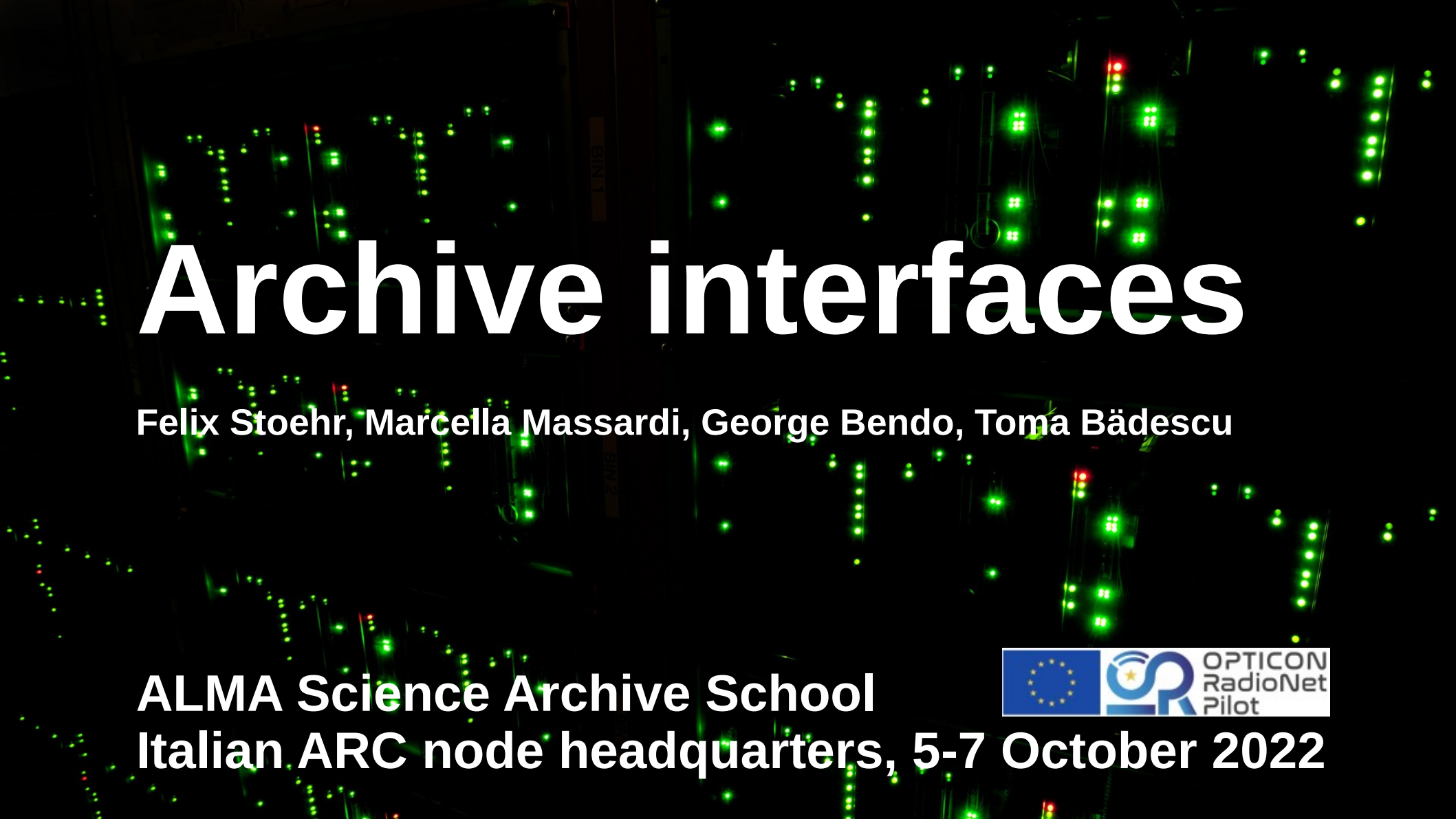


Archive interfaces

Felix Stoehr, Marcella Massardi, George Bendo, Toma Bädescu

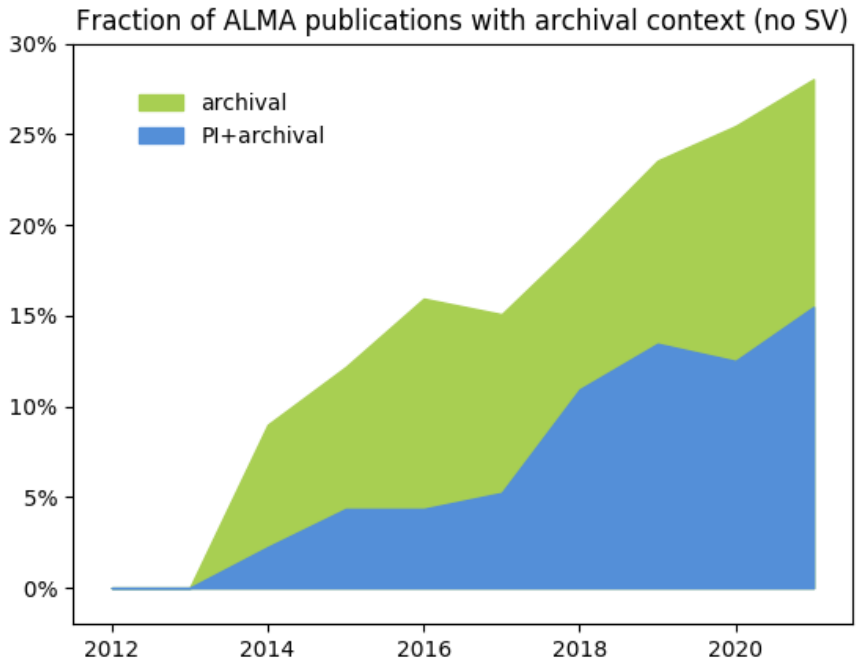
ALMA Science Archive School
Italian ARC node headquarters, 5-7 October 2022

