swissuniversities

UNIVERSITÉ DE GENÈVE FACULTÉ DES SCIENCES

EPFL

n|w Fachhochschule Nordwestschweiz

AstroORDAS

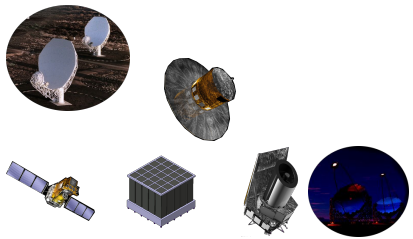
Astronomical Open Research Data Analysis Services

Swiss Open Research Data Grants Track B: Establish projects

https://www.astro.unige.ch/astroordas/

Volodymyr Savchenko

Carlo Ferrigno, Gabriele Barni, Mohamed Maharga, Stéphane Paltani,
Andrii Neronov, Oleg Kalashev, Andrei Variu, François Antoine Morier-Genoud, Emma Tolley, Jean-Paul Kneib,
Andreas Wassmer, Lukas Gering, André Csillaghy



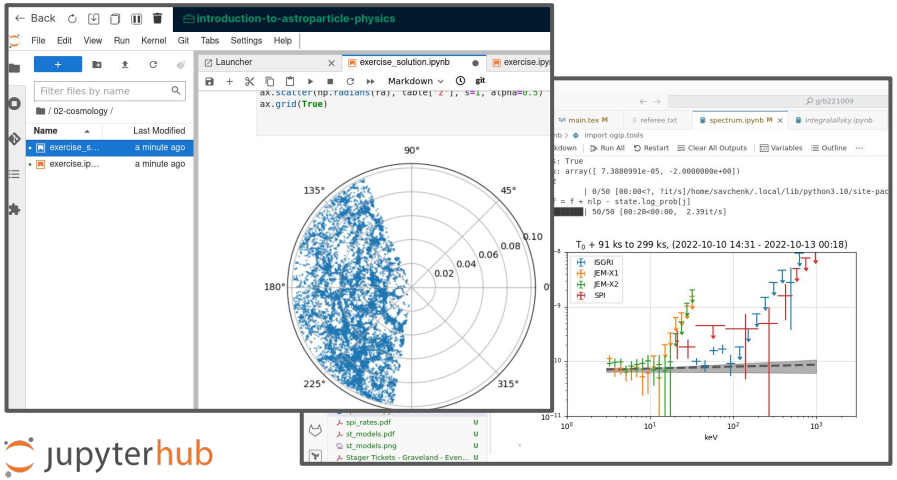
Swiss SKA days 2024
EPFL, Geneva Campus Biotech

https://github.com/oda-hub/

Platforms for Coding and for using Tools and Workflows

Development environment allows to write code, e.g. in python.
But this is never enough, need existing assets: data and data transformation tools.

