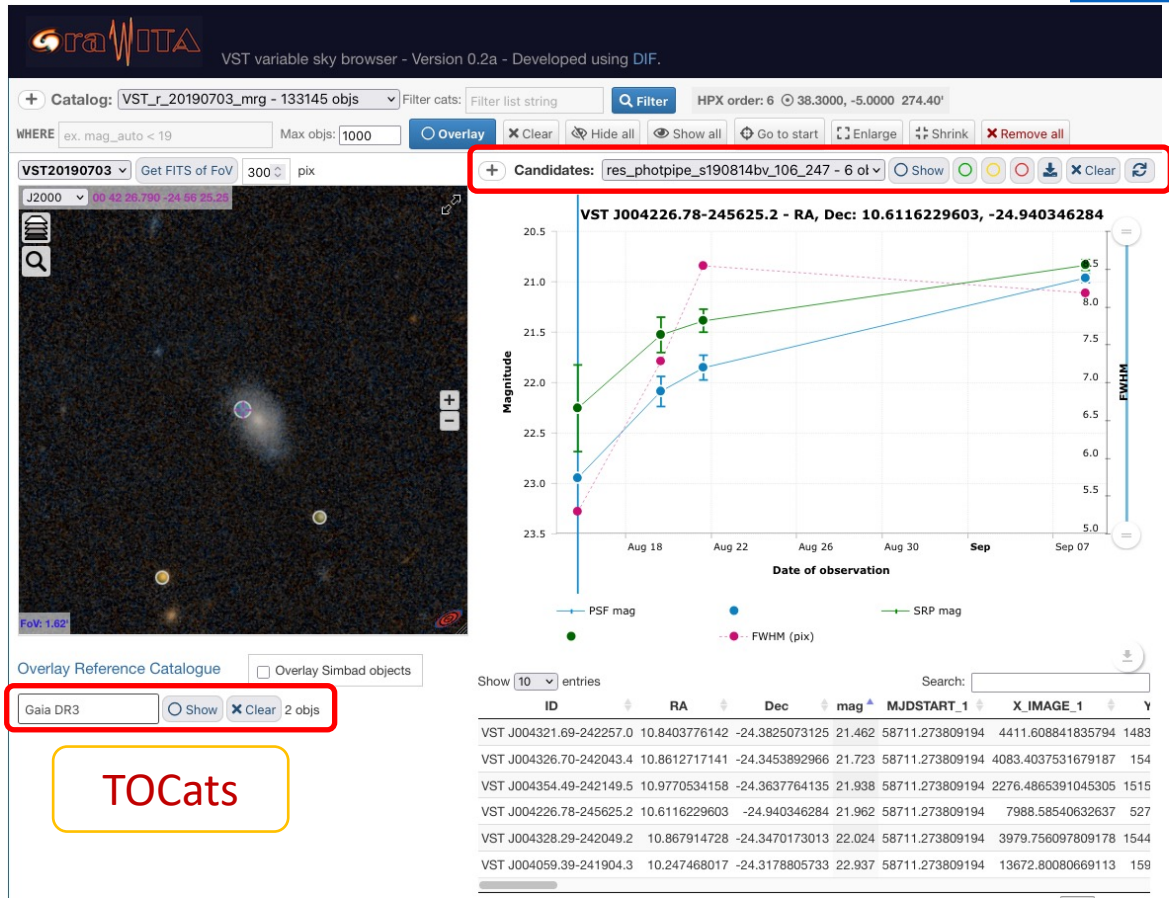
grawita.inaf.it/VSTbrowse/GW170814/



JS9

GRAWITA: VSTbrowse – G337515

grawita.inaf.it/VSTbrowse/G337515/

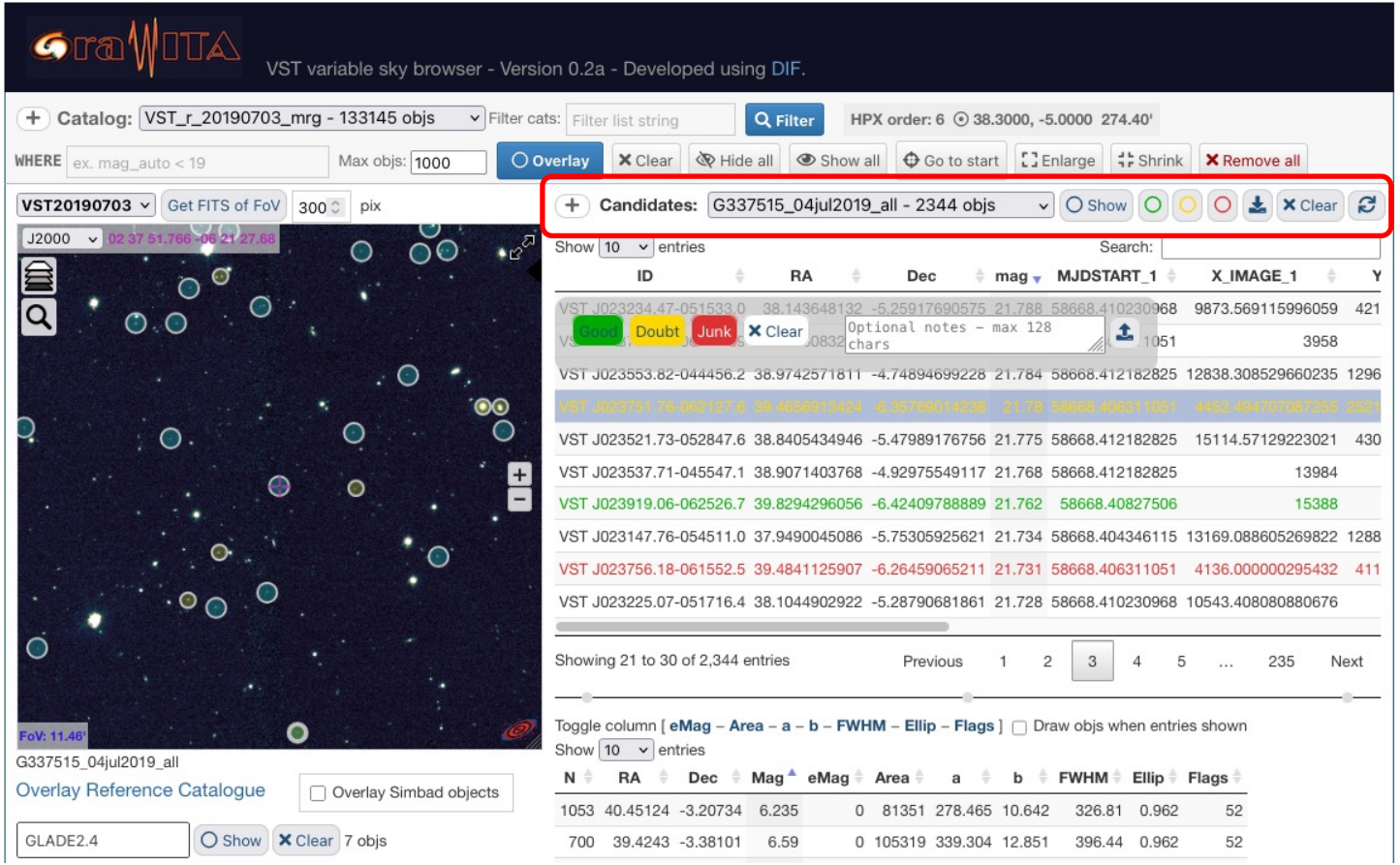


amCharts

TOCats

GRAWITA: VSTbrowse – G337515

grawita.inaf.it/VSTbrowse/G337515/



JS9 – web-based DS9

