Data mining:
Making the most of the ALMA archive

Benjamin Magnelli
German ALMA Regional Centre

The ALMA archive in a nutshell

What can I do with the ALMA archive ?

- See the list of all observed ALMA projects, i.e., previous and current cycles
- Download publicly available datasets (i.e., delivered more than a year ago)

Why should I use the ALMA archive ?

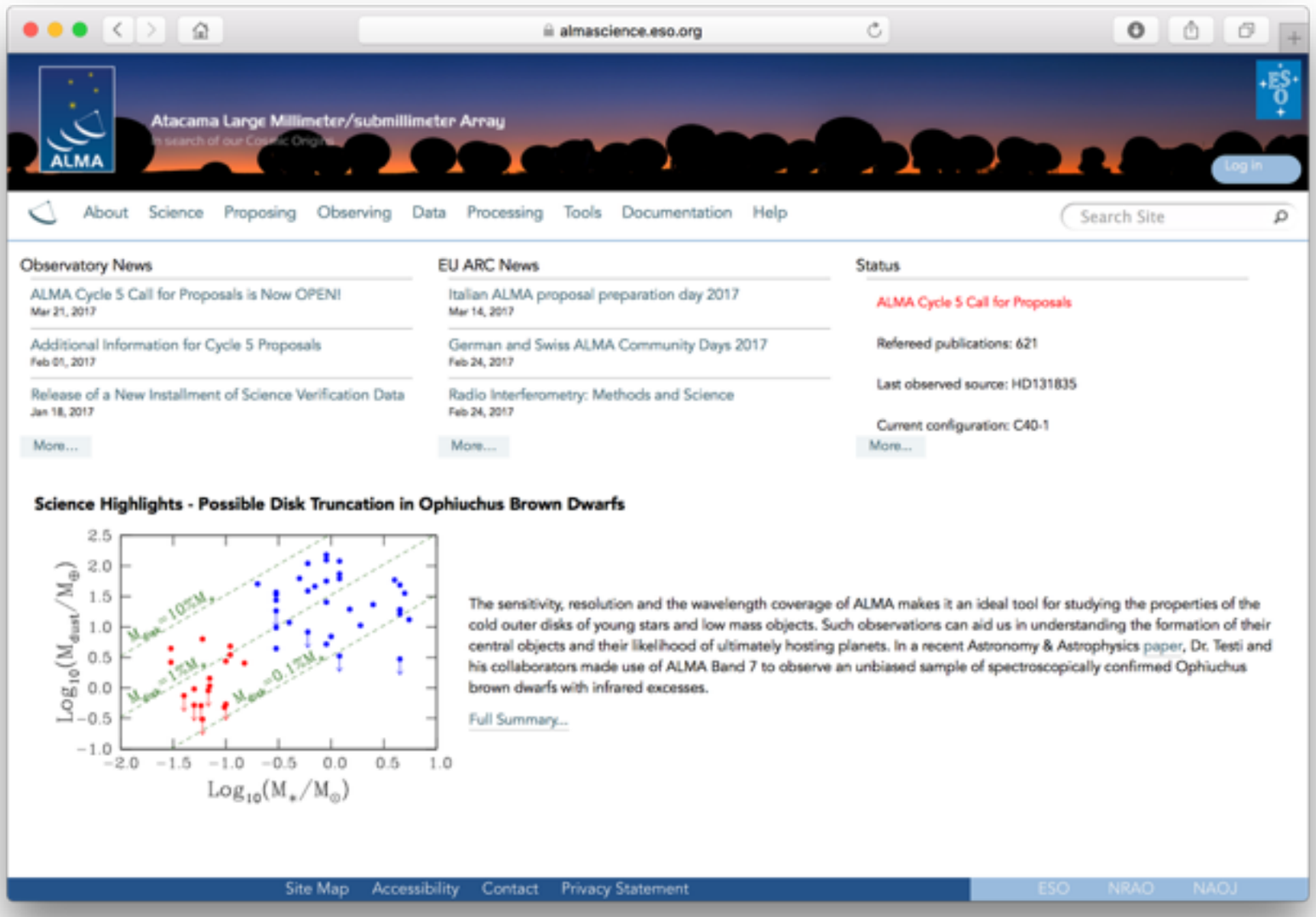
- To verify possible duplication issues before submitting your proposal
- To retrieve your proprietary ALMA observations
- To make the most of the ALMA observations not yet exploited by their PI

How can I access the ALMA archive ?

- via its web interface on the ALMA science portal (**WWW.ALMASCIENCE.ESO.ORG**)
- via the python package **ASTROQUERY**

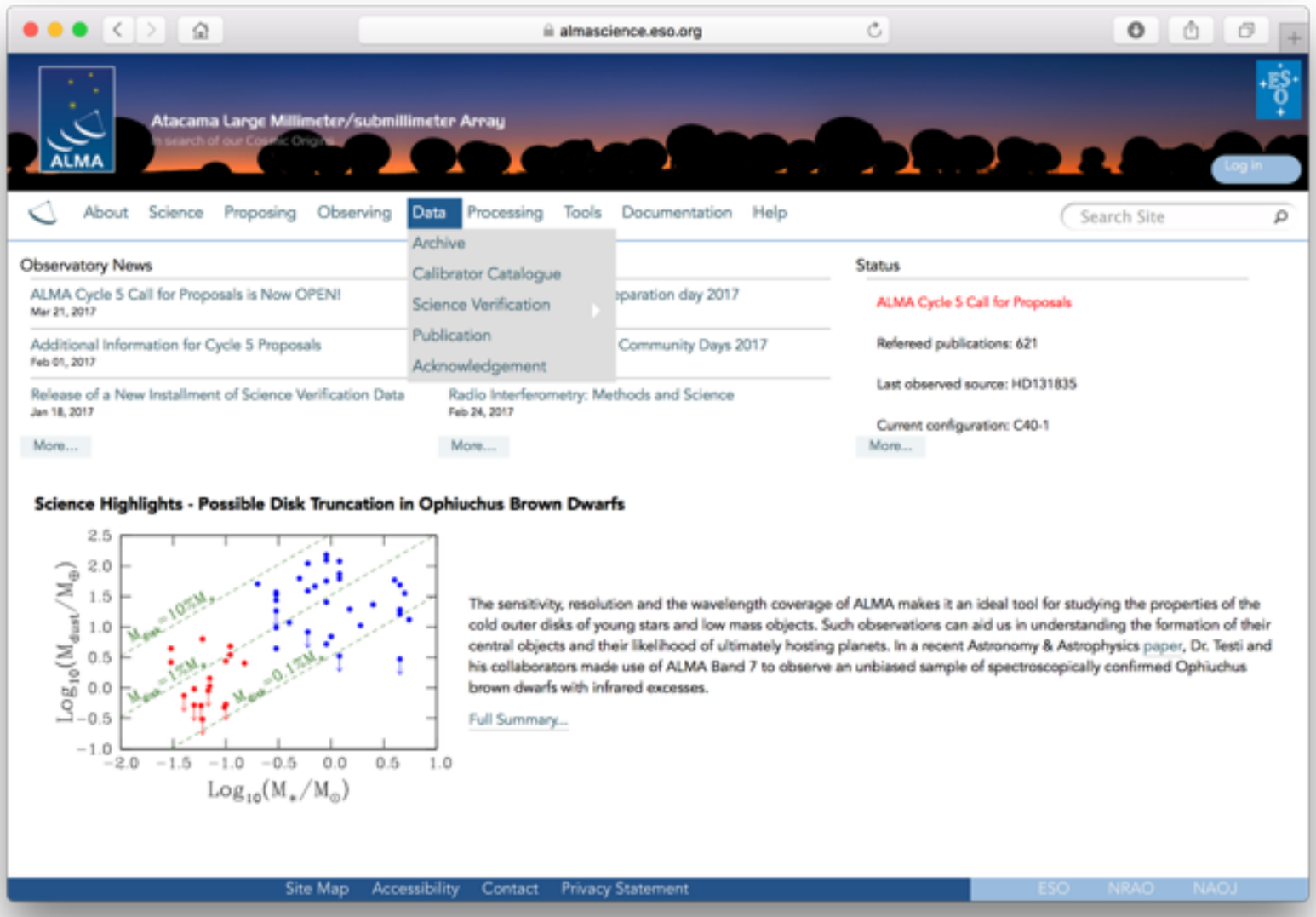
Exploring the ALMA archive

Everything start, of course, from the ALMA science portal !



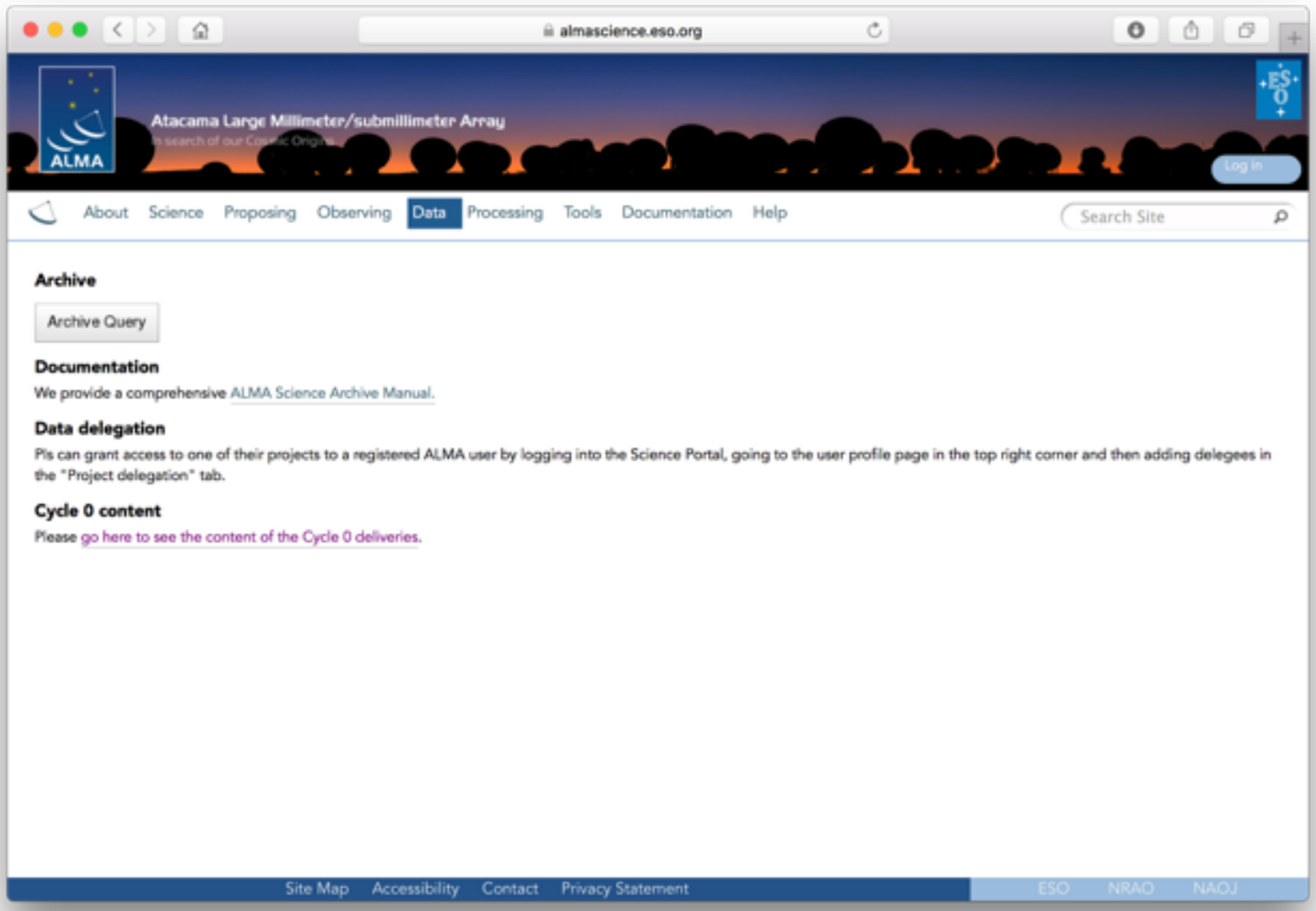
Exploring the ALMA archive

Everything start, of course, from the ALMA science portal !