environment to develop research code



jupyterhub

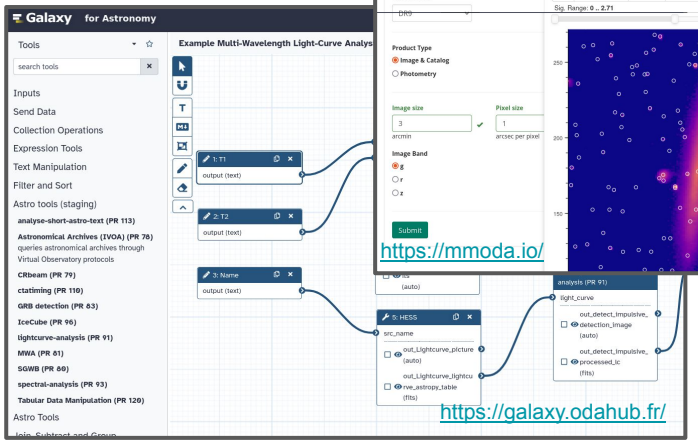
VRE at CERN

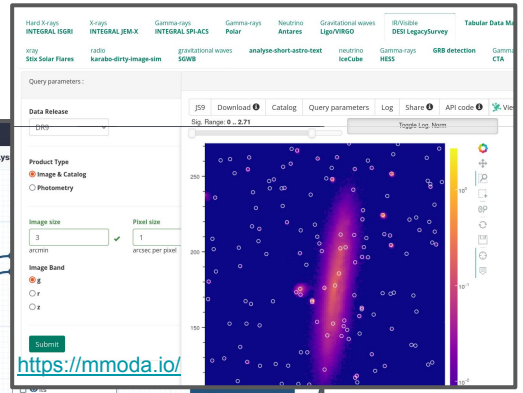
renku

colab

OPEN OnDemand

analysis tools and workflows





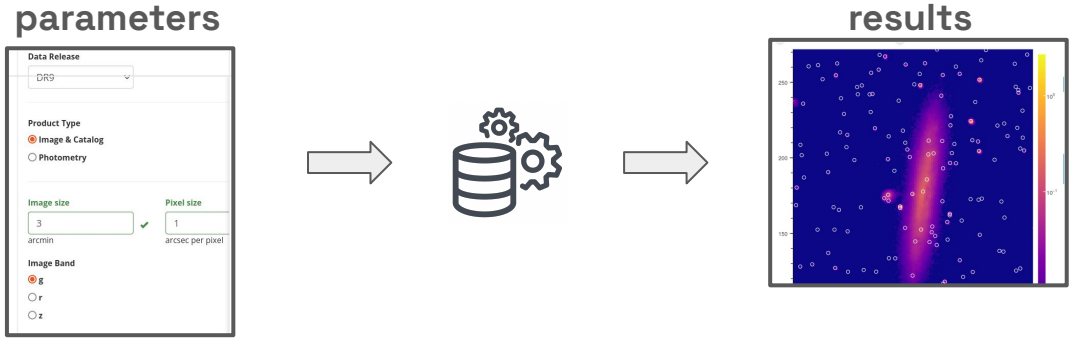
MMODA
Multi-Messenger Online Data Analysis

AstroORDAS Workflows and **Tools** are **functions** producing **data** given some **input**. Workflows are linked look.

Tools may be accessible from dedicated UI or from an API from another platform.

An Analysis Service in AstroORDAS

“Services” in AstoORDAS are live services producing research data.
Not data storage services, compute, TAP, etc - although all these are used in alignment with the live services.



Typically are containers annotated as **functions** with **input** and **output**.
They be realised live a **web service endpoint** (e.g. swagger), a **tool in a workflow language** (e.g. CWL) or a **function** general purpose language, and can be converted between these forms.

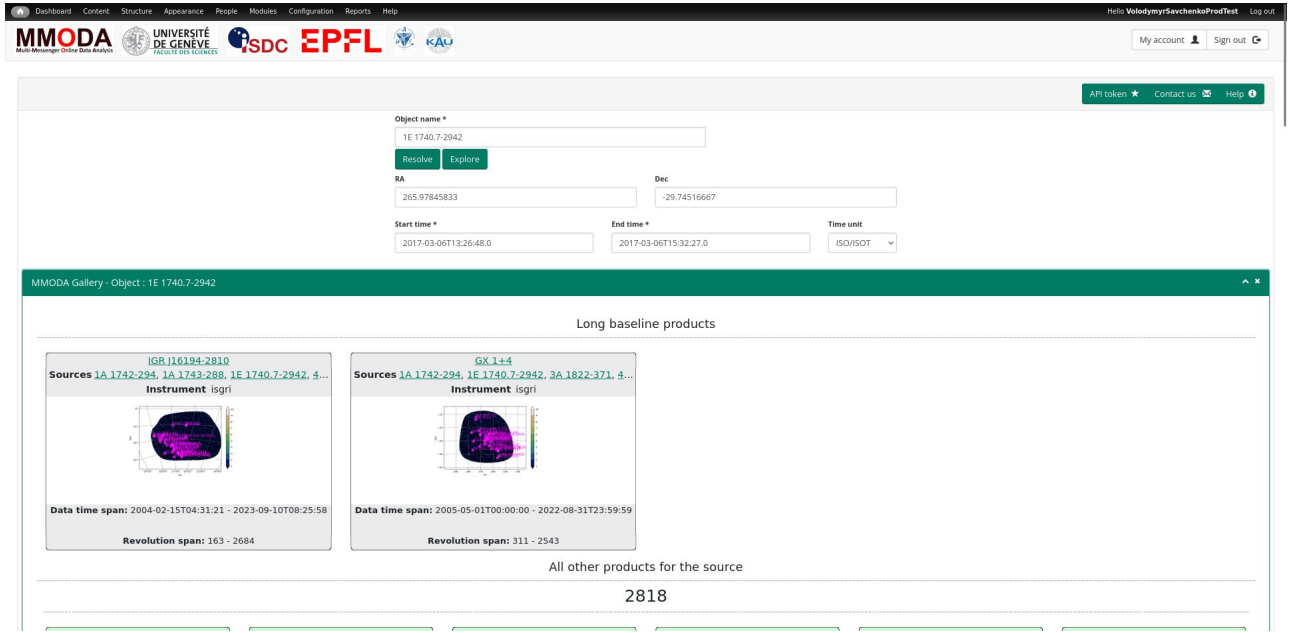
VO provides interfaces to manage data but not functions (note though Execution Planner).