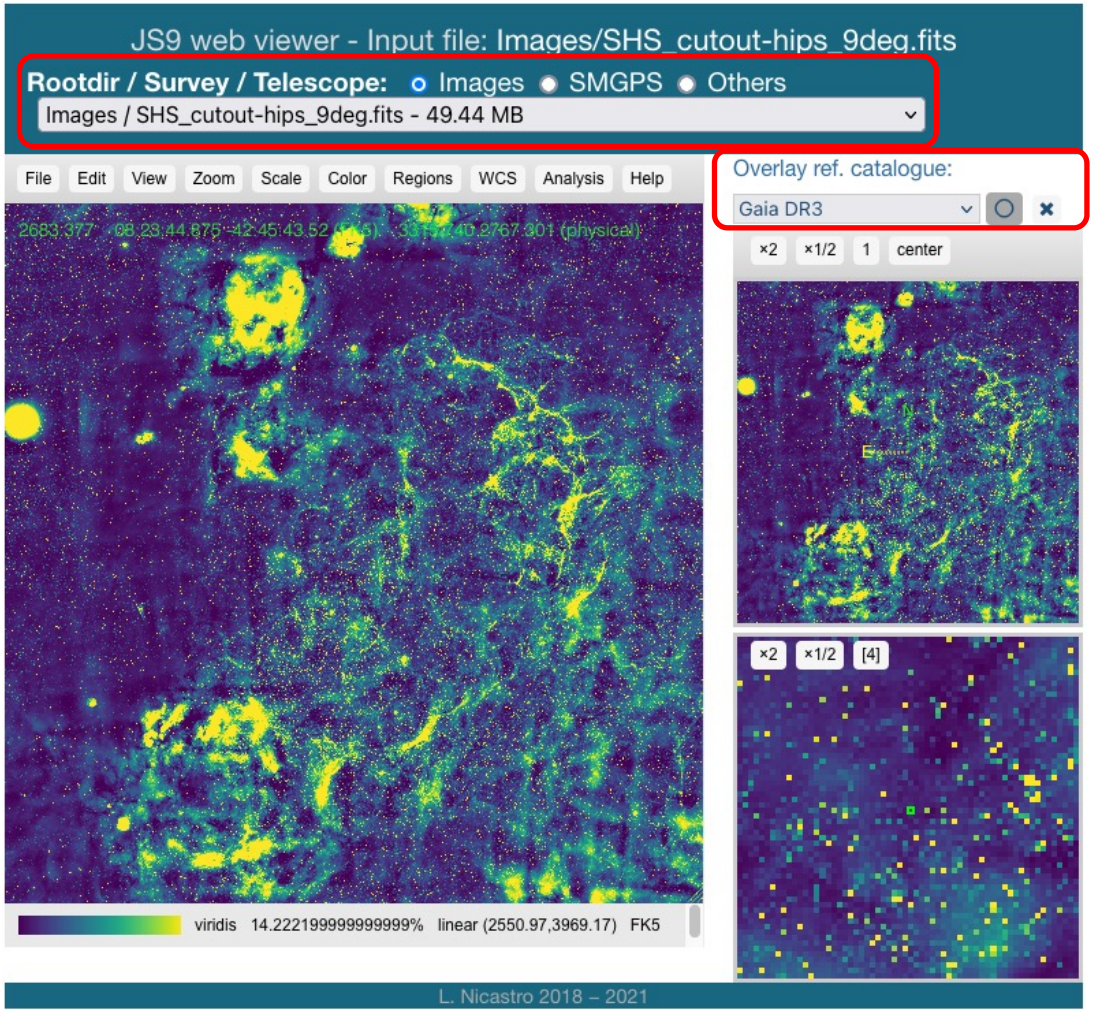
js9.si.edu

A custom implementation

cats.oas.inaf.it/js9/

Direct server image link

cats.oas.inaf.it/js9/?file=Images/SHS_cutout-hips_9deg.fits



JS9 – web-based DS9

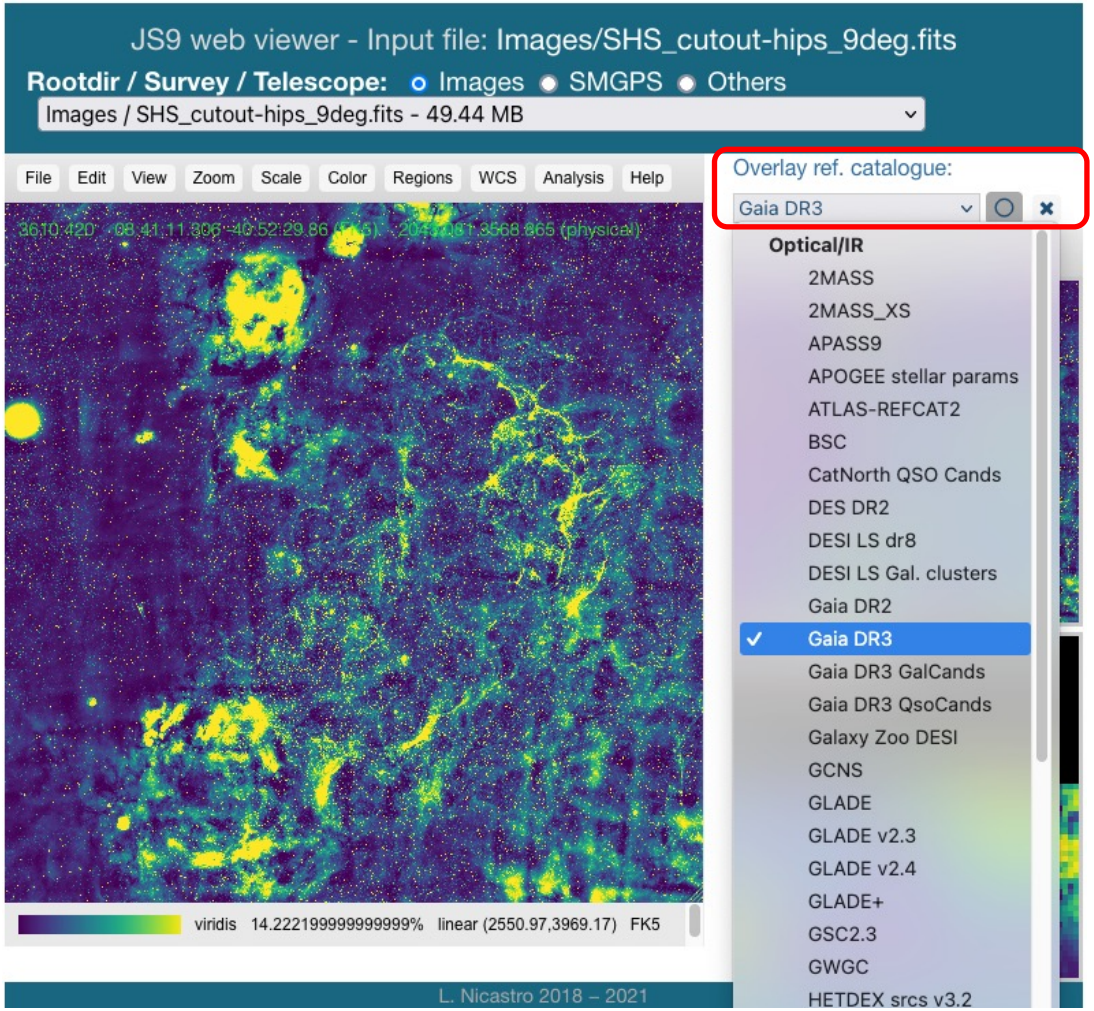
js9.si.edu

A custom implementation

cats.oas.inaf.it/js9/