OPTICON RadioNet Pilot

1. introduction

Context
Science Archives

- The Principle Investigators suggest observations and if granted time, get those observed.
- After typically 12 months they become public
- Science Archives are a major part of modern astronomical facilities
- Enables archival researches to work on huge amounts of data immediately



In a nutshell
ALMA Science Archive

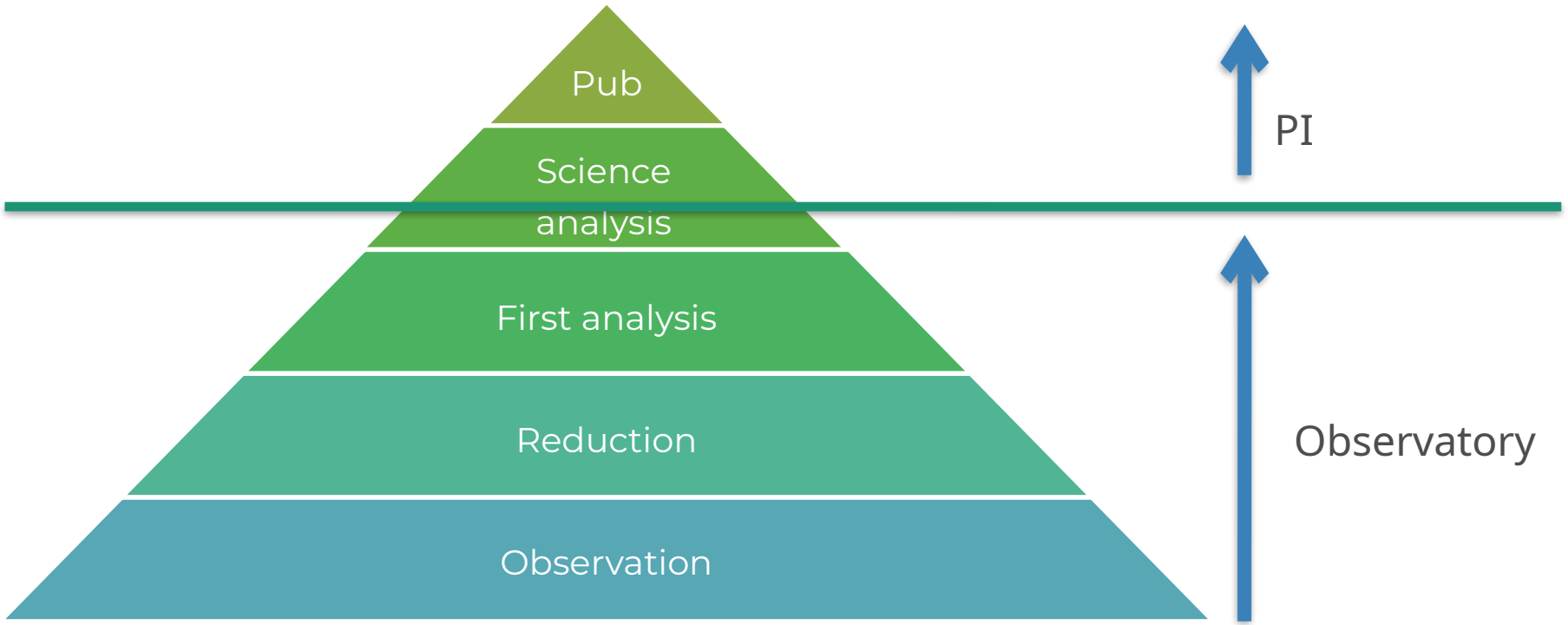
- 10 years of observations collected
- Science categories from the solar system to cosmology
- 1.4 PB of data
- 45 000 observations are already publicly accessible
- >10 000 of those have not yet been published

- The ASA is a **treasure-trove** for archival research!