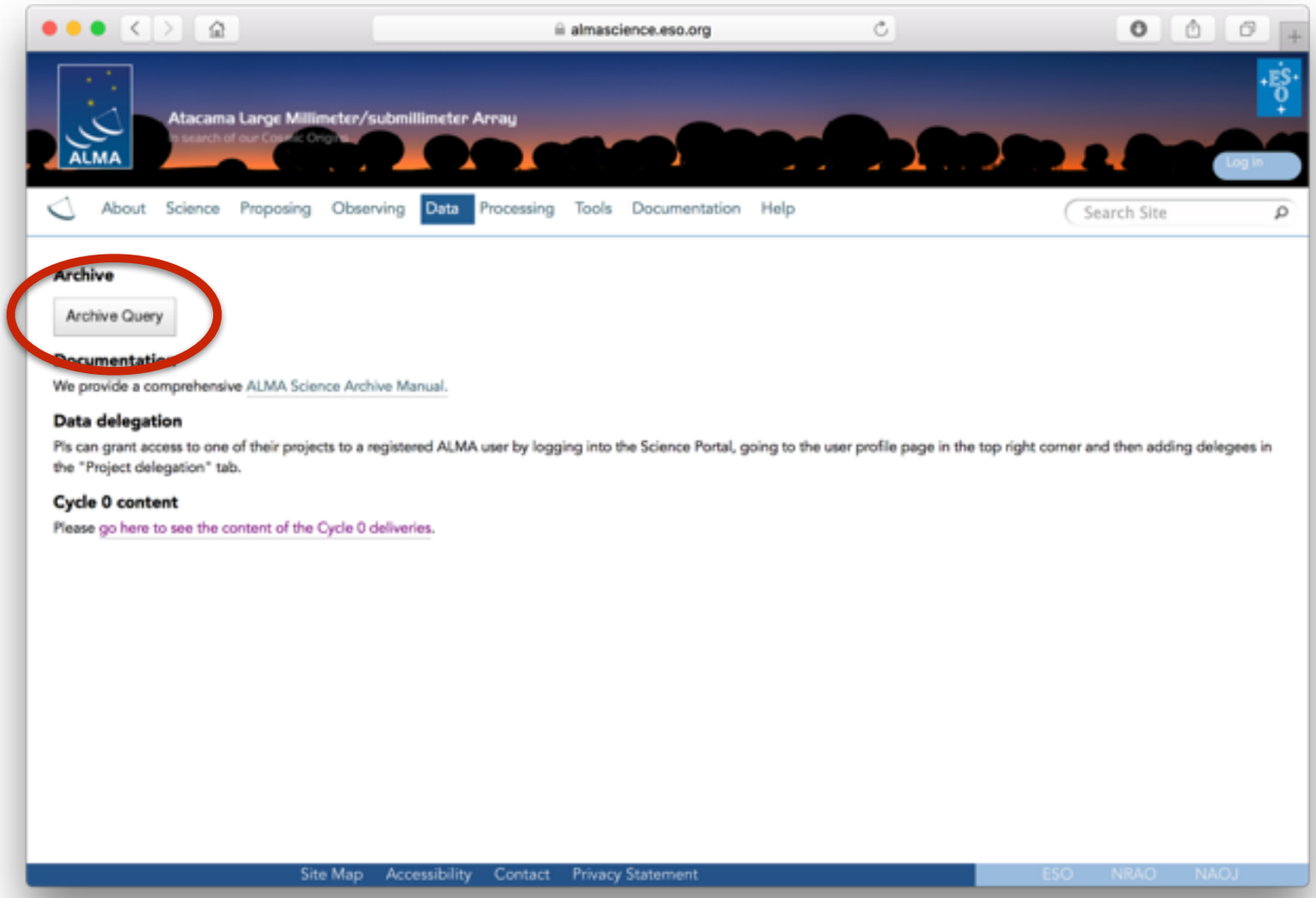
Exploring the ALMA archive

Everything start, of course, from the ALMA science portal !



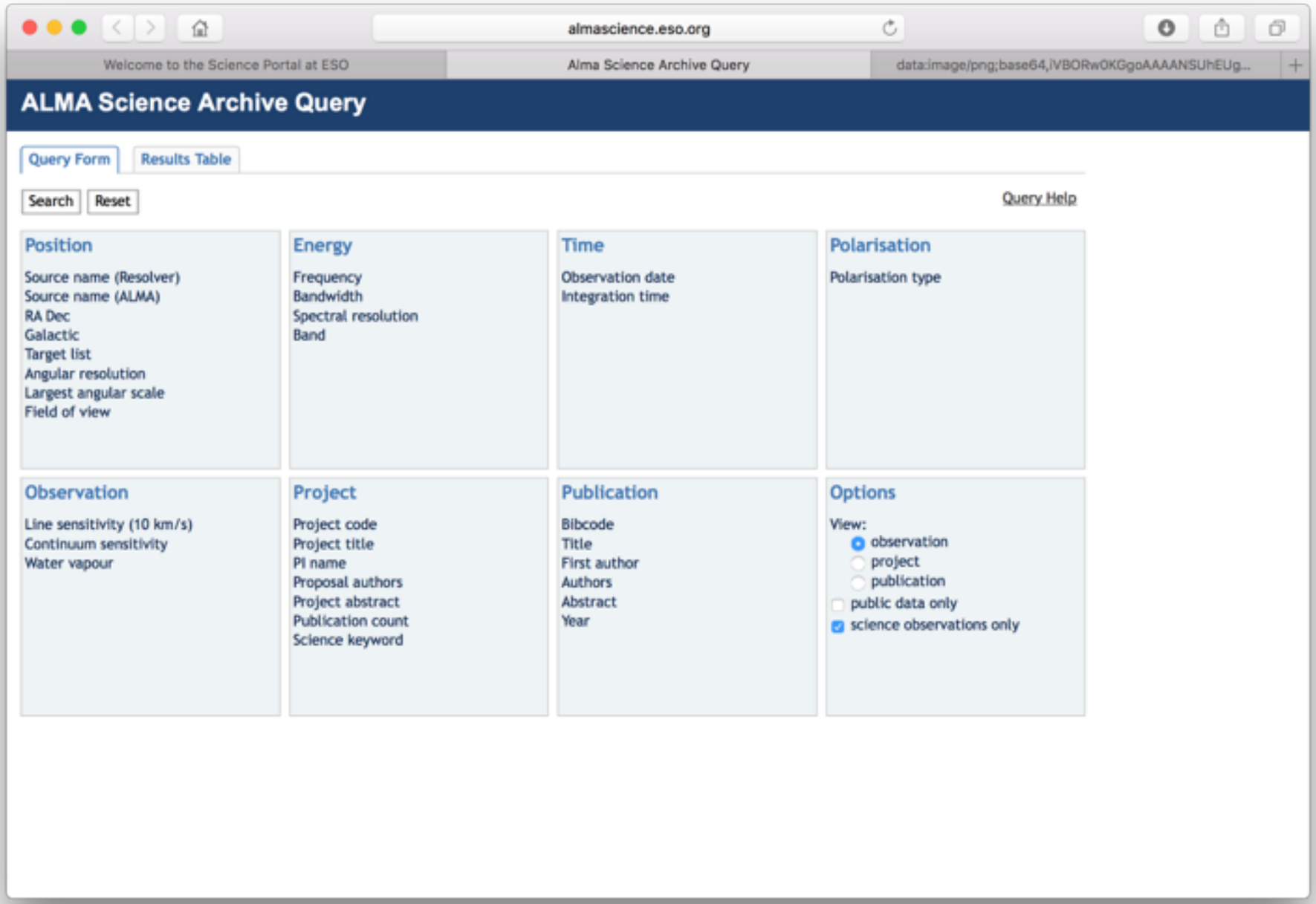
Exploring the ALMA archive

Everything start, of course, from the ALMA science portal !



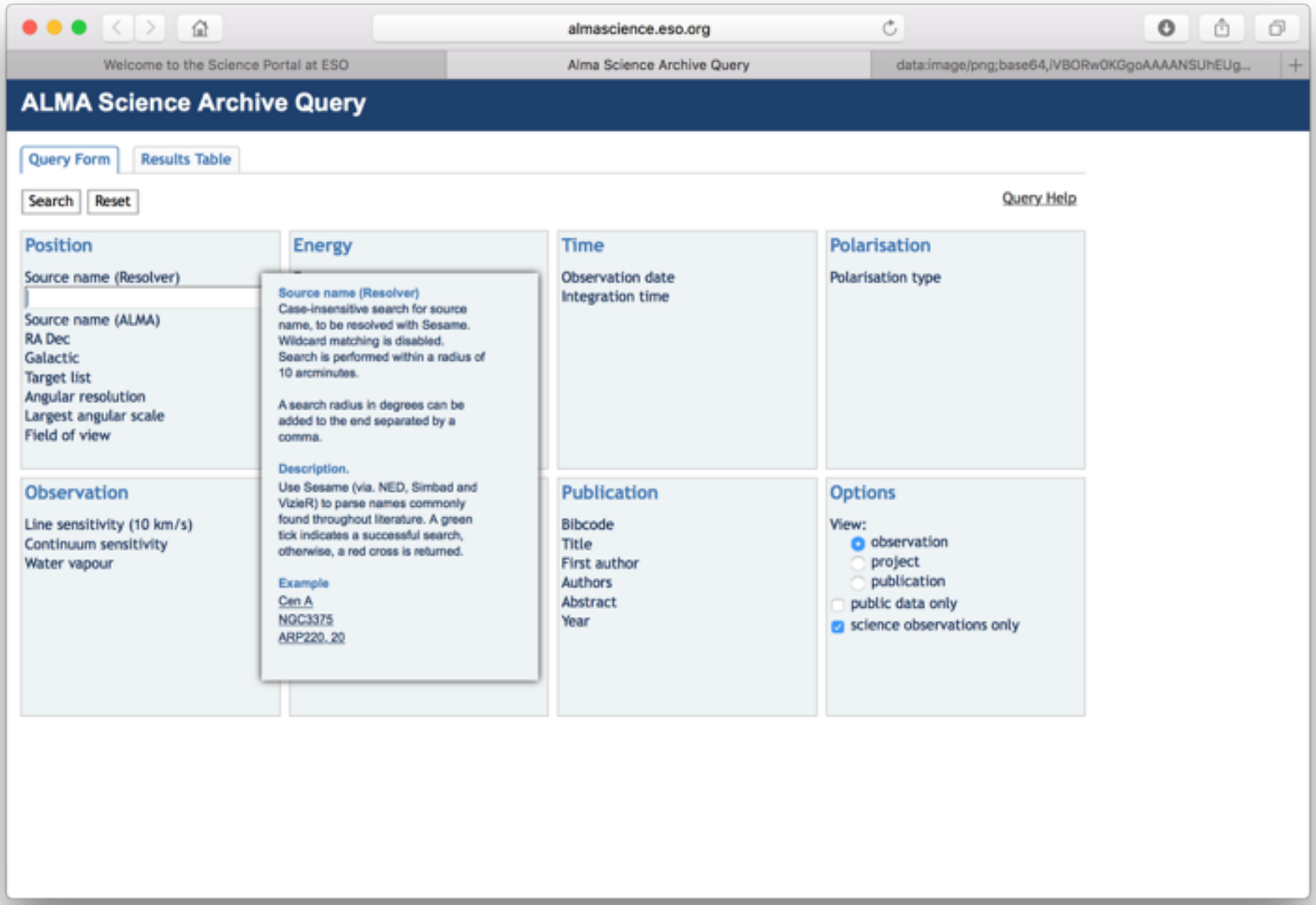
Exploring the ALMA archive

Queries can be performed in many different ways and can be combined.