Direct server image link

cats.oas.inaf.it/js9/?file=Images/SHS_cutout-hips_9deg.fits



JS9 – web-based DS9

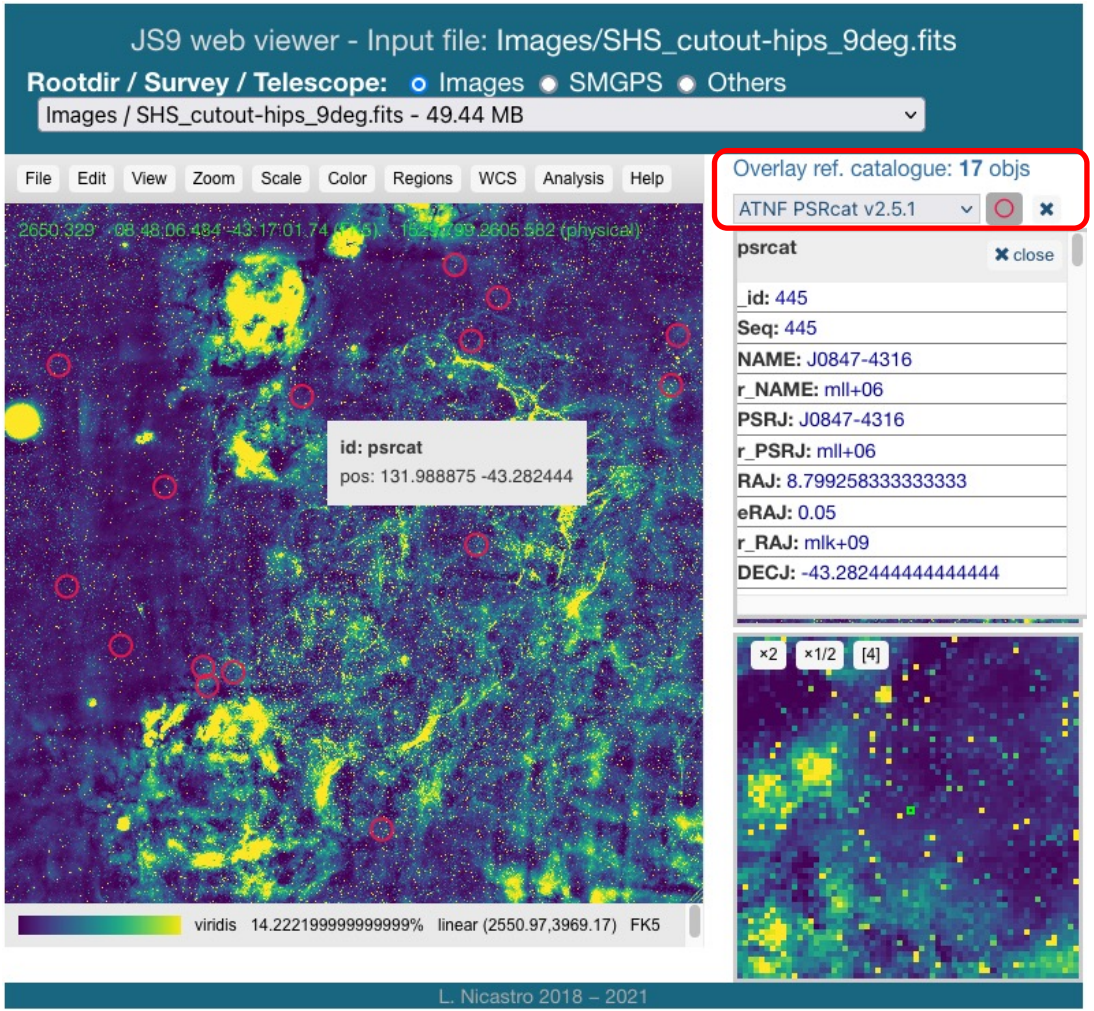
js9.si.edu

A custom implementation

cats.oas.inaf.it/js9/

Direct server image link

cats.oas.inaf.it/js9/?file=Images/SHS_cutout-hips_9deg.fits



Discussion items for this meeting (part.)

- Development of web tools (JS libs, node.js, WebGL, …)
- Development of *Instrument agnostic* libraries / pipelines
- Co-existence of relational and No-SQL DBs
- Move the code, not the data (products ok)
- Is it Python good for all? It is the present, and the future?
- Semantic queries ⟹ AI for astronomy?

Projects and Facilities drive (custom) SW development strategies, but:

- Does (or shall) INAF have the resources / capabilities to invest in SW development / homogenization?
- Is INAF attractive for SW engineers?

rem.oas.inaf.it/DB

catsweb.oas.inaf.it

cats.oas.inaf.it/js9