Context
Best practices 2022

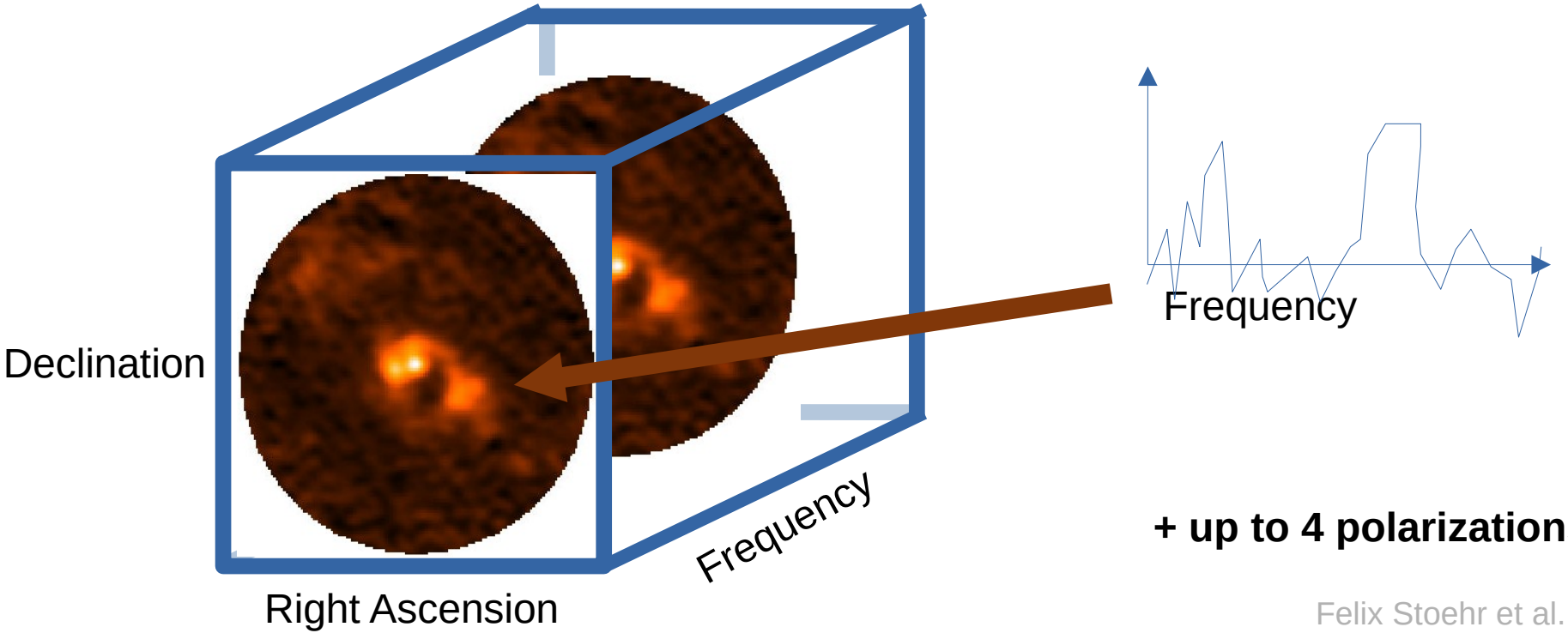
- Physical quantities
- Unscoped search
- Observations, Proposals, Publications
- Target-list upload
- Previews
- Modern user-experience
- Programmatic access (VO)
- Metadata are public
- Science-grade products + PL
- Anonymous downloads
- Self-describing FITS files
- Parallel downloads
- Authors must cite data-use
- Frequent Reprocessing
- **NEW**: Science platforms

Vision
Observatories: more responsibility



Vision
3D data challenge

- ALMA is producing 2D images and **3D data cubes**



+ up to 4 polarizations

Vision
3D data challenge

- ALMA is producing 2D images and **3D data cubes**
- More and more other facilities from x-ray to optical, too
- Huge challenges for visualization and analysis (3D source-extraction)

- ALMA Science Archive:
  - Spectral coverage viewer
  - Previews
  - CARTA visualization

Vision
fastronomy

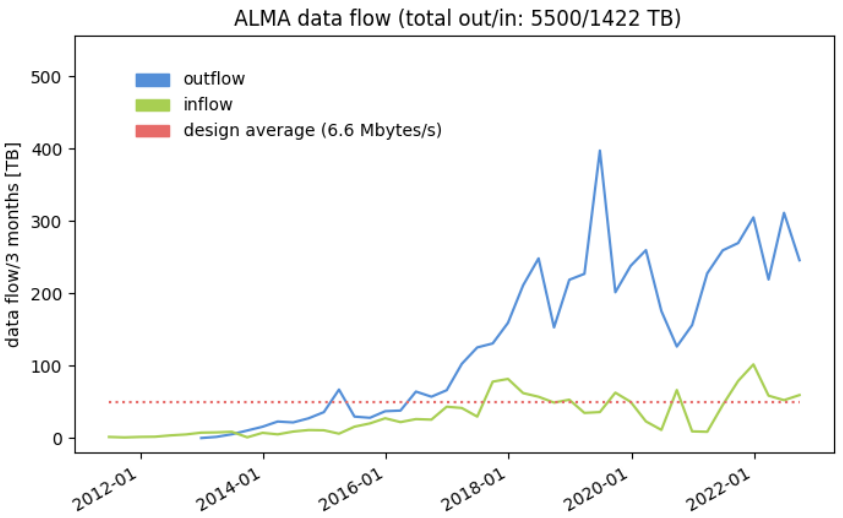
It is not enough that people can do what they need to do.
They need to be able to do it **fast**!