Compare to AWS Lambdas, WikiFunctions, the spirit of serverless compute.
If you know about **catalog of analysis functions** a thing in academia let me know! The closest I know is HEASoft, and Galaxy.

Functions with parameters are reproducible data products

Unlike static data, functions with no parameters can be simply recomputed (giving identical result) or recomputed with some changes, reproducing the result in new conditions.

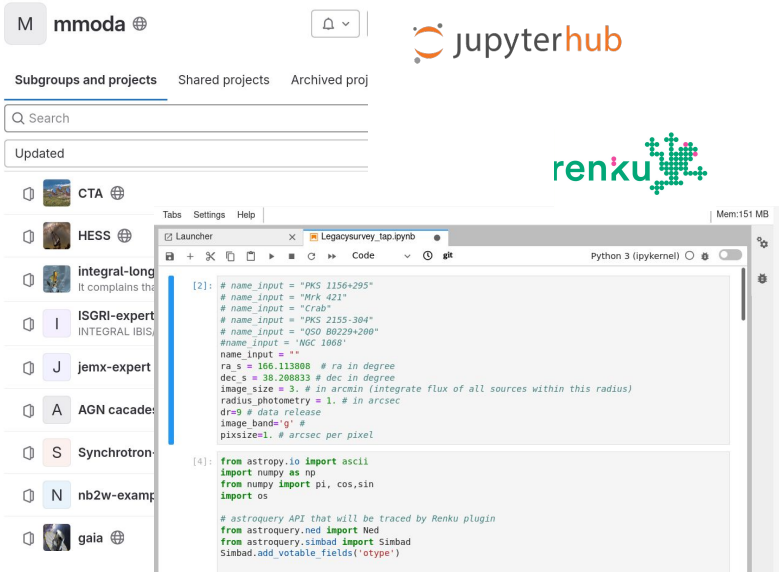
MMODA Gallery keeps track of user-selected computations done by MMODA services.

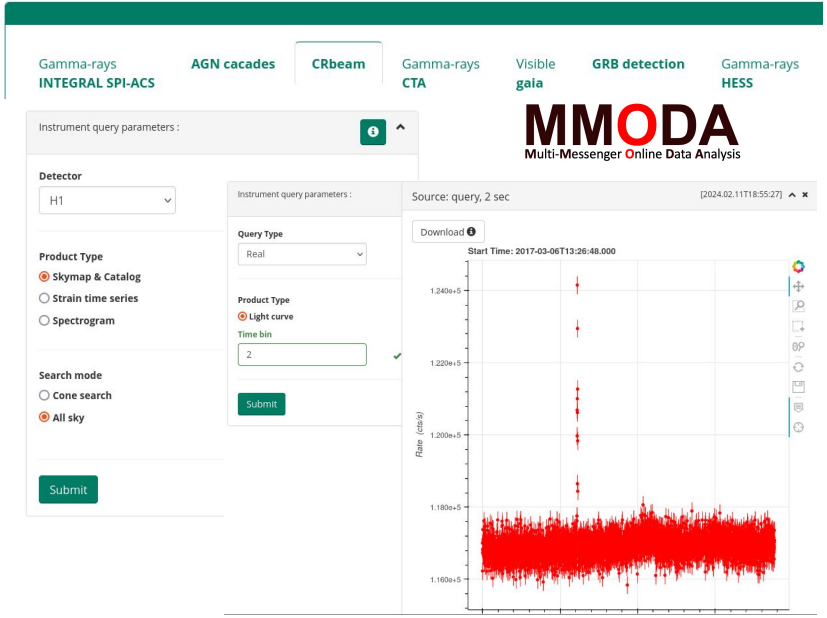


Service development environment

Developing tools is hard. Developing domain-specific analysis is very hard. **MMODA** answer is division of labor, we create JupyterLab, (in Jupyterhub, renkulab) development environment allowing **expert users** to **easily contribute analysis workflows**, adding semantics of **input parameters**, **output data**, and any additional dependencies.