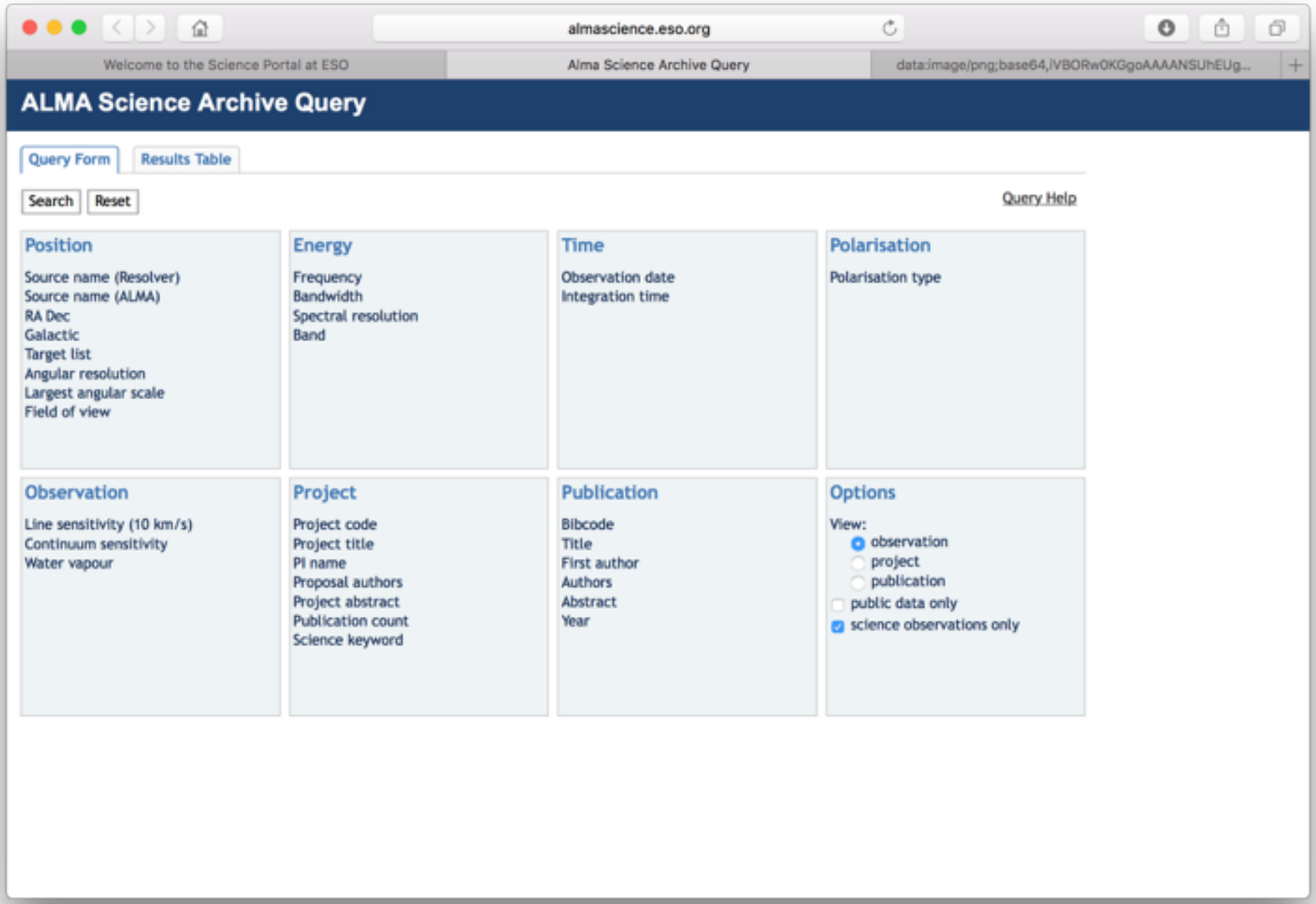
Exploring the ALMA archive

To start just hover over the appropriate search box
(Tips and examples will also appear)



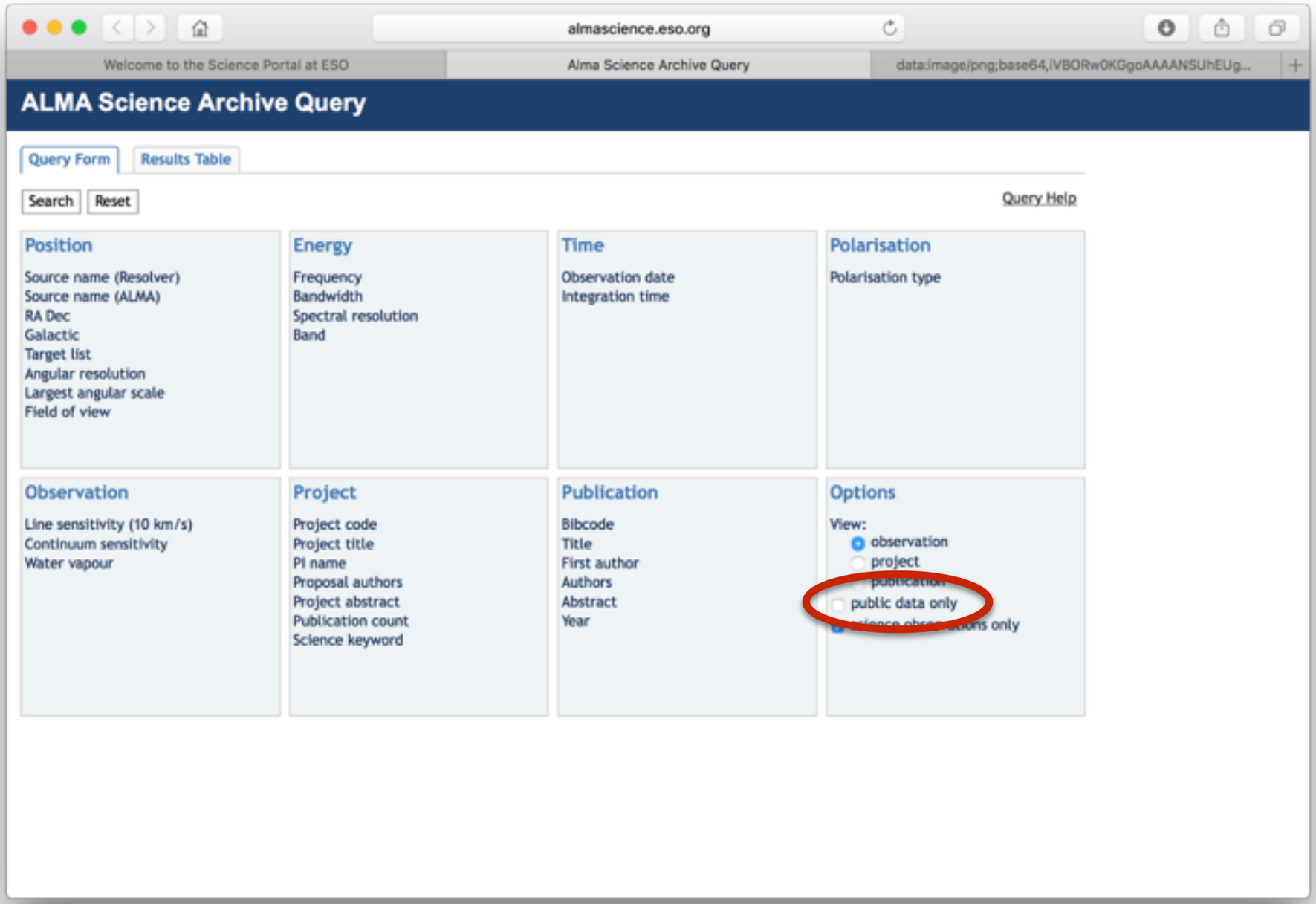
Exploring the ALMA archive

To start just hover over the appropriate search box
(Tips and examples will also appear)



Exploring the ALMA archive

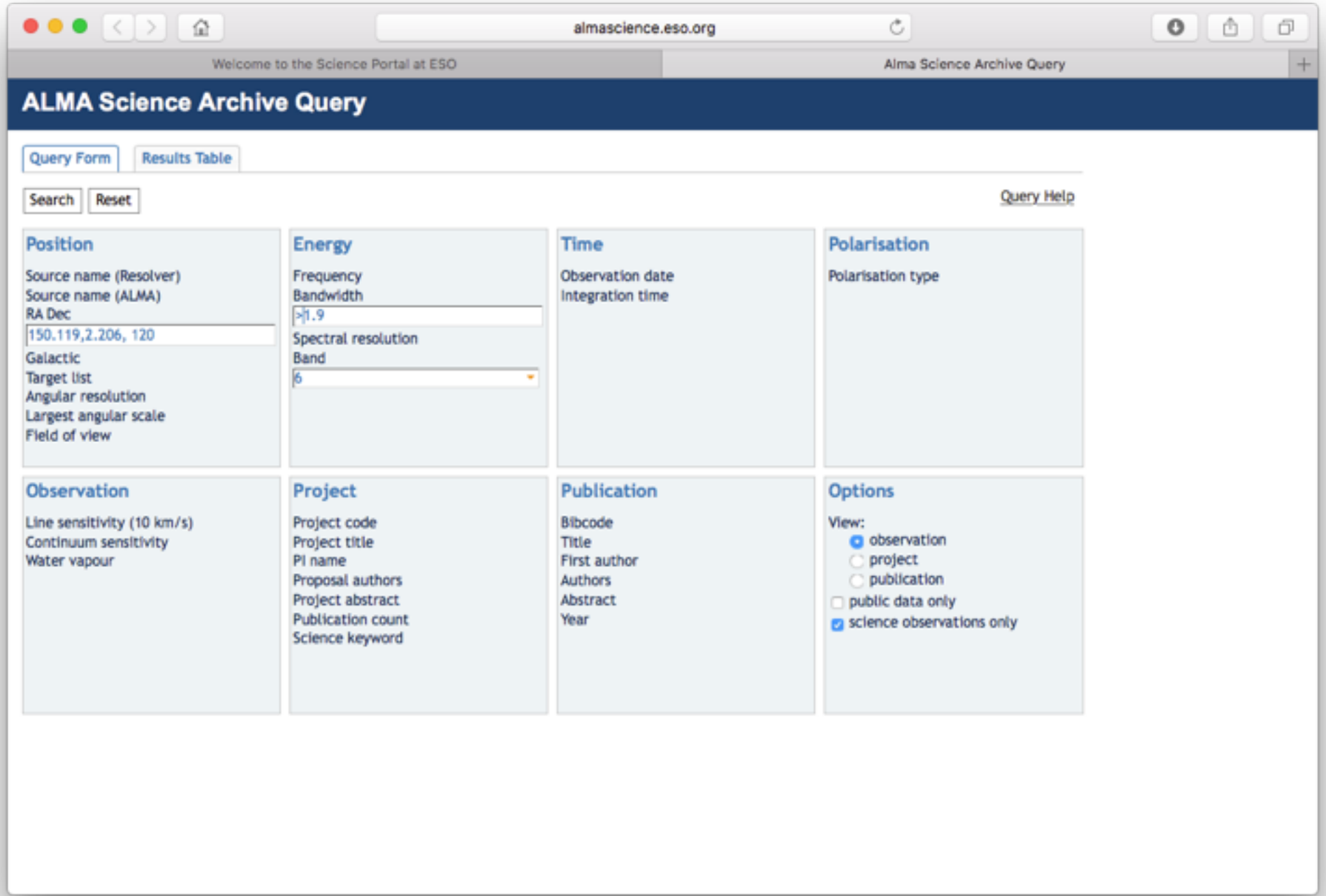
To start just hover over the appropriate search box
(Tips and examples will also appear)