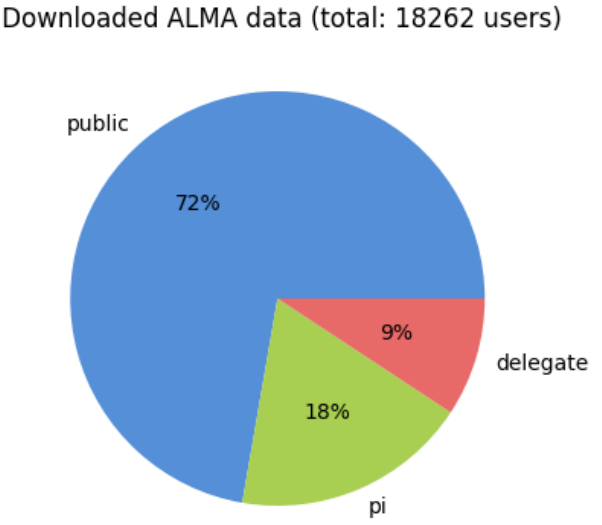
fastronomy

Let scientists concentrate on science

2. ALMA Science Archive

ALMA Science Archive Access





ALMA Science Archive
Towards ALMA Science Archive 1.0

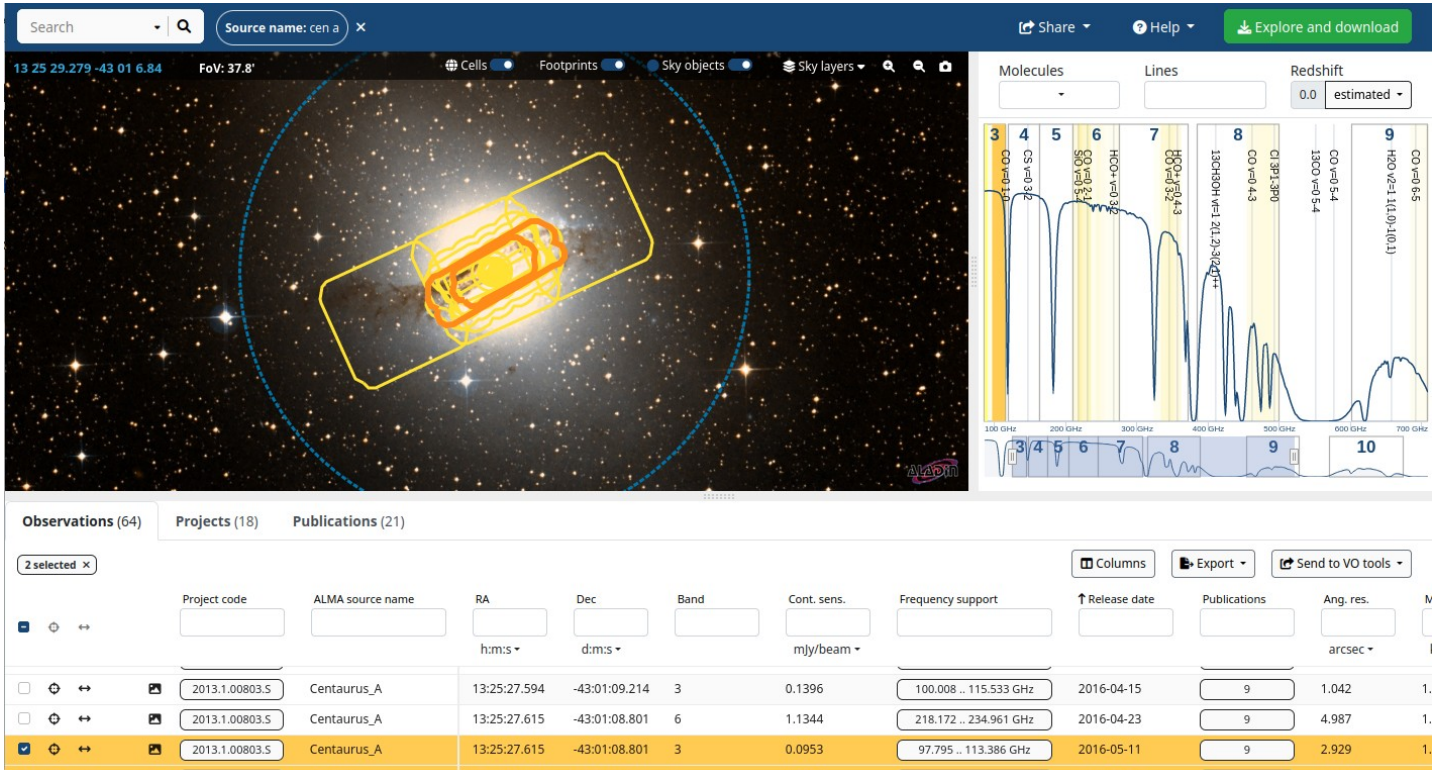
Recent developments

- CARTA
- Text-based similarity search
- Complete VO suite
- Jupyter Notebook tutorials
- Object-type search
- Previews

