

Artemix : ALMA REMOTE MINING EXPERIMENT



EUROPEAN ARC
ALMA Regional Centre || IRAM

ARTEMIX

Yaye Awa Ba, Philippe. Salomé, Michel. Caillat
(LERMA)

with credits to : L. Loria, N. Kasradze

Archive and Data Mining

Artemix

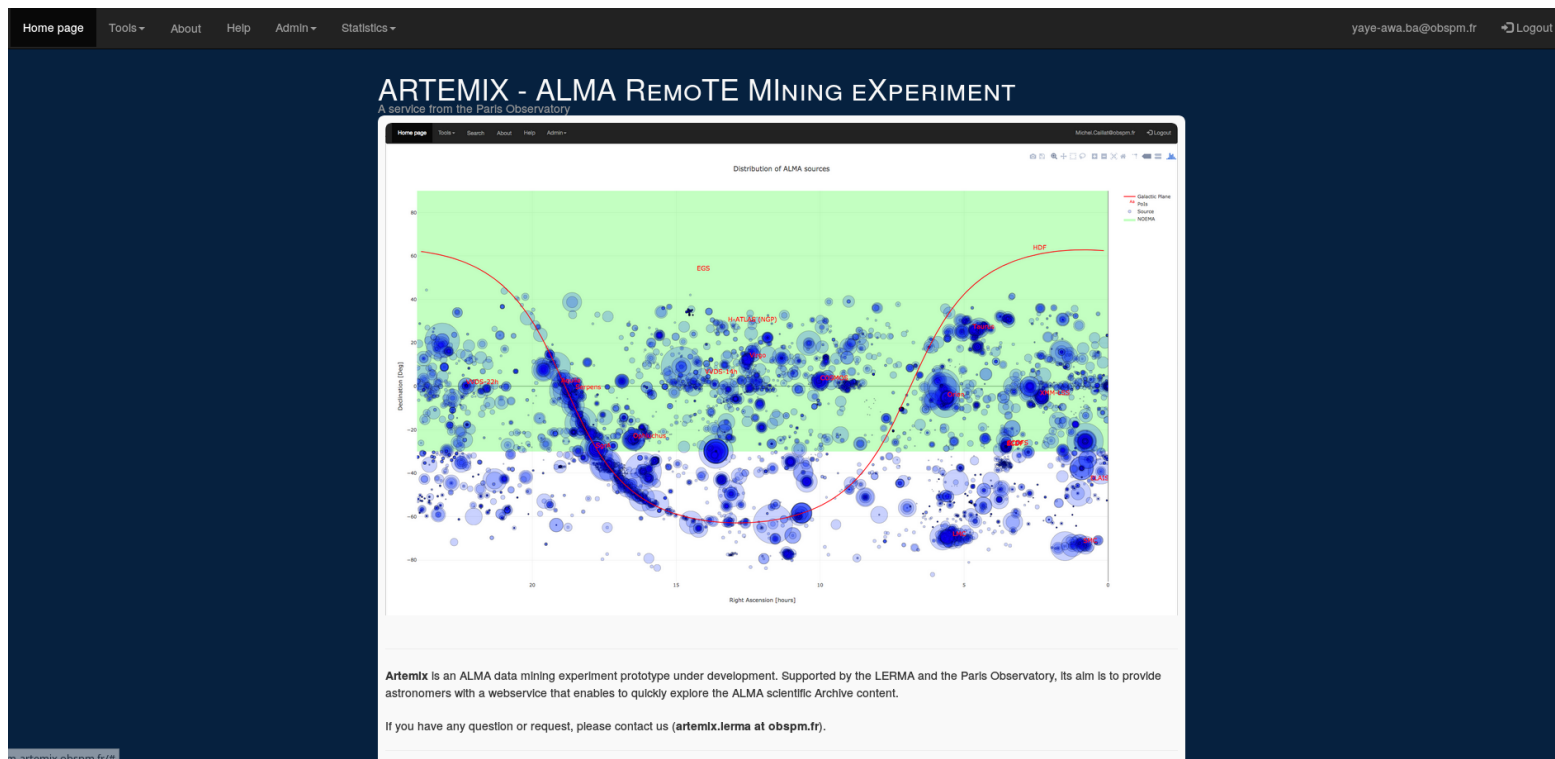
- Software development project developed in LERMA at Paris Observatory
- Pilote study for remote data inspection
- On-line data mining tools for an optimize / easy exploitation of the ALMA science archive