We **do not reinvent** workflow language, but instead map **easy tool construction and update in a notebook** into other tool forms.





https://gitlab.renkulab.io/astronomy/mmoda/

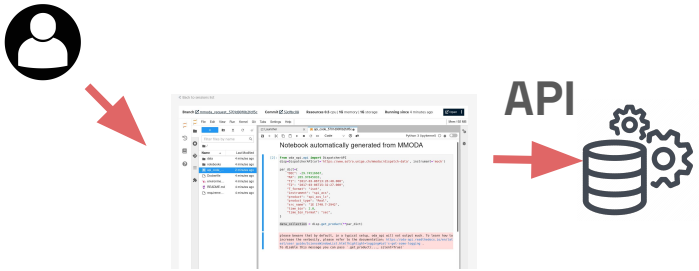
Compute and Data access

SDSC collaboration

Tools need access to storage and distributed compute.

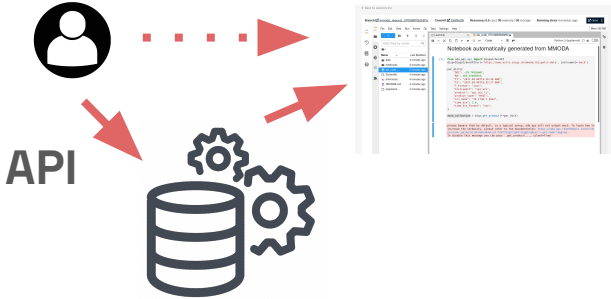
Two ways to access capacity outside simple code/tool:

1. Talking to a service by accessing **API from within interactive session**: dask cluster, S3 storage, HTTP or posix.
2. Launching the **entire session** or **batch tool** on the **resource next to the data**. Workflow from pure functions ensures controlled access to data, HTC.



We also developed runtime introspection intercepting external resource queries to create annotations.

This allows access to private and public infrastructures of the developers, with a focus on **CSCS**.



Publications integration:

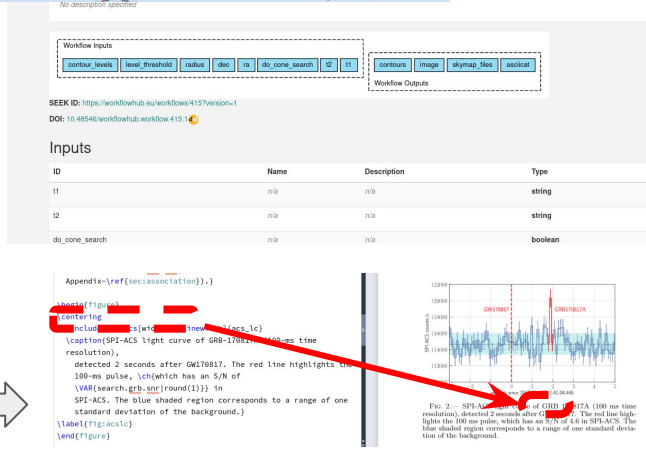
Paper **production:**

MMODA platform and the AstroORDAS project has features for publishing research assets.

- All **queries** are **reproducible**, and semantically annotated.
- **Publishing** research data but also **analysis tools**

We are working with “almost traditional” publishers and journals especially innovative like EPIScience.

Including tight data-paper connection in a form of a “live paper”.

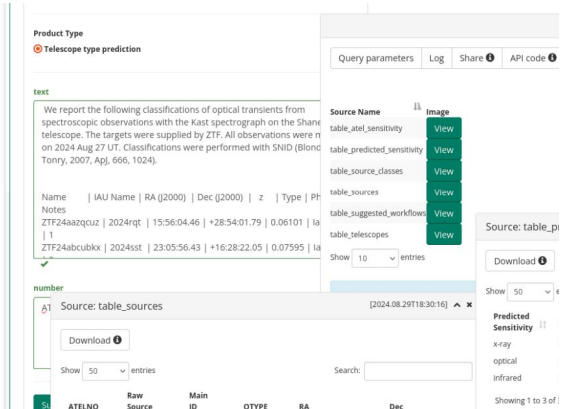


And also paper **consumption**:

We made a tool for **paper analysis**, extracting semantic data, based on ontologies from IVOA + expansion.

Based in this semantic data, made **predictions for telescope followup**.