Exploring the ALMA archive

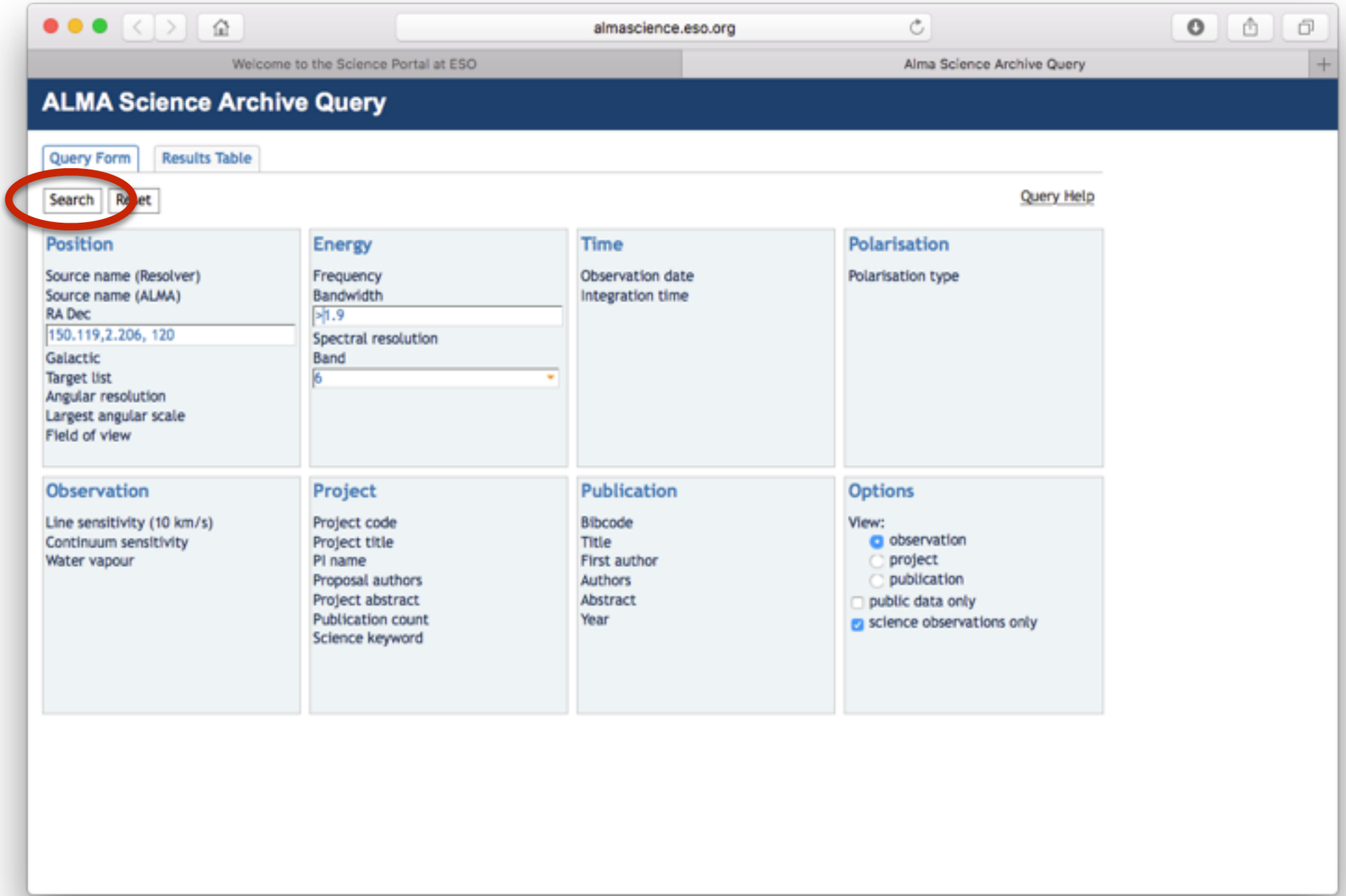
To start just hover over the appropriate search box

(Tips and examples will also appear)

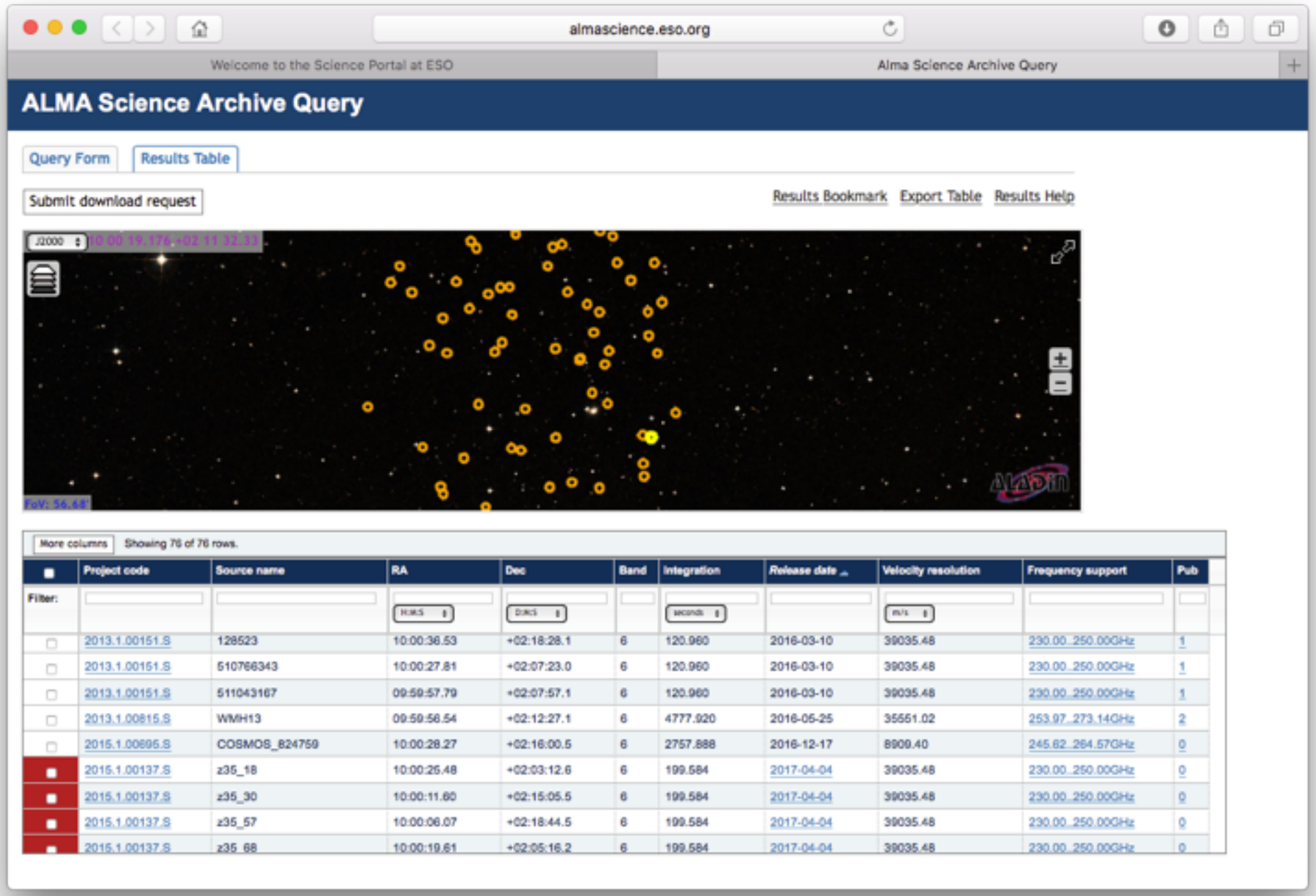


Exploring the ALMA archive

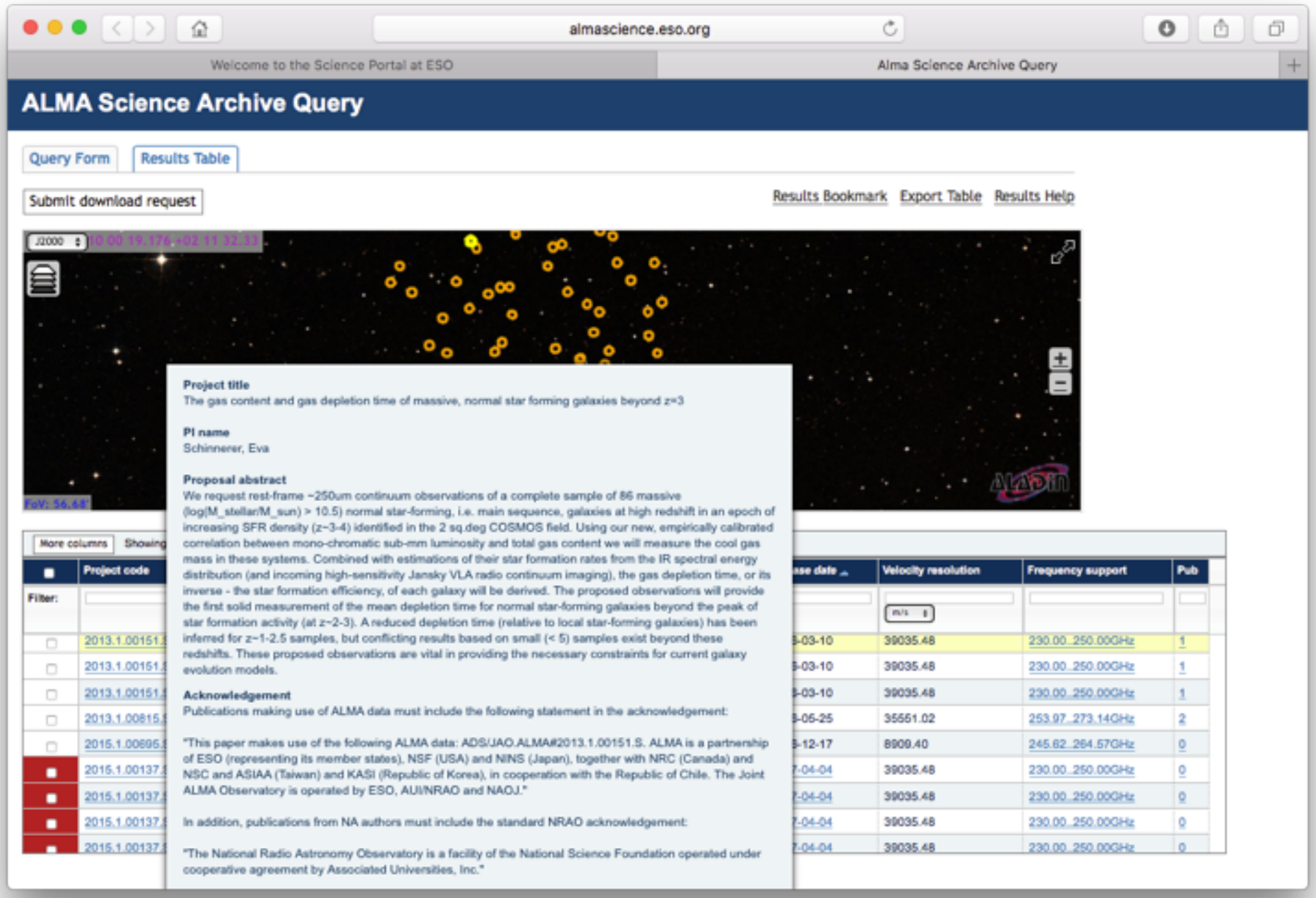
To start just hover over the appropriate search box
(Tips and examples will also appear)