User survey 2019
statistics and 43 pages of comments

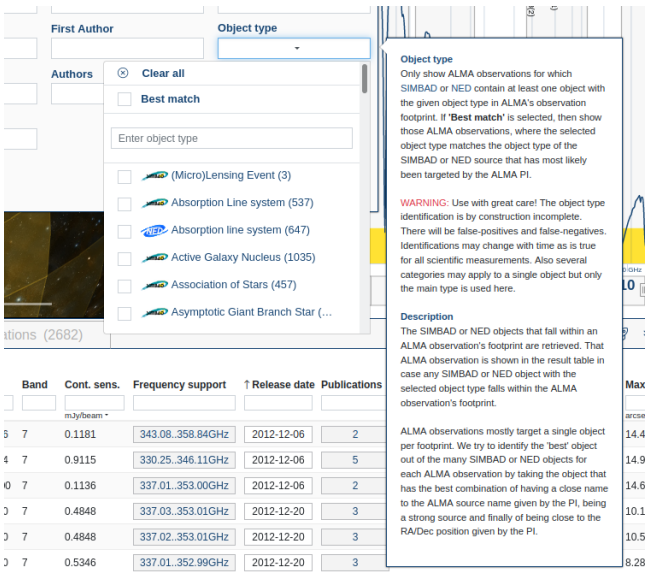
ALMA Science Archive

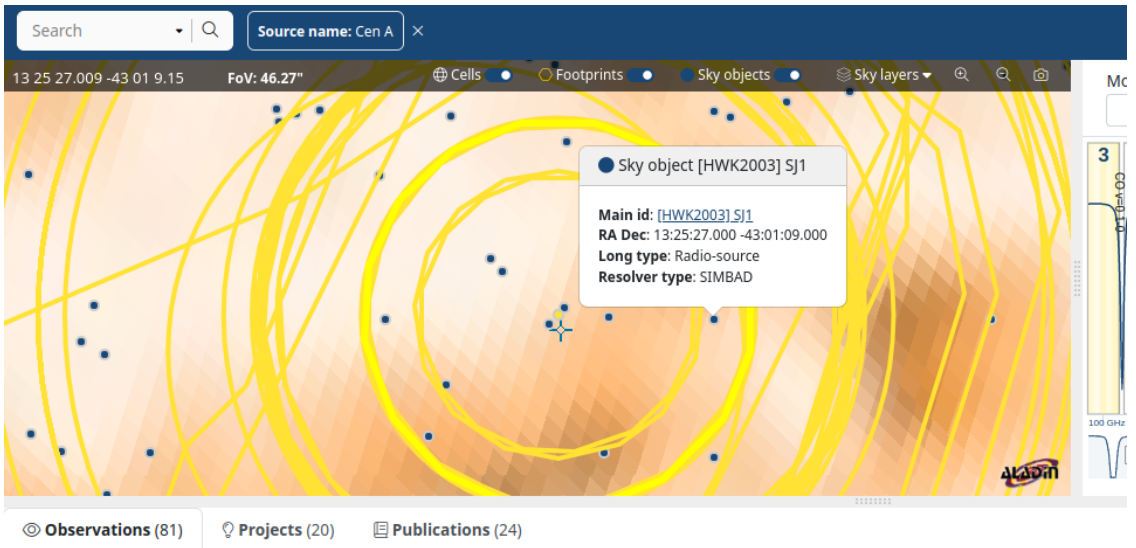
almascience.org/aq (live demo)



ALMA Science Archive
Object-type search

All SIMBAD and NED objects that fall into ALMA observations can be **queried** for on the ASA and are **displayed** on the interface.





ALMA Science Archive
VO tools: Aladin and Topcat

- Start Aladin, Topcat or both
  - java -jar topcat-full.jar (download link)
  - java -jar Aladin.jar (download link)
- On the ALMA Science Archive Interface click on the rocket symbol and choose between all displayed rows and the selected rows
- In order to see the observation footprints in Aladin, click on the “ALMA Science Archive” entry on the right and then scroll until the FoV appears. Click on those.