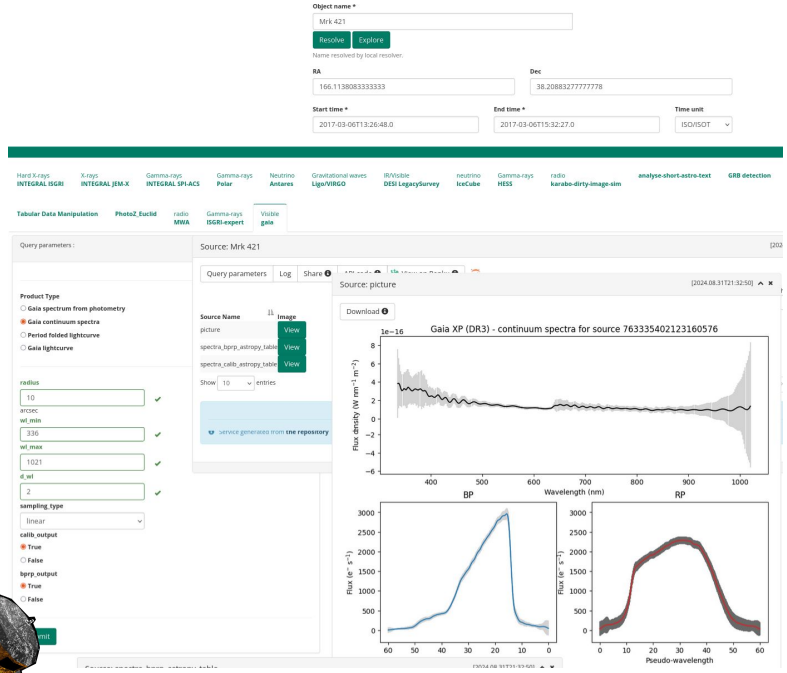
Publication analysis engine is itself a MMODA tool.

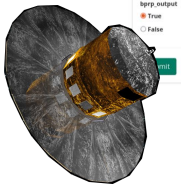


AstroORDAS new telescopes for MMODA and Galaxy

Gaia

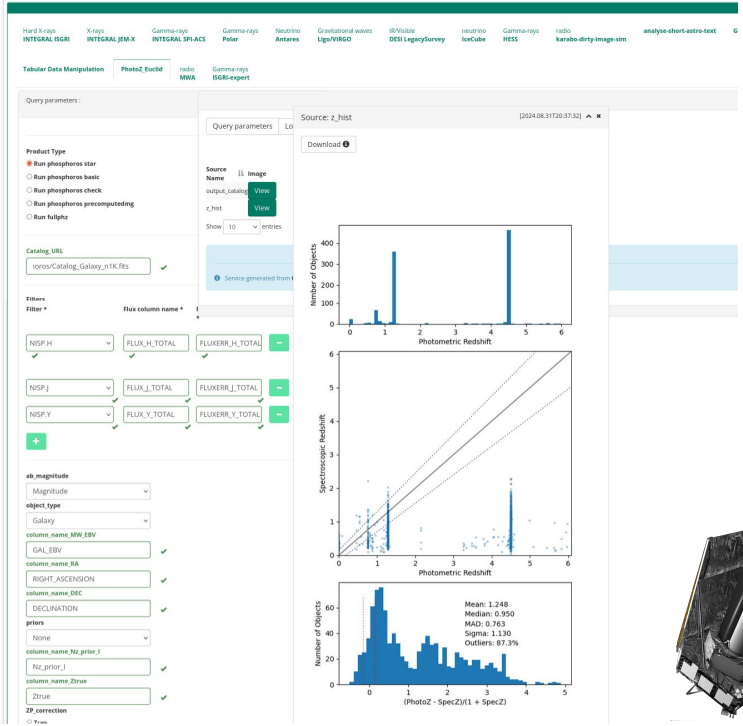
Milestone astrometric mission, we include tools for accessing is