ALMA (Atacama Large Millimeter Array)

- Is a millimeter sub-millimeter reconfigurable interferometer
- Inaugurated in March 2013 on the Chajnantor plateau (5000m altitude) in Chile
- It operates 66 antennas
- Data are distributed to the partners (Chile, USA, Japon and Europe)
- Public scientific archives available every one and half year by a web service «ALMA Science Archive » or by python library « Astroquery »
- Delivers :
 - data cubes (fits format) with RA/DEC and the frequency as third axis produced by the reduction pipeline
 - metadata

Artemix Goals

- A web service for exploration and visualization of public archives of ALMA observations
 - Visualization interface
 - Search tool (by sesame name, project code ...)

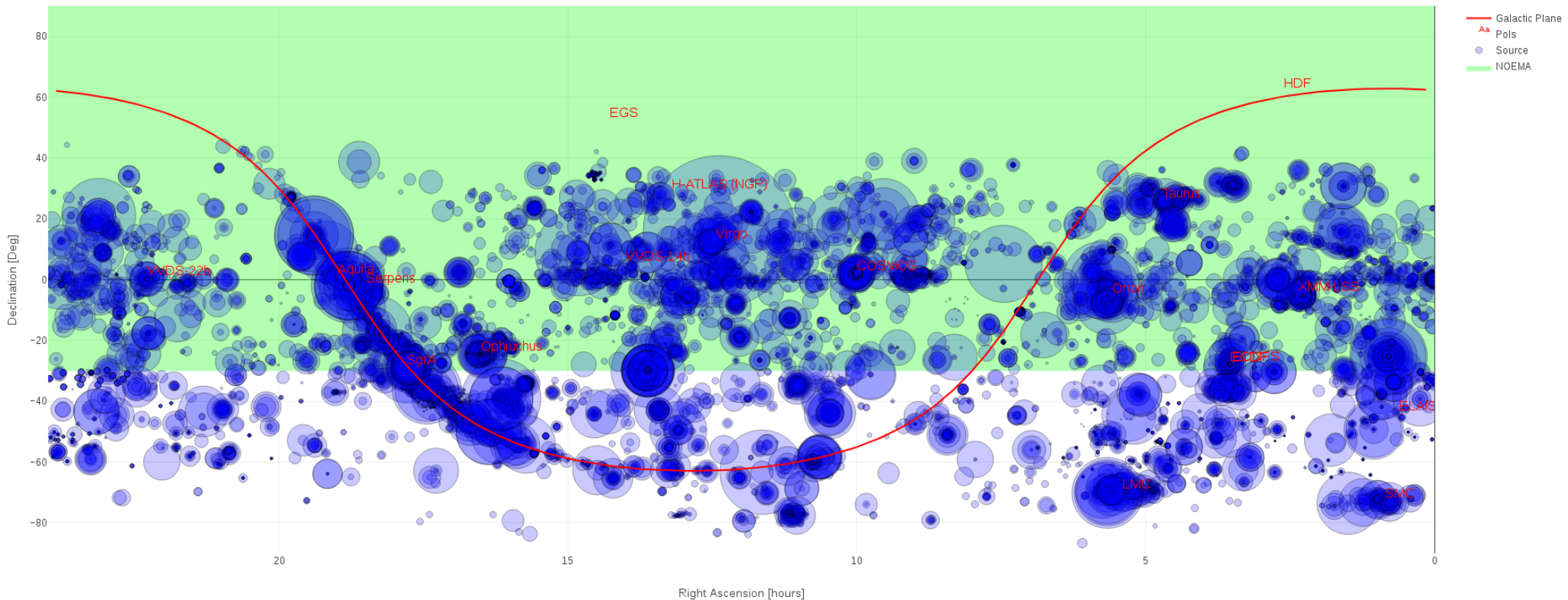


Demo : all observed sources

Blue circles are all the observed sources

The circle radius depends on the observation time

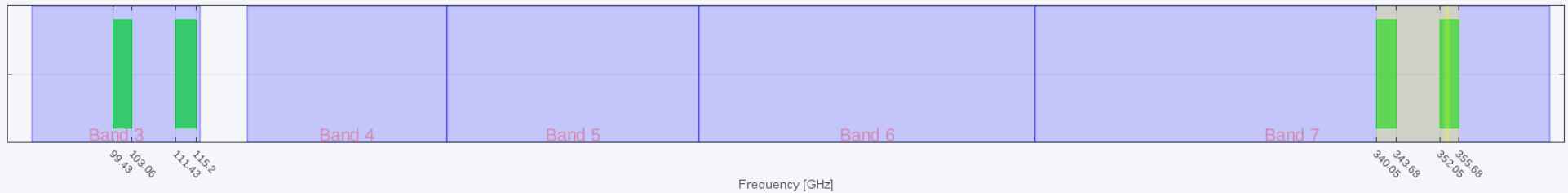
Distribution of ALMA sources



Demo : Frequency coverage

Green bands : frequency range for which observation are planned