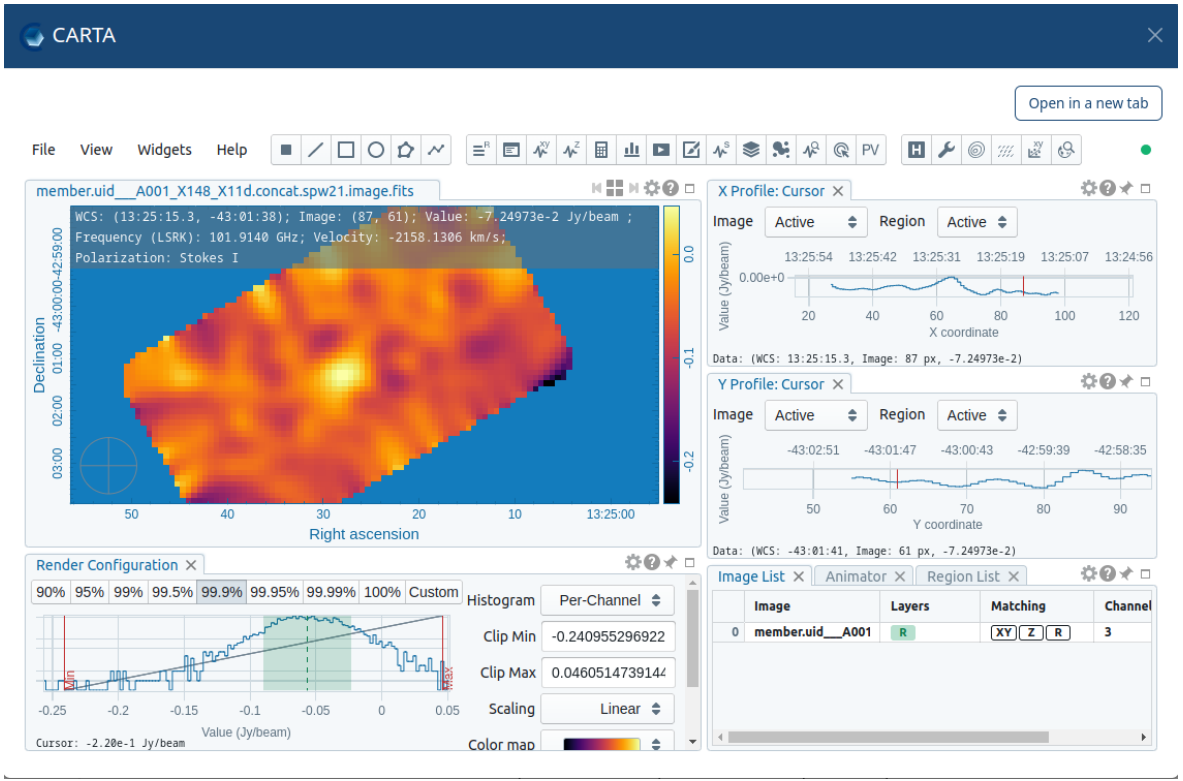
ALMA Science Archive
Complete VO suite

We now have a **full suite of VO tools** in place including

- TAP (ObsCore)
- SIAv2
- DataLink (now used in TAP and SIAv2 outputs)
- SODA for cut-outs

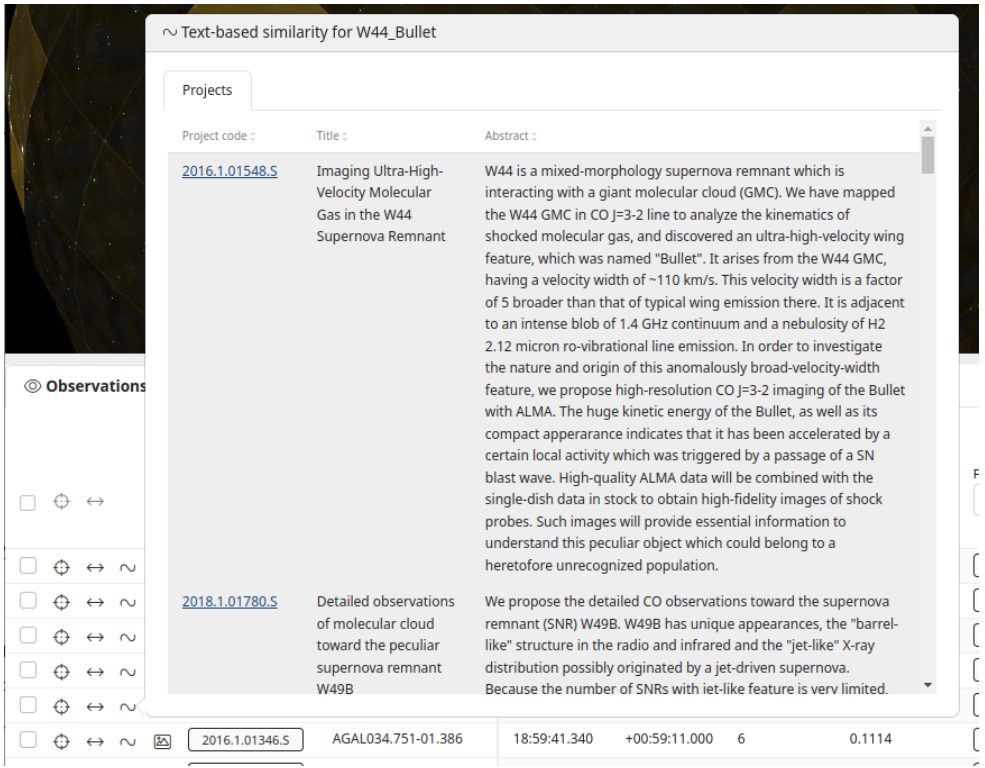
ALMA Science Archive
CARTA

CARTA 3.0 is available on the ASA on the query page as well as on the download page.