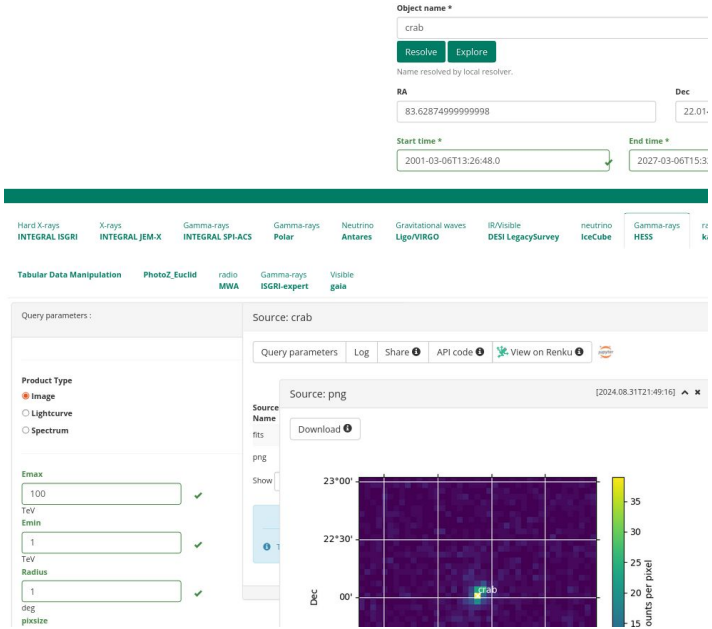
Euclid

Recently launched cosmology mission, we include tool computing Photometric Redshift.



Cherenkov Telescope Array Observatory (CTAO)

H.E.S.S.: older telescope, public data.



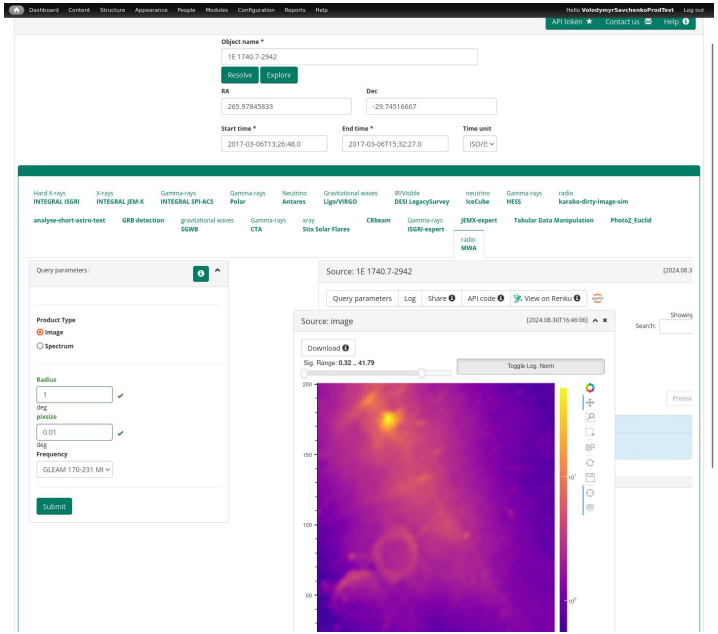
For CTAO, CRbeam generates model prediction and, Gammapy simulates synthetic data.

Here for a change Galaxy is shown.

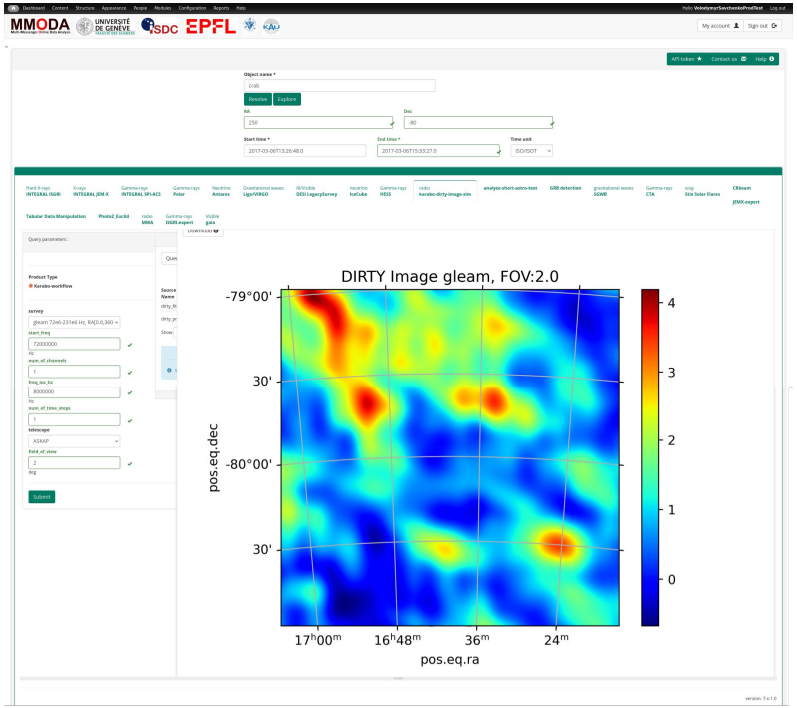
Planned to be used as one of the offered components in Science Data Challenge

AstroORDAS Radio

MWA GLEAM: an interface to **SkyView** incorporating it in **AstroORDAS/MMODA** and **Galaxy**