Yellow bands : correspond to data which have been actually imaged



Redshift Velocity (km/s) [All Info](#) [Show Metadata](#)

Show all data. Initially shown are the data of which the filename ends with ".pbcor.fits", ".pbcorr.fits", ".image.fits", ".line.fits", ".cont.fits" or ".clean.fits"

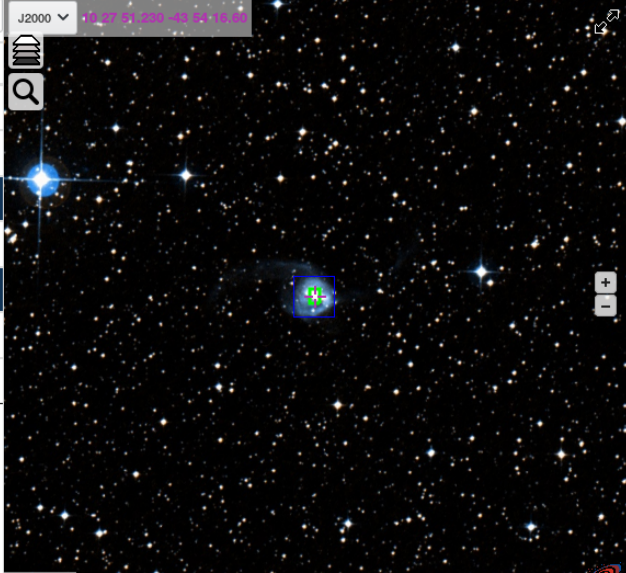
Show entries

Search:

#	Fits file	Target	RA	DEC	Cube size	Freq. range	Proj. code	uid
1	Info	NGC3256	10:27:51.23	-43:54:16.60	200x200x1	99.563 --- 115	2011.0.00525.S	N/A
2	Info	NGC3256	10:27:51.23	-43:54:16.60	200x200x88	342.047 --- 343.051	2011.0.00525.S	N/A
3	Info	NGC3256	10:27:51.23	-43:54:16.60	200x200x1	340.157 --- 355.659	2011.0.00525.S	N/A
4	Info	NGC3256	10:27:51.23	-43:54:16.60	200x200x36	353.234 --- 353.65	2011.0.00525.S	N/A
5	Info	NGC3256	10:27:51.23	-43:54:16.60	360x360x57	353.157 --- 353.835	2011.0.00525.S	N/A
6	Info	NGC3256	10:27:51.23	-43:54:16.60	360x360x1	340.148 --- 355.65	2011.0.00525.S	N/A
7	Info	NGC3256	10:27:51.23	-43:54:16.60	360x360x88	342.067 --- 343.082	2011.0.00525.S	N/A