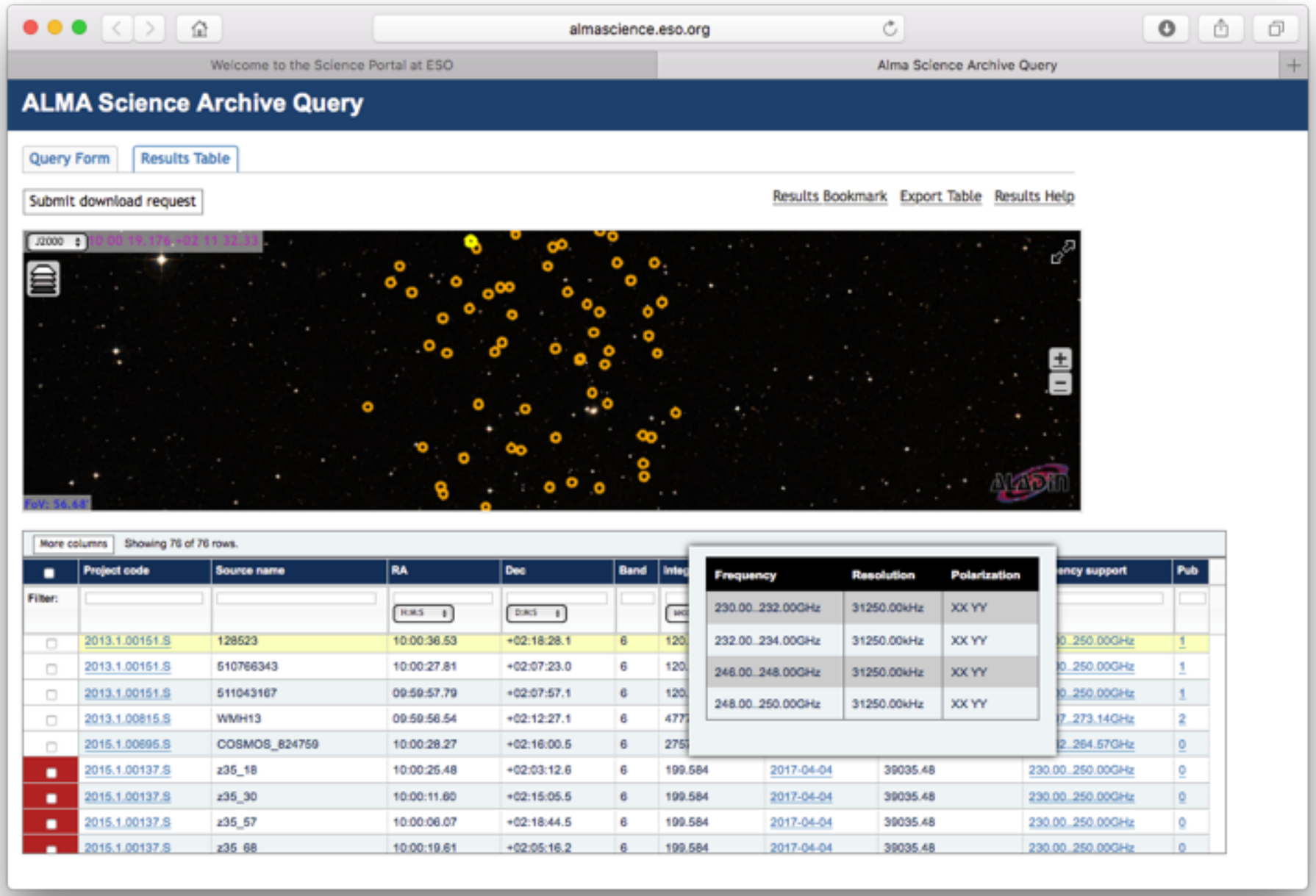
Exploring the ALMA archive



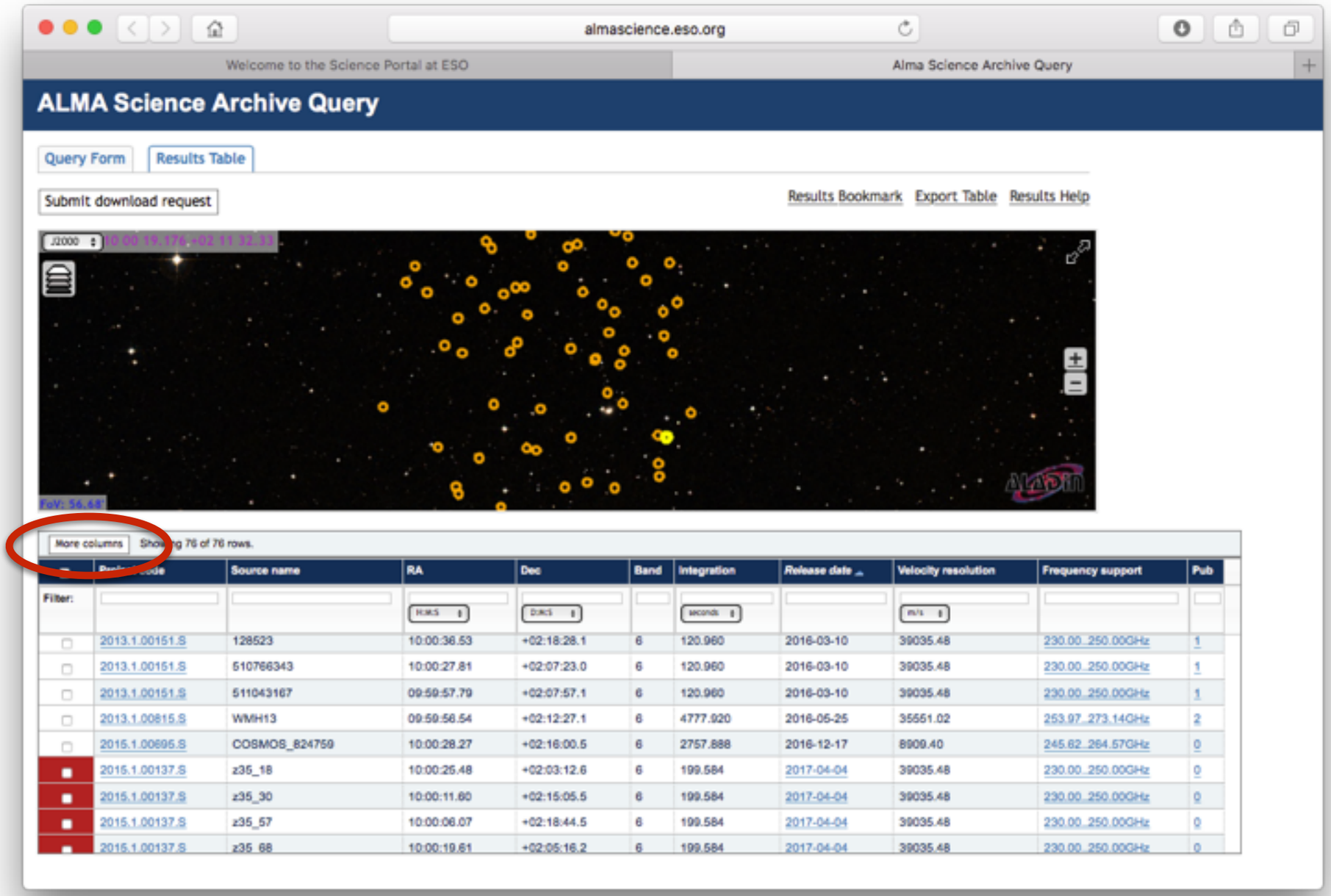
Exploring the ALMA archive



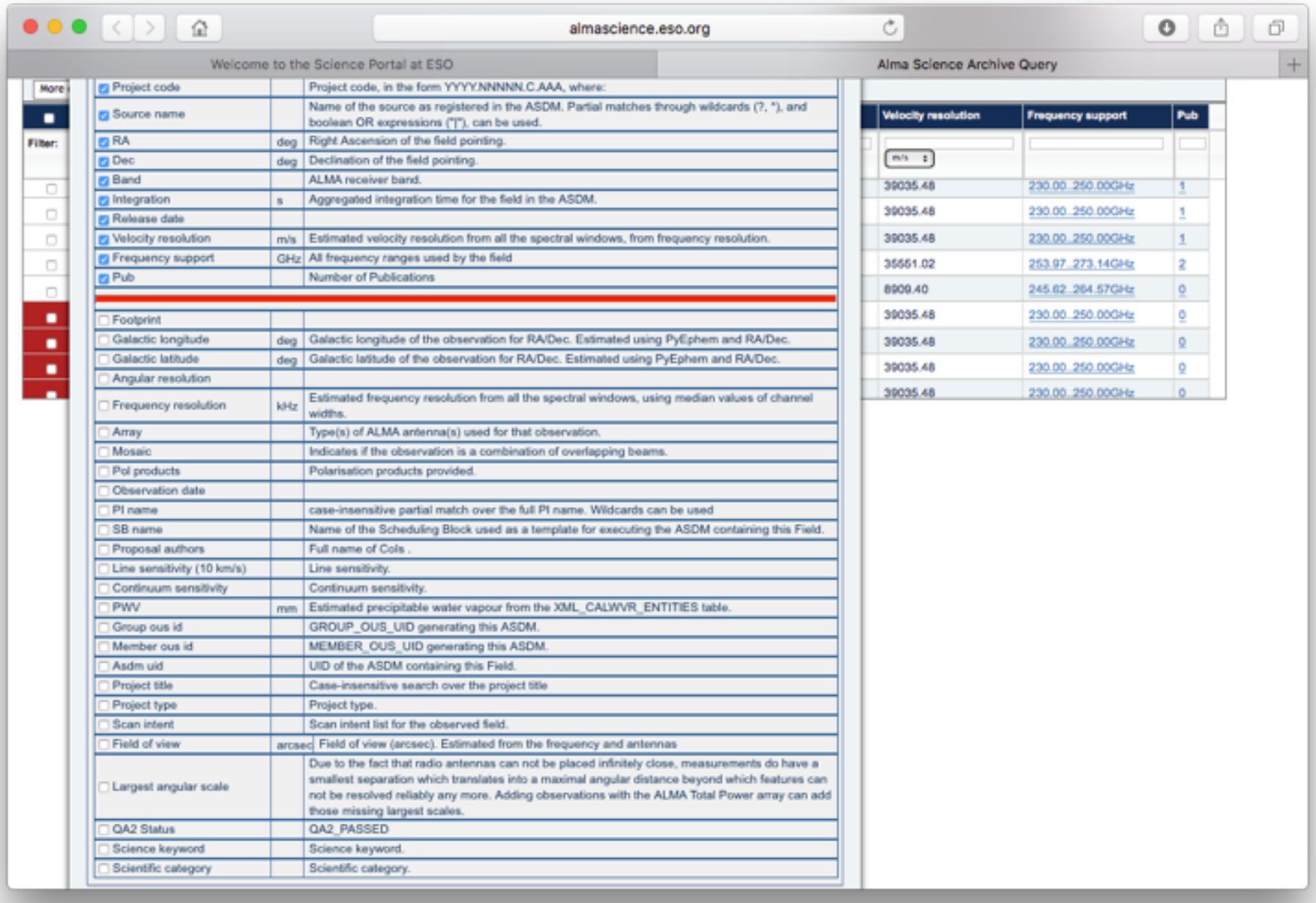
Exploring the ALMA archive



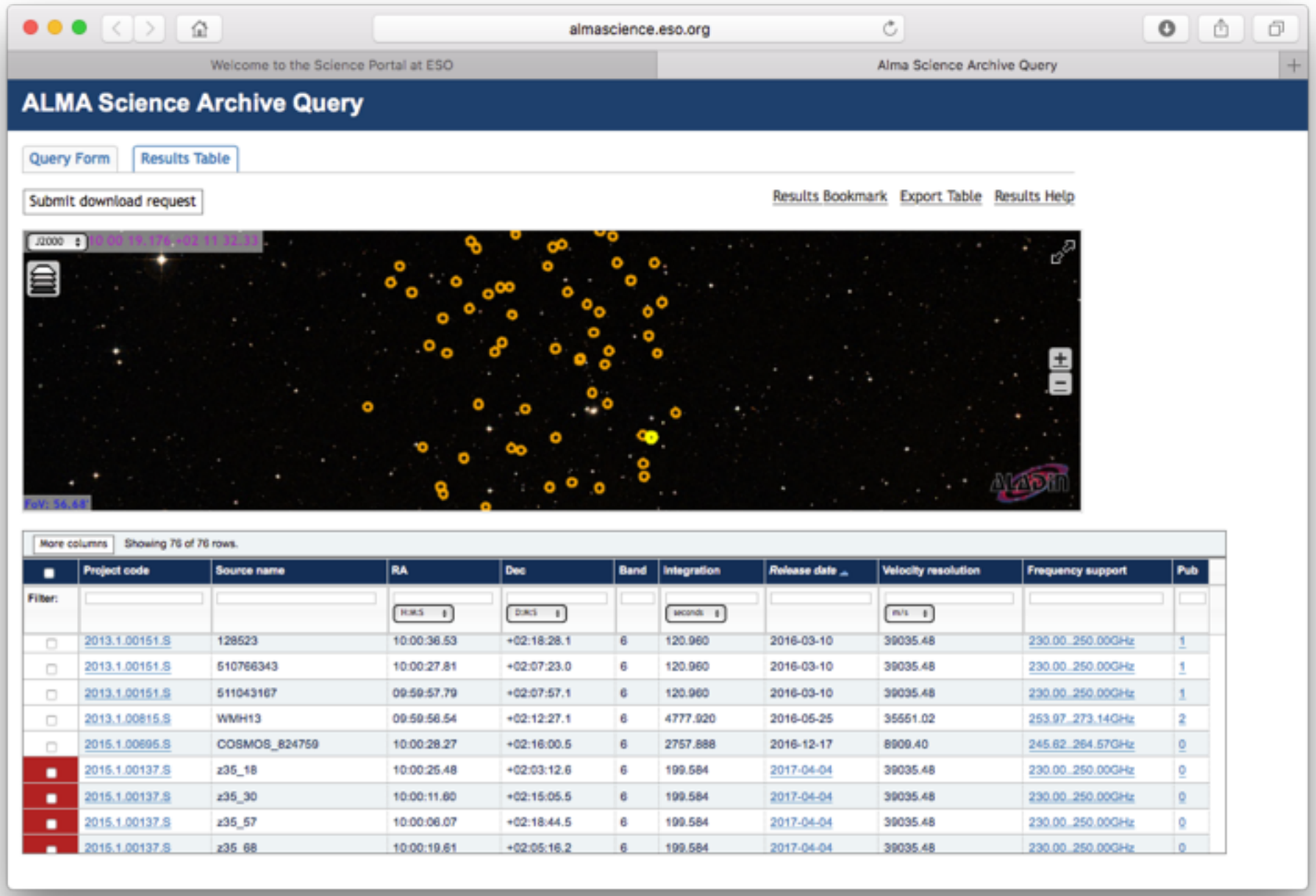
Exploring the ALMA archive