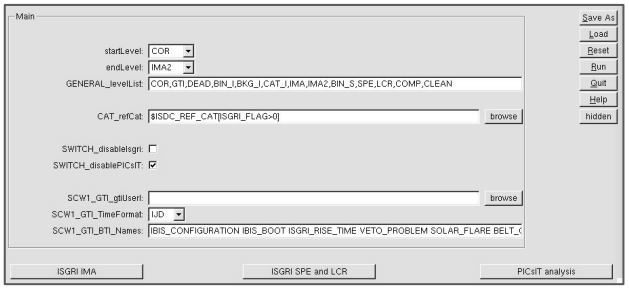
Karabo Simulation, dirty image.



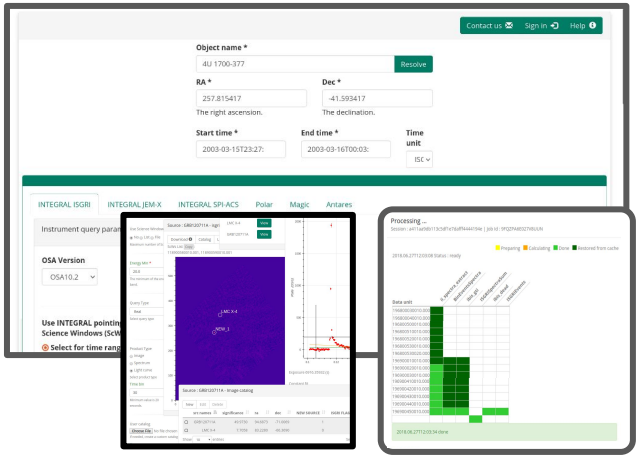
Use Case for new GUI for existing tooling

As much as the GUI for tools is needed (for simple analysis), it can be provided in Web platforms like MMODA.
Lot's of modern GUI Desktop Apps are web apps anyway (e.g. electron-based). Typically also supports mobile.

INTEGRAL OSA

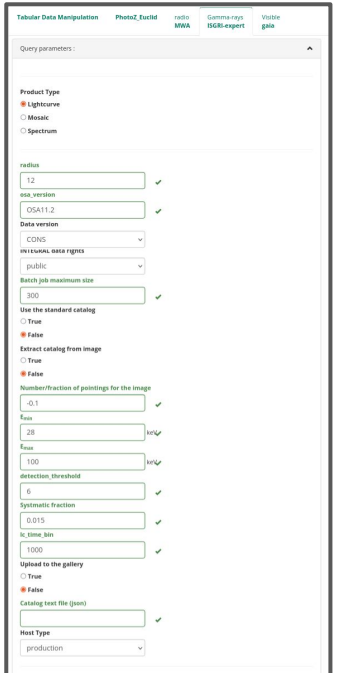


INTEGRAL ODA



MMODA
Multi-Messenger Online Data Analysis

High-level analysis

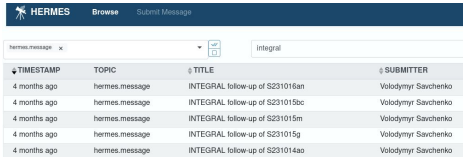


And because it's Web, no need to install and no need to copy the data. Can rent resources owned by someone else.
Now deployed at **ESA**!

Use Case of multi-messenger transients

Web brings diverse resources together **fast**. This matches the world of transients, since we have all the tooling in the same place. One the most used MMODA services is all-sky gamma-ray observations with INTEGRAL.

SciMMA, MOSSAIC, VOEvent, GCN, ATel, Kafka, ...