Showing 1 to 7 of 7 entries

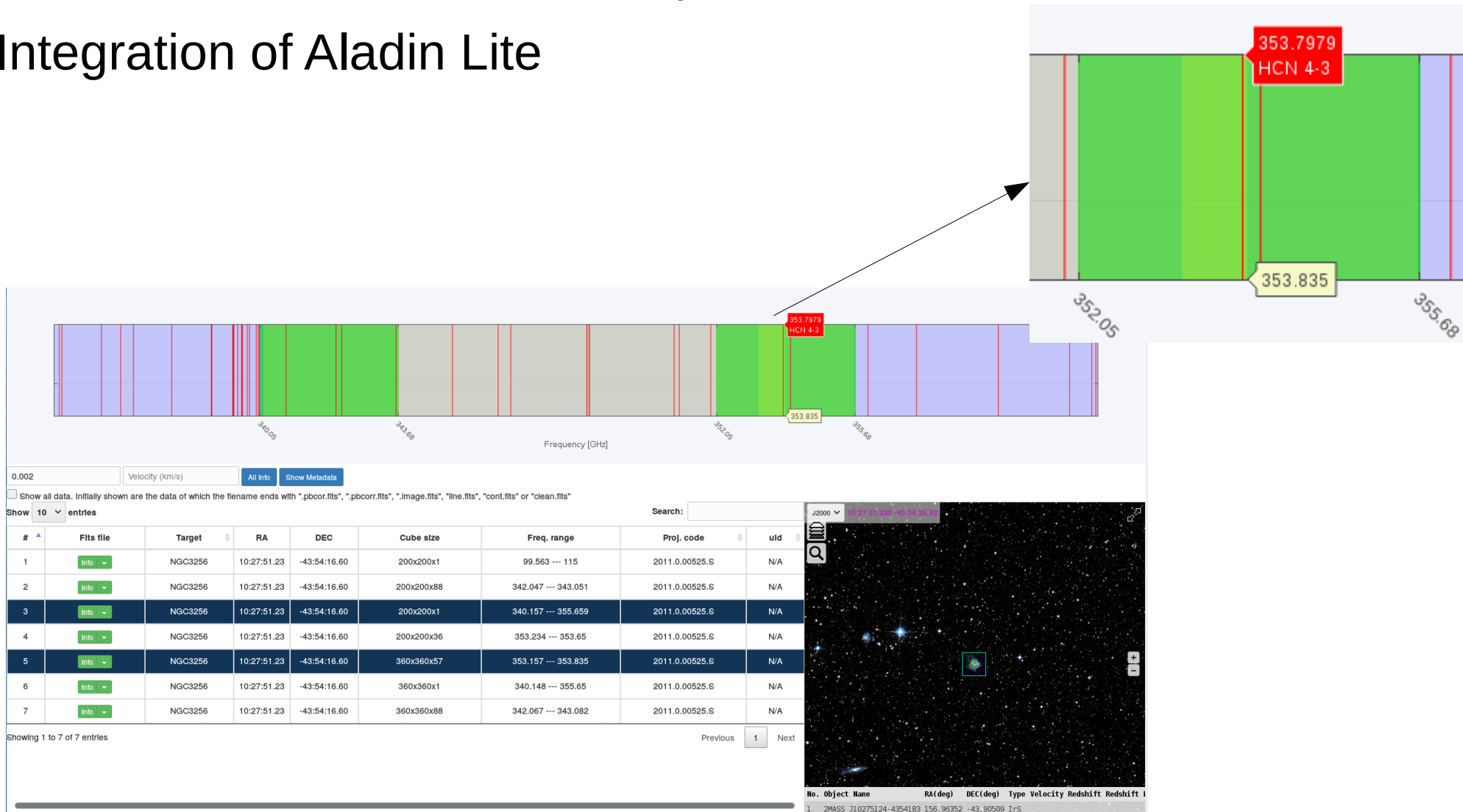
Previous Next



Demo : Frequency coverage

With the redshift we can see spectral lines and molecules

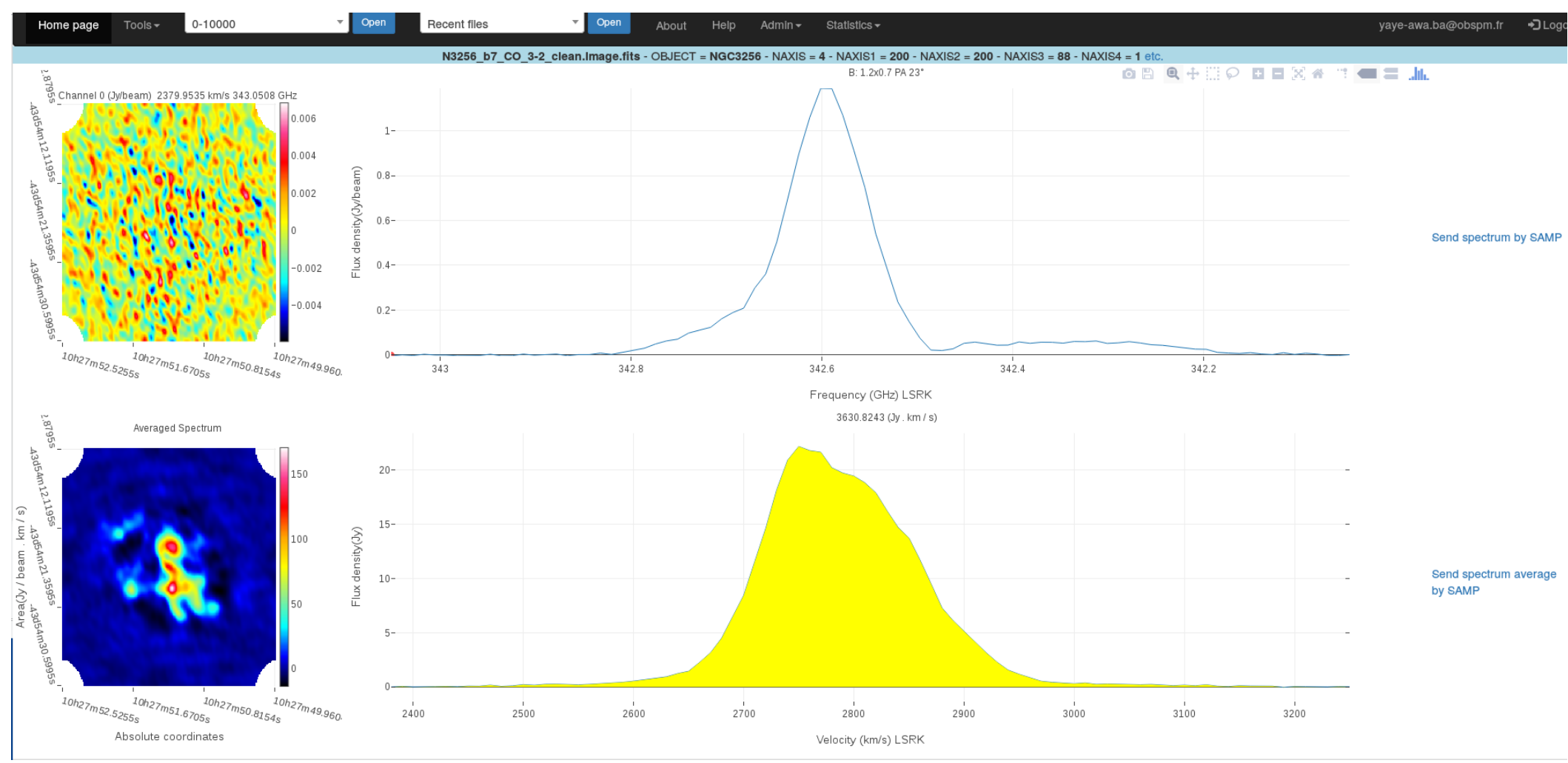
Integration of Aladin Lite



Demo: Data cube exploration

Two representations of spectrum

Spectrum can be send to Cassis by the SAMP protocol



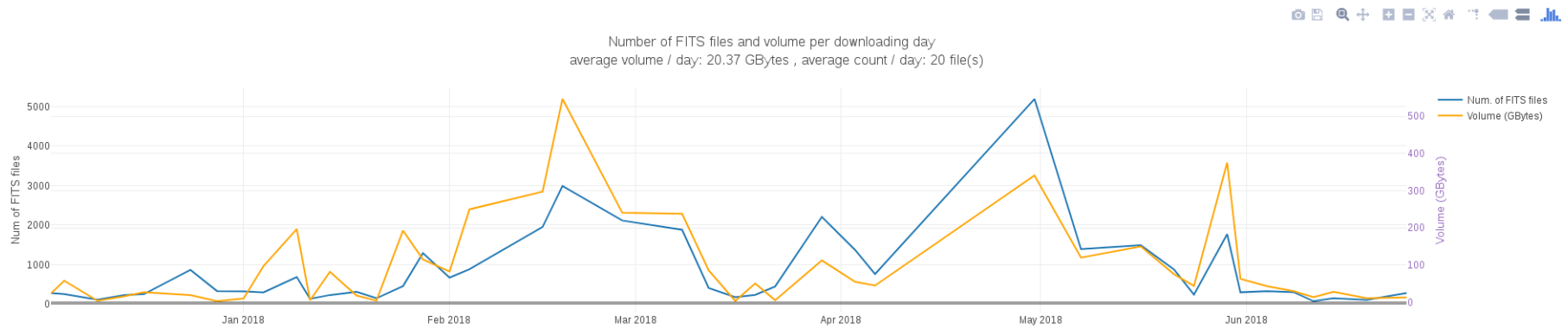
Demo: Statistics

FITS files - Statistics