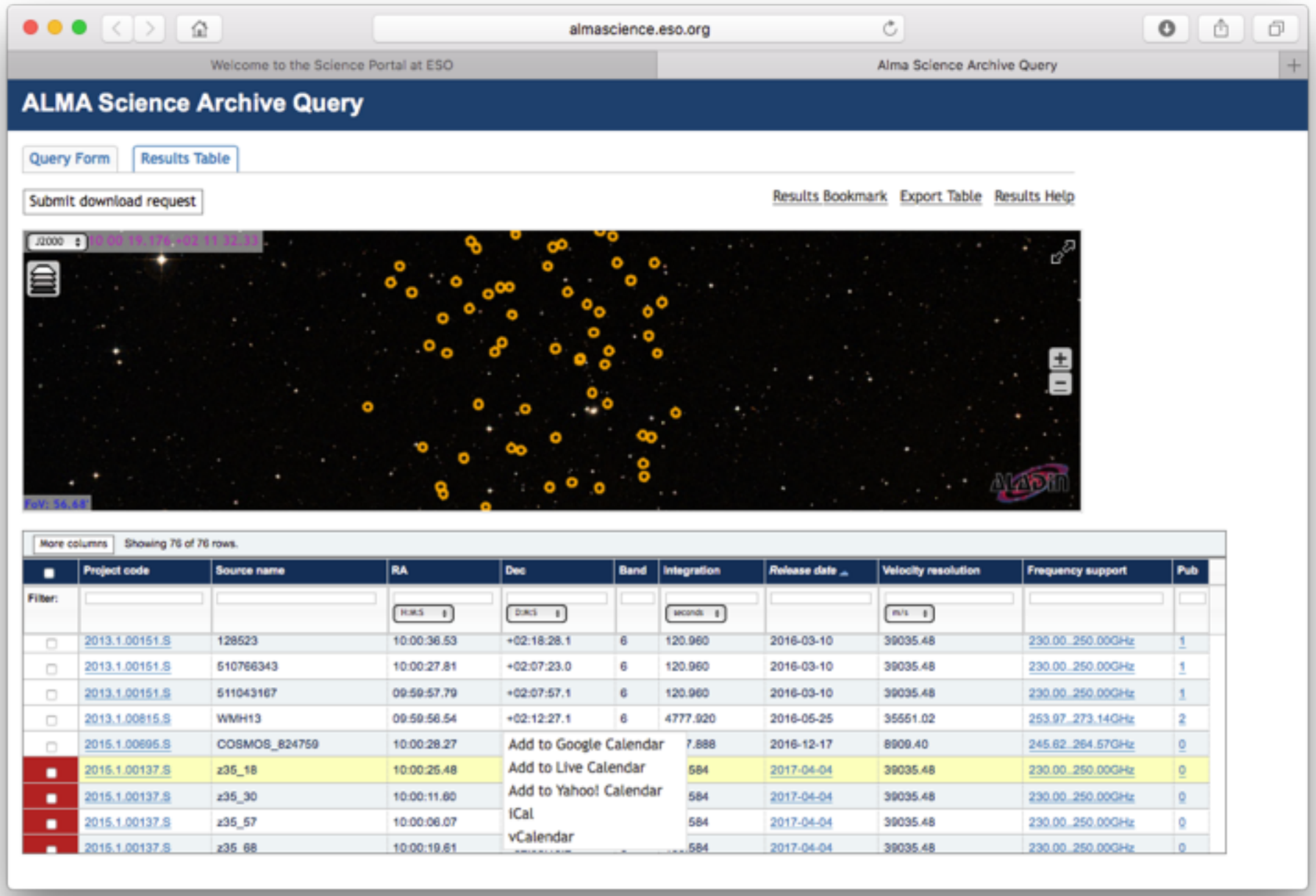
Exploring the ALMA archive



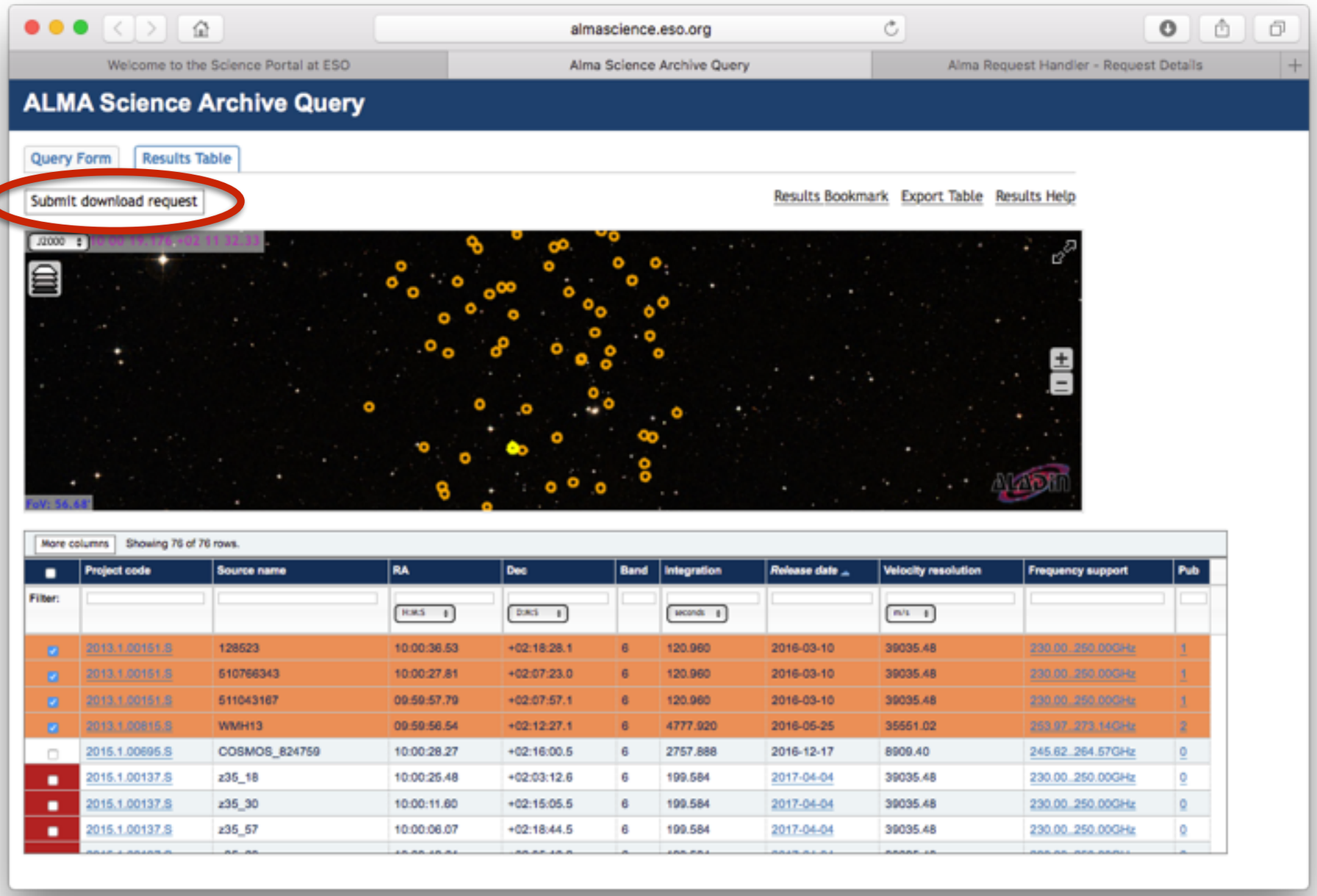
Exploring the ALMA archive



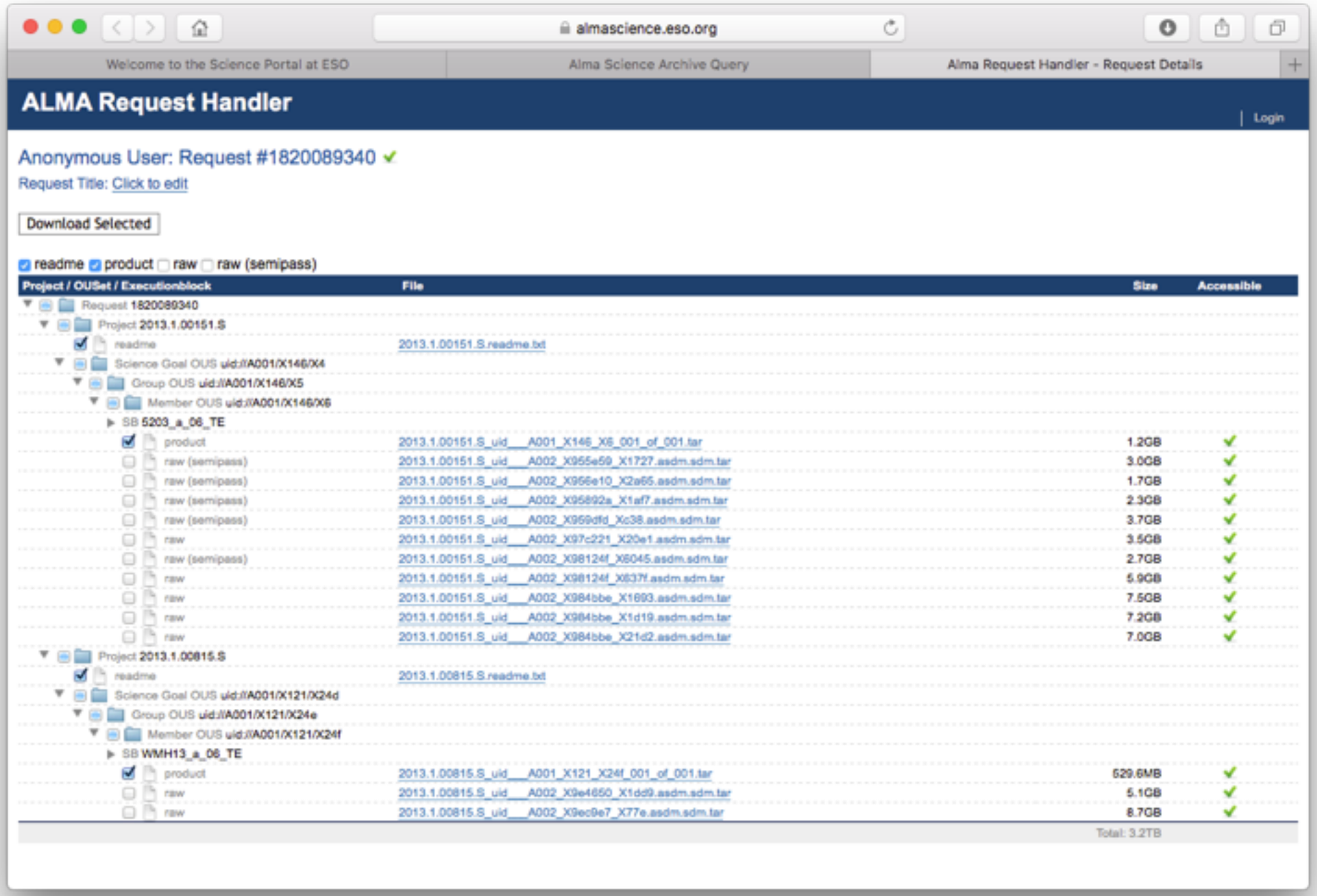
Exploring the ALMA archive



Exploring the ALMA archive



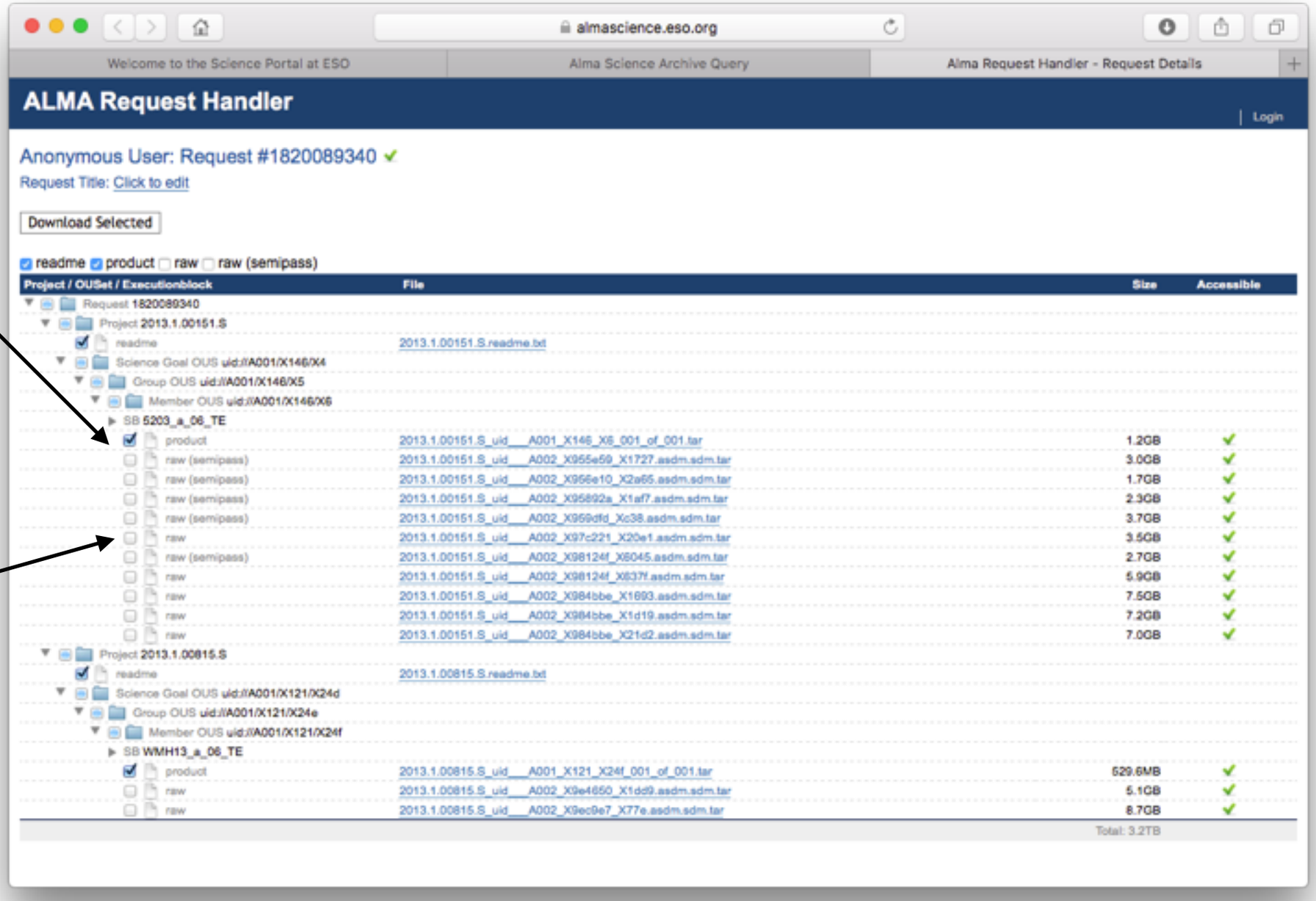