Conclusions, Sustainability, Dissemination

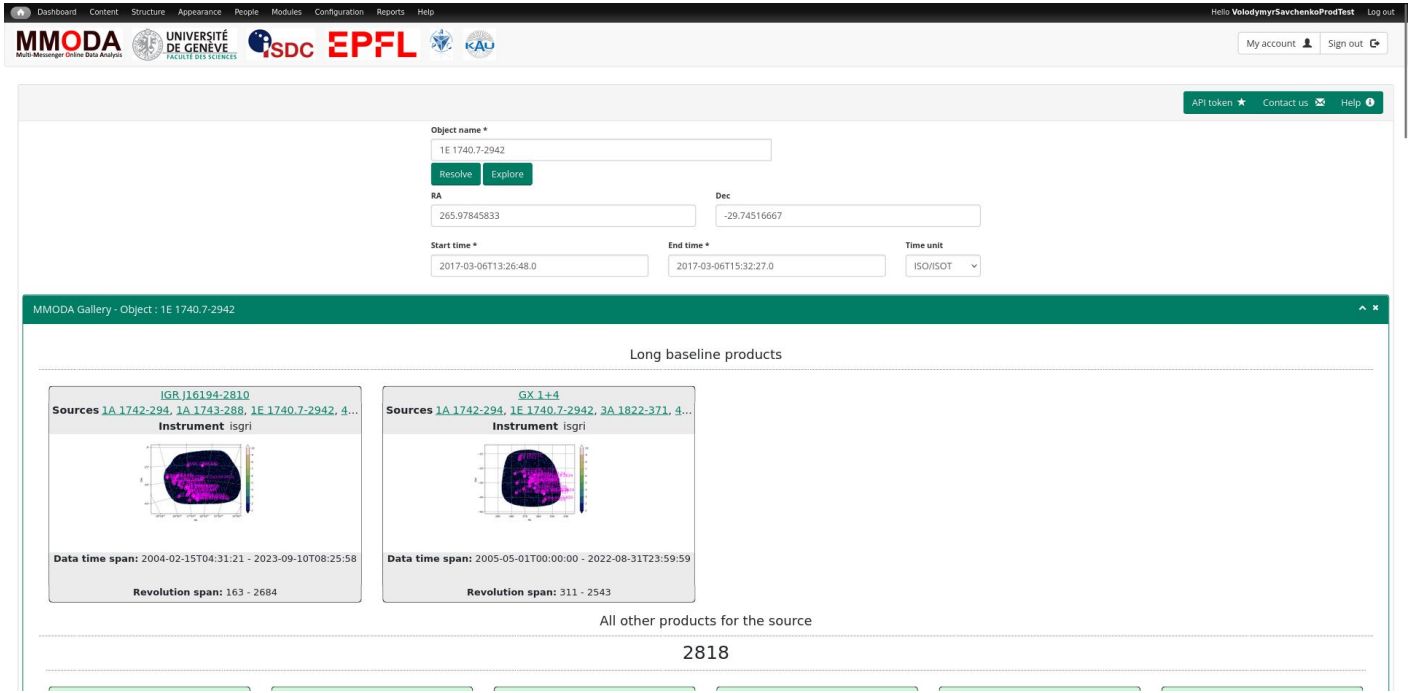
- AstroORDAS aims to develop **an ecosystem of cloud-based services and technologies** to provide **added value to data** from science data centers for astronomy, astroparticle and cosmology projects.

- **Sustainability** enabled by:
  - modular open-source design (see code)
  - community engagement by crowdsourcing workflows.
  - coordination with many similar Open Research projects and engagement in particular projects.

- Assets living past AstroORDAS:
  - analysis tools integrated in multiple **MMODA** and **Galaxy** instances
  - interactive visualisation libraries, JS libraries
  - technologies for connection to data and storage
  - publication technologies

- **EuroScienceGateway** continues connecting **Galaxy** platform to Astronomy. Galaxy has very large support (>100k unique users recently). Small new Galaxy EU project starting soon.
- **EU ACME** funding is just starting and will last for 5 years, including **MMODA** with CTAO, SKAO, KM3Net, ET etc.
- New ORD project **Solidipes** for enhancing reproducible publishing.
- MMODA/AstroORDAS already being used in a growing list of publications

Functions


Dashboard Content Structure Appearance People Modules Configuration Reports Help
Hello VolodymyrSavchenkoProdTest Log out
MMODA
Multi-Messenger Online Data Analysis
UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
ISDC
EPFL
My account
Sign out
API token
Contact us
Help
Object name *
1E 1740.7-2942
Resolve
Explore
RA
265.97845833
Dec
-29.74516667
Start time *
2017-03-06T13:26:48.0
End time *
2017-03-06T15:32:27.0
Time unit
ISO/ISOT
MMODA Gallery - Object : 1E 1740.7-2942
Long baseline products
IGR J16194-2810
Sources 1A 1742-294, 1A 1743-288, 1E 1740.7-2942, 4...
Instrument isgri
Data time span: 2004-02-15T04:31:21 - 2023-09-10T08:25:58
Revolution span: 163 - 2684
GX 1+4
Sources 1A 1742-294, 1E 1740.7-2942, 3A 1822-371, 4...
Instrument isgri
Data time span: 2005-05-01T00:00:00 - 2022-08-31T23:59:59
Revolution span: 311 - 2543
All other products for the source
2818