The table below shows the number and the accumulated size of Alma Science Archive FITS files whose sizes are above different thresholds for each year between 2011 and 2017.

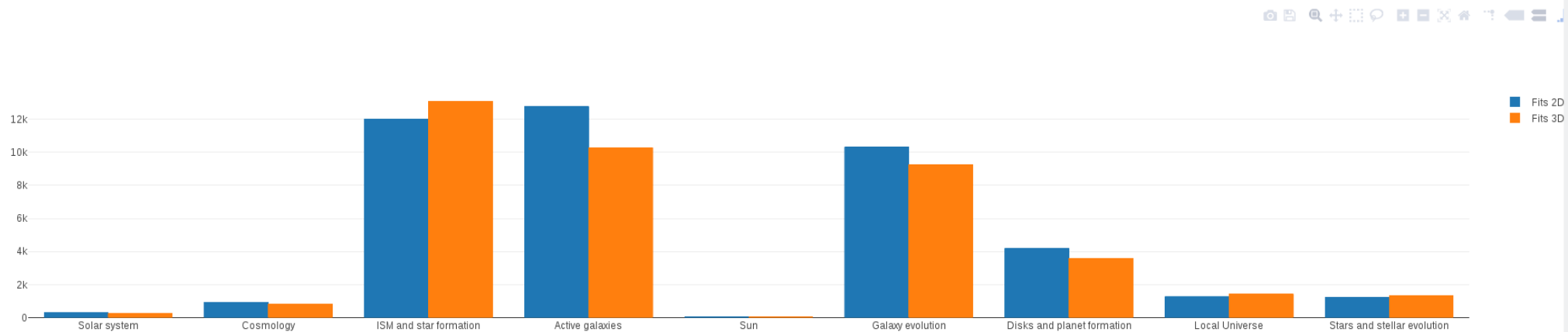
	2011	2012	2013	2014	2015	2016	2017	Total per threshold
>0Gb	2037 files / 88.94Gb	6245 files / 791.78Gb	22899 files / 3383.07Gb	0files / 0Gb	35322 files / 6736.68Gb	35704 files / 4506.03Gb	0files / 0Gb	102207 files / 15506.5 Gb
>1Gb	11 files / 19.3Gb	136 files / 414.46Gb	833 files / 2234.03Gb	0files / 0Gb	1081 files / 5802.71Gb	1106 files / 3275.4Gb	0files / 0Gb	3167 files / 11745.9 Gb
>2Gb	0files / 0Gb	74 files / 327.23Gb	332 files / 1538.84Gb	0files / 0Gb	711 files / 5260.72Gb	432 files / 2355.33Gb	0files / 0Gb	1549 files / 9482.12 Gb
>3Gb	0files / 0Gb	54 files / 276.17Gb	198 files / 1199.44Gb	0files / 0Gb	452 files / 4667.26Gb	205 files / 1825.3Gb	0files / 0Gb	909 files / 7968.17 Gb