Exploring the ALMA archive



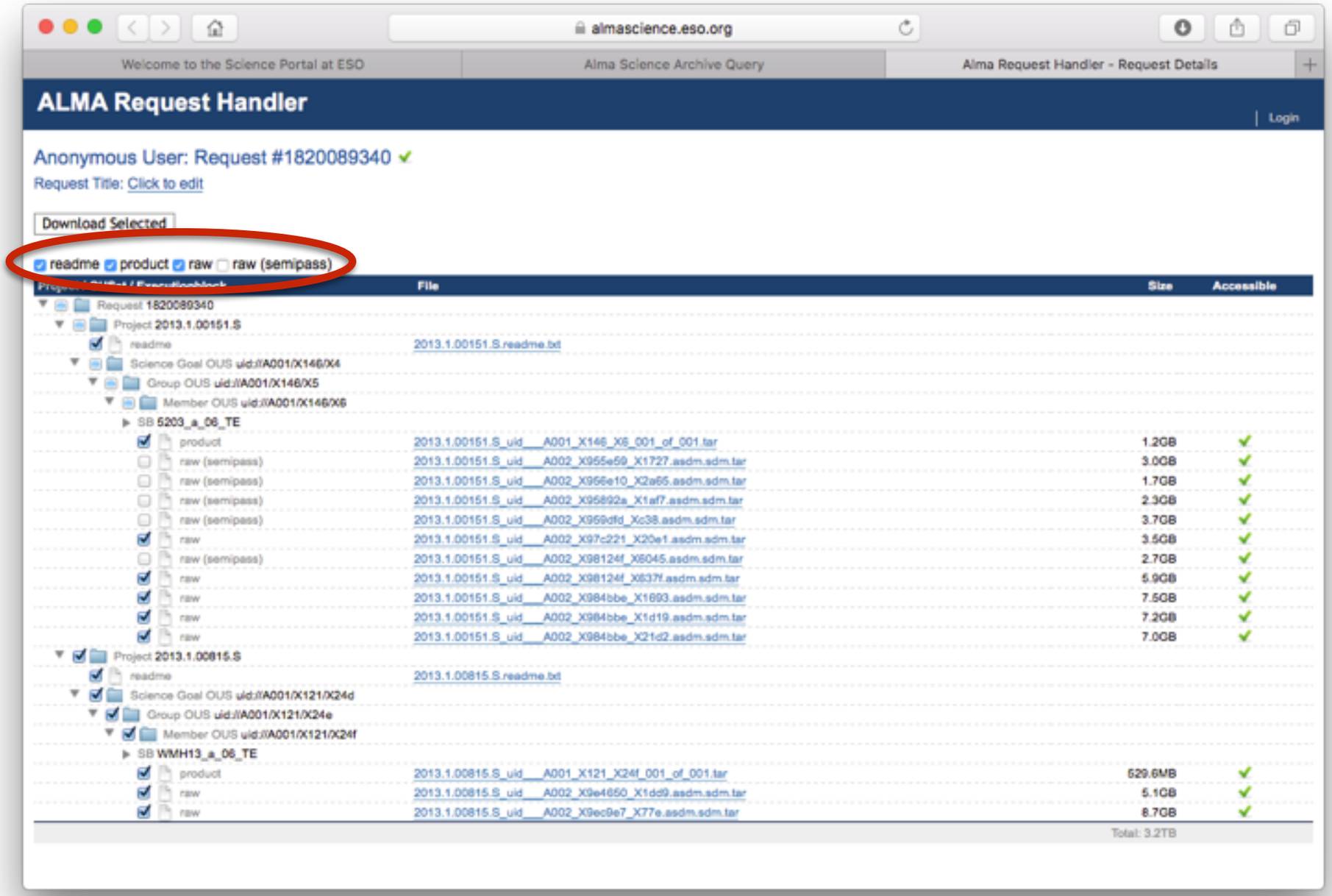
Exploring the ALMA archive

“Product” contains the scripts, images and cubes produced by the QA2 analyst

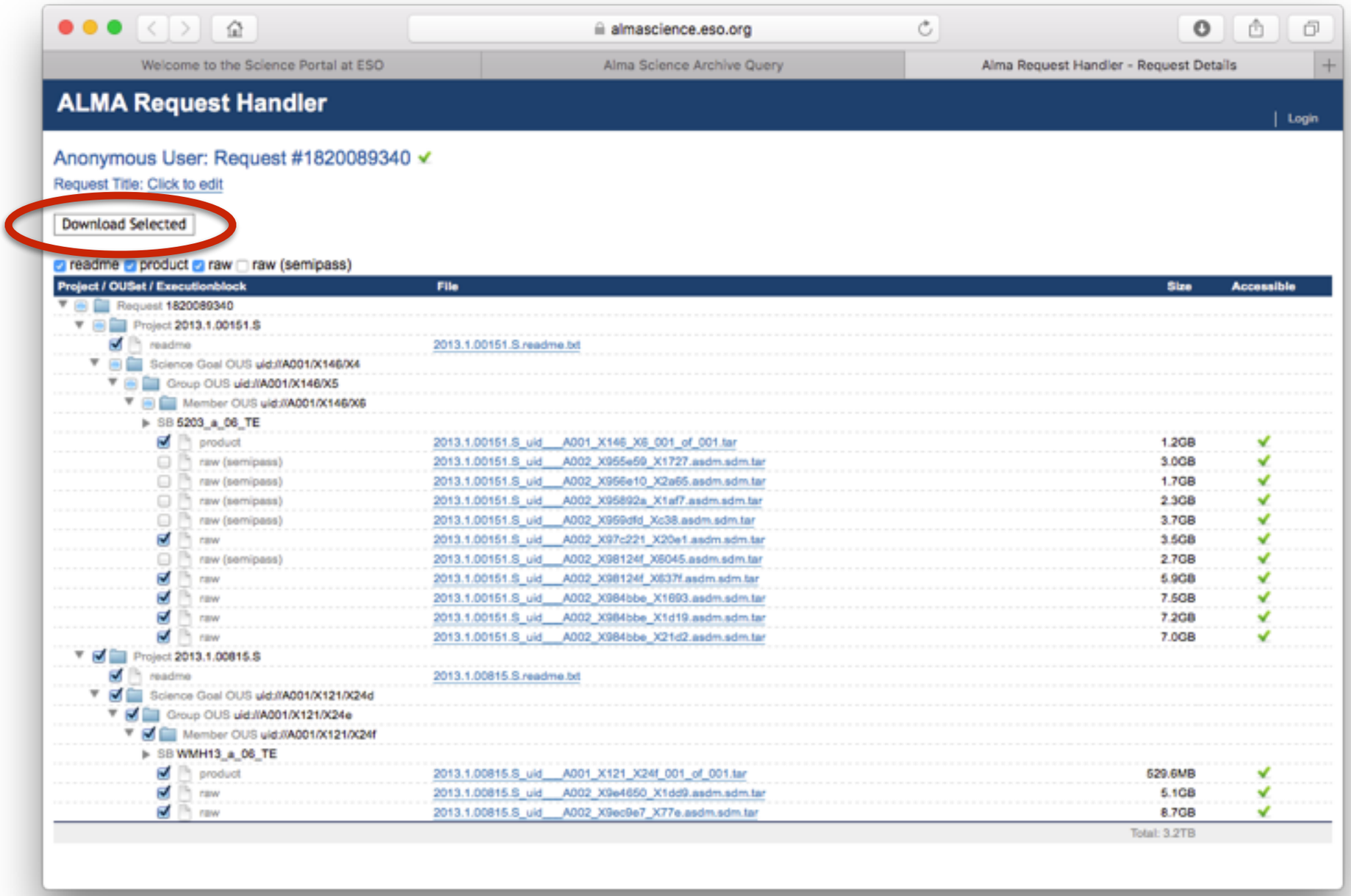
“Raw” contain the raw data, it is needed if you want to re-calibrate and/or re-image