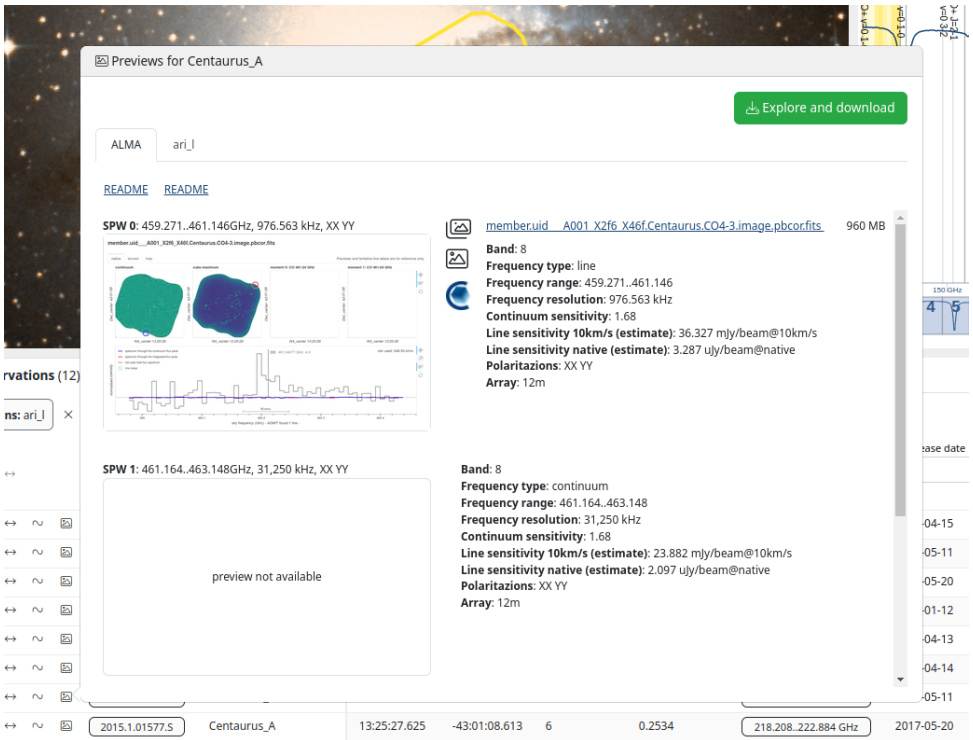
ALMA Science Archive
Text-based similarity

The ASA can show projects or publications that are similar to a given one based on **state-of-the-art text similarity**. (Idea and proof-of-concept: Alejandro Barrientos)

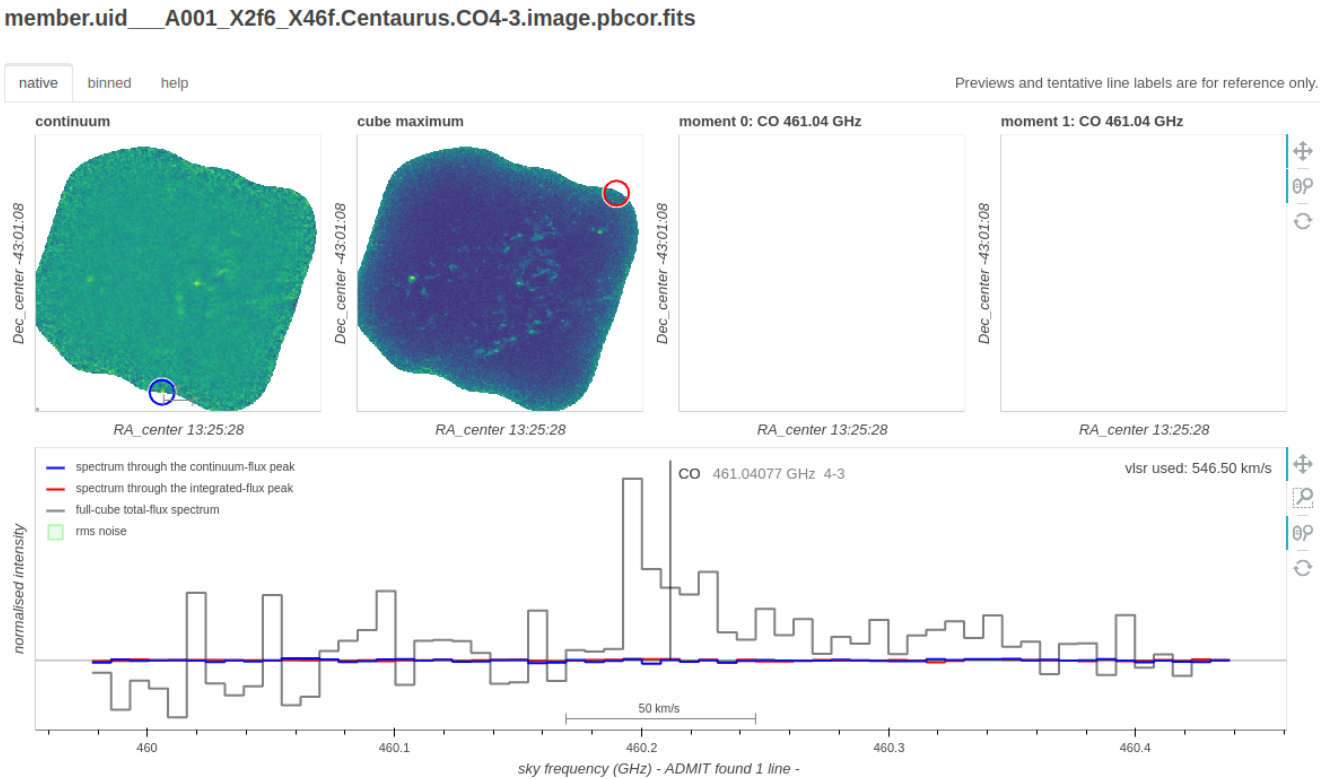


ALMA Science Archive
Previews

Previews for (nearly all) ALMA FITS files are available directly from the query form. The previews are static files but nevertheless fully **interactive** . We use the ALMA Data Mining Tool-Kit (ADMIT) to run line-finding and tentative **line-identification** on the ALMA images.