Projects of year 2016

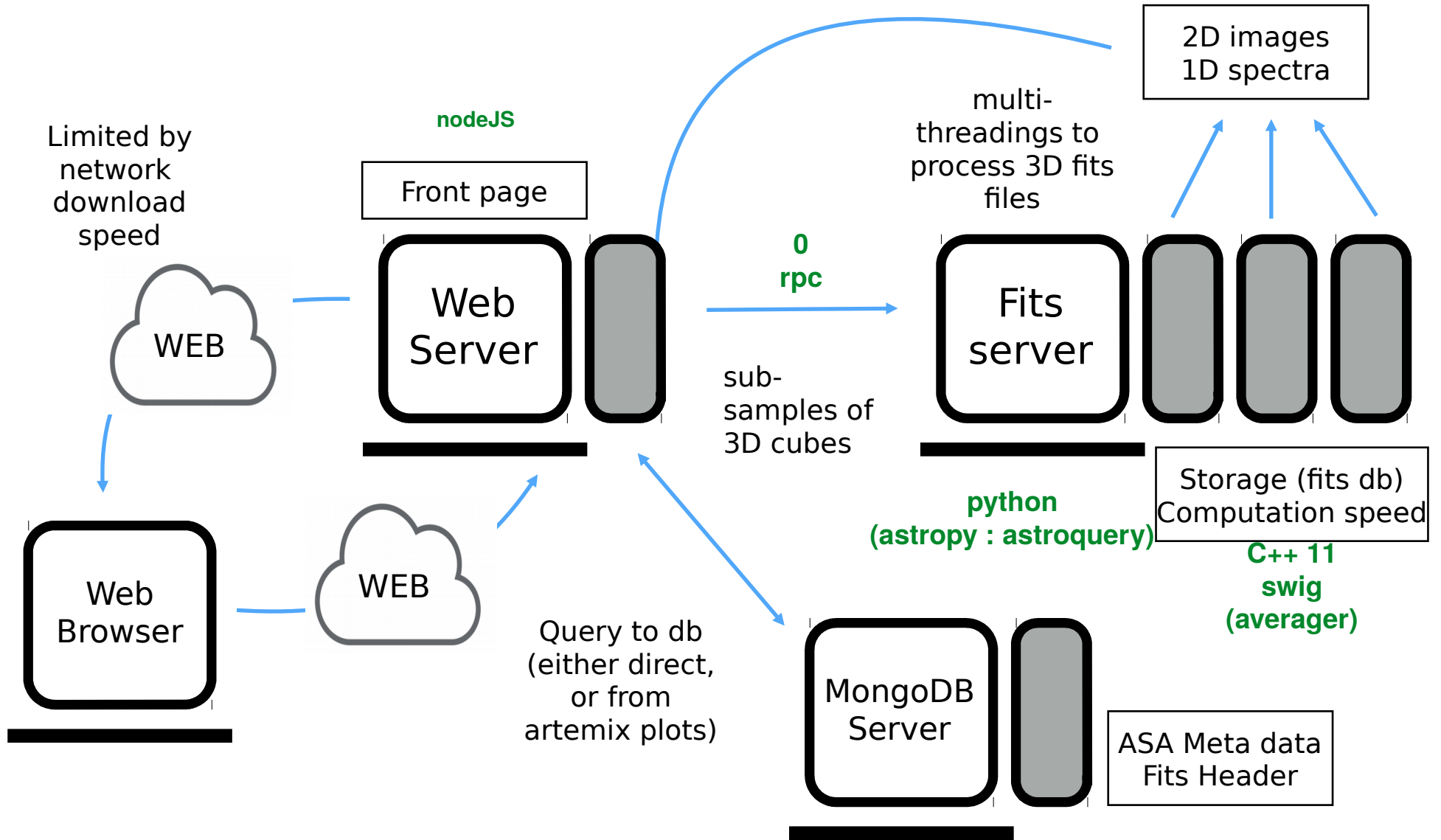


FITS files - category statistics

The graphics below show the type and the number of Alma Science Archive FITS files in each science category.



Artemix : architecture



Next steps

- Improve performance with large datasets. Ex : 2500x2500x500
- Integration of machine learning techniques to detect automatically sources, flux, energies ...
- Connexion with VAMDC project for the retrieval of spectral lines and molecule names.

Let's work together

We are interested in working with other VO-tools

- Artemix web service : <http://artemix.obspm.fr/>
- Contact : philippe.salome[at]obspm.fr

Thank you