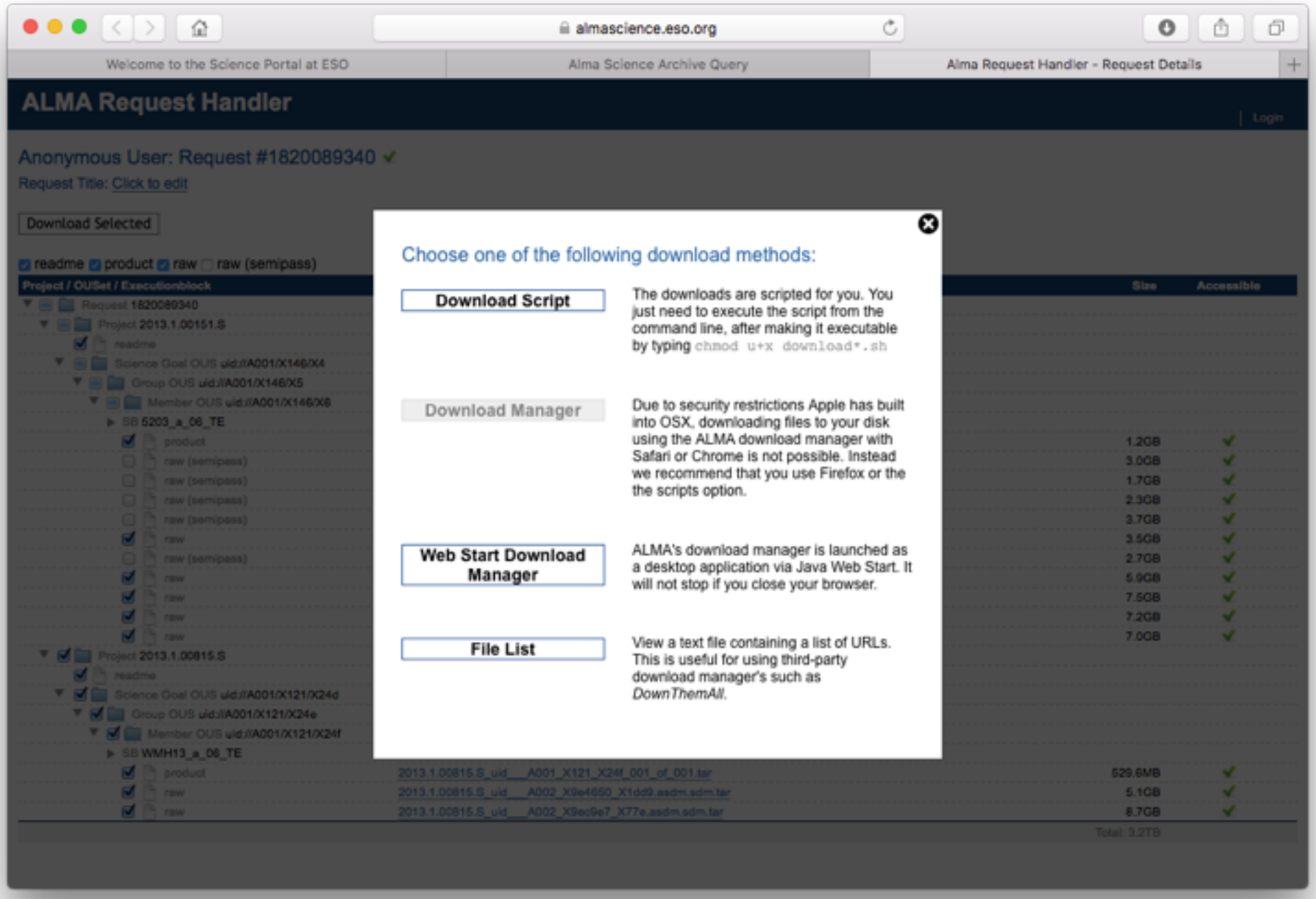
Exploring the ALMA archive



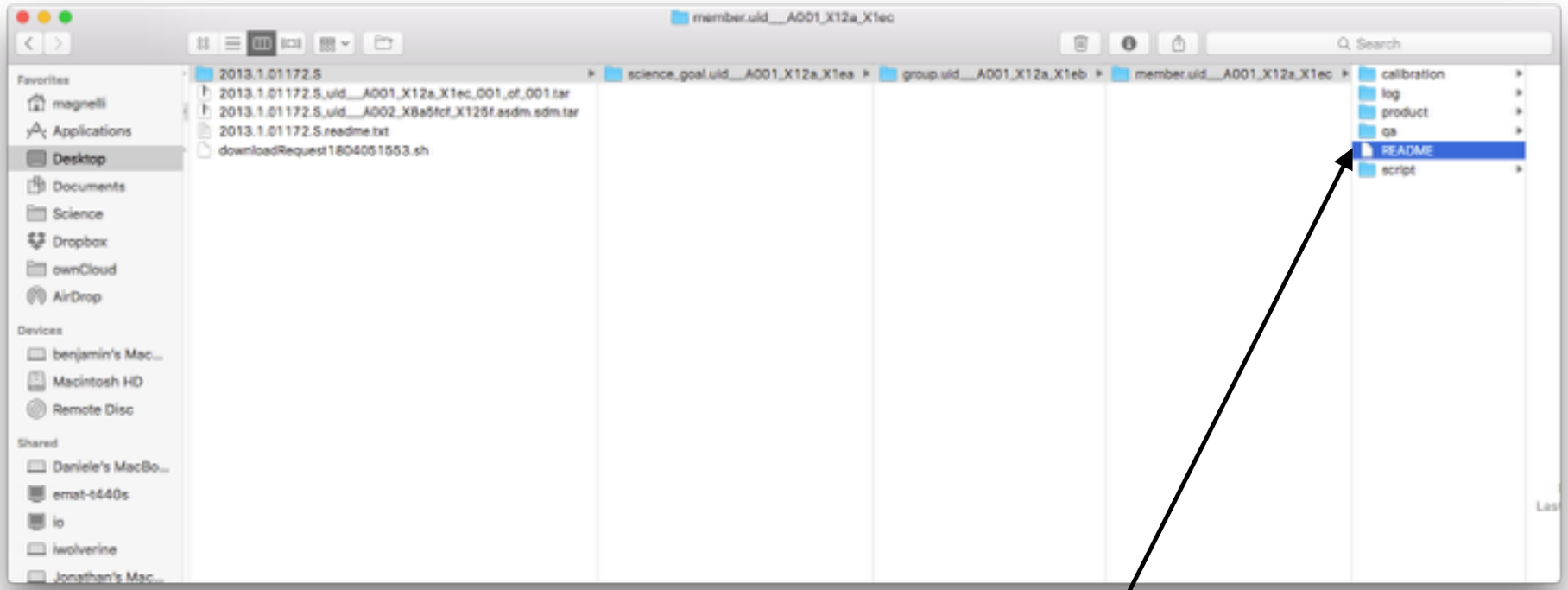
Exploring the ALMA archive



Exploring the ALMA archive

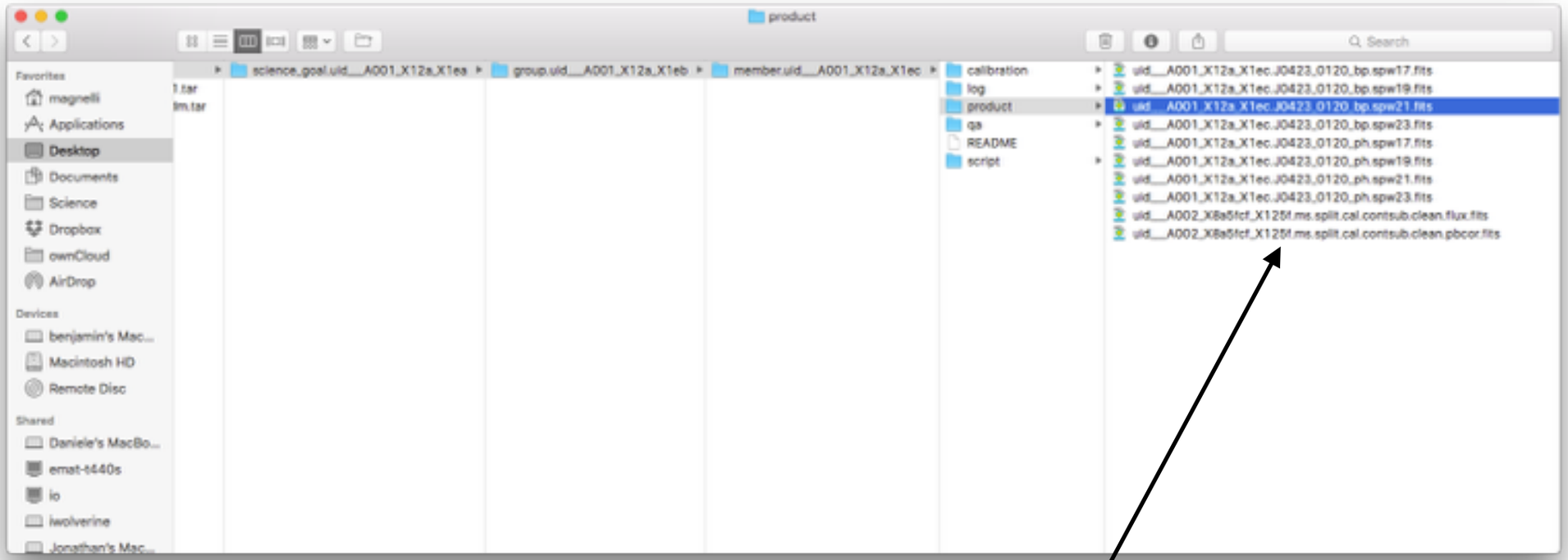


Exploring the ALMA archive



This “README” contains information about the delivered products as well as about the way to re-calibrate and/or re-image the data

Exploring the ALMA archive



Cube produced by the QA2 analyst

Exploring the ALMA archive: ASTROQUERY

The ALMA archive can also be access programatically using the external python package **ASTROQUERY**

https://astroquery.readthedocs.org/en/latest/

ASTROQUERY allows querying and retrieving data from many large observatories including ALMA.

As a simple example, querying the ALMA Archive for the source M83 and retrieving all the corresponding data is done by

```
from astroquery.alma import Alma
import numpy
# Querying the ALMA Science Archive for source M83
result = Alma.query_object('M83')
# Extracting a list of Member ObsUnitSets
member_ous = numpy.unique(result['Member ous id'])
# Creating an ALMA astroquery instance and downloading all data the identified Member OUS
myAlma = Alma()
myAlma.cache_location = '/big/external/drive/'
myAlma.retrieve_data_from_uid(member_ous[0])
```

ASTROQUERY is the preferred way to query the ALMA archive for long lists of sources. Examples on the usage of the package for ALMA data are available at :

https://astroquery. readthedocs.org/en/latest/alma/alma.html

SUMMARY

- **The ALMA archive is a ever growing database, containing already several thousands of hours of observations**

- **The web interface and astroquery package render queries and retrieval of data from the ALMA archive very easy**

- **Many scientific breakthroughs are certainly hidden is the ALMA archive !**