ALMA Science Archive
Previews



ALMA Science Archive
Some other noteworthy features

- Spectral coverage viewer (see 3D challenge), including lines that might be covered
- Search for real physical quantities (e.g. sensitivity, resolutions)
- Search the entire n-dimensional data cube. I.e. observations, projects and publications together “show all public but not yet published observations”)
- Slide smoothly through sky backgrounds
- Get a calendar event when proprietary data will become public
- …
- Data downloads will be discussed in George’s presentation

4. Documentation

Documentation
ASA manuals

- Archive Manual

  https://almascience.eso.org/alma-data/documents-and-tools/latest/science-archive-manual
- Video tutorials

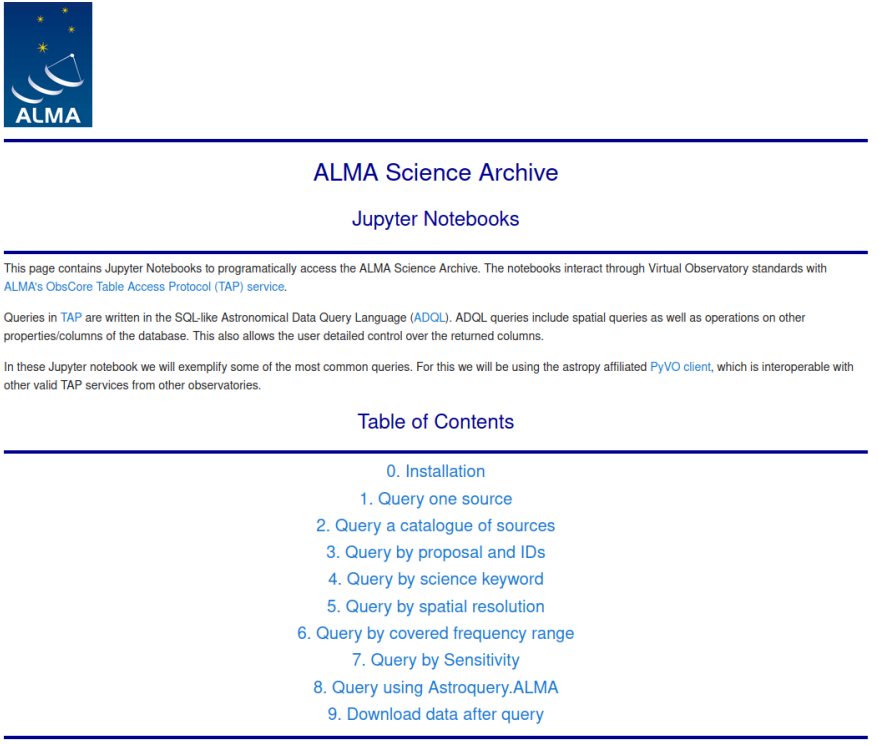
  https://almascience.eso.org/alma-data/archive/archive-video-tutorials
- ALMA Archival data – a Primer

  https://almascience.eso.org/documents-and-tools/cycle9/archive-primer
- Jupyter notebooks

  https://almascience.eso.org/alma-data/archive/archive-notebooks

Documentation
Jupyter notebook tutorials

Tutorials for programmatic access to the ASA have been published on the Science Portal as Jupyter Notebooks.



ALMA Science Archive

Jupyter Notebooks

This page contains Jupyter Notebooks to programatically access the ALMA Science Archive. The notebooks interact through Virtual Observatory standards with ALMA's ObsCore Table Access Protocol (TAP) service.

Queries in TAP are written in the SQL-like Astronomical Data Query Language (ADQL). ADQL queries include spatial queries as well as operations on other properties/columns of the database. This also allows the user detailed control over the returned columns.

In these Jupyter notebook we will exemplify some of the most common queries. For this we will be using the astropy affiliated PyVO client, which is interoperable with other valid TAP services from other observatories.

Table of Contents

4. Additional tools

Additional tools
ALMiner

https://github.com/emerge-erc/ALminer

alminer is a Python-based code to effectively query, analyse, and visualize the ALMA science archive. It also allows users to directly download ALMA data products and/or raw data for further image processing.

Additional tools
ALMA Data Mining Tookit (ADMIT)

https://admit.astro.umd.edu

git clone --branch python3 https://github.com/astroumd/admit.git

A toolkit that works on ALMA data using CASA to

- detect sources
- find lines
- identify lines

The ALMA previews make use of ADMIT for the line finding and line identification.