



DATA MINING TOOLS

S.BURKUTEAN & M. MASSARDI



Request access to the ARC cluster!

arc.ia2.inaf.it/arccluster

The screenshot shows the top navigation bar of the website. On the left are browser controls (back, forward, refresh, home). The address bar contains the URL "arc.ia2.inaf.it/arccluster.html". To the right of the address bar are zoom controls (80%), a three-dot menu, a download icon, and a star icon. A search bar on the far right contains the text "CARTA ALMA viewer" with a right-pointing arrow.

The main header features the ALMA logo and the text "EUROPEAN ARC ALMA Regional Centre | Italian". Below the header is a navigation menu with links: homepage, tools, activities, support, arc cluster (which is highlighted in blue), and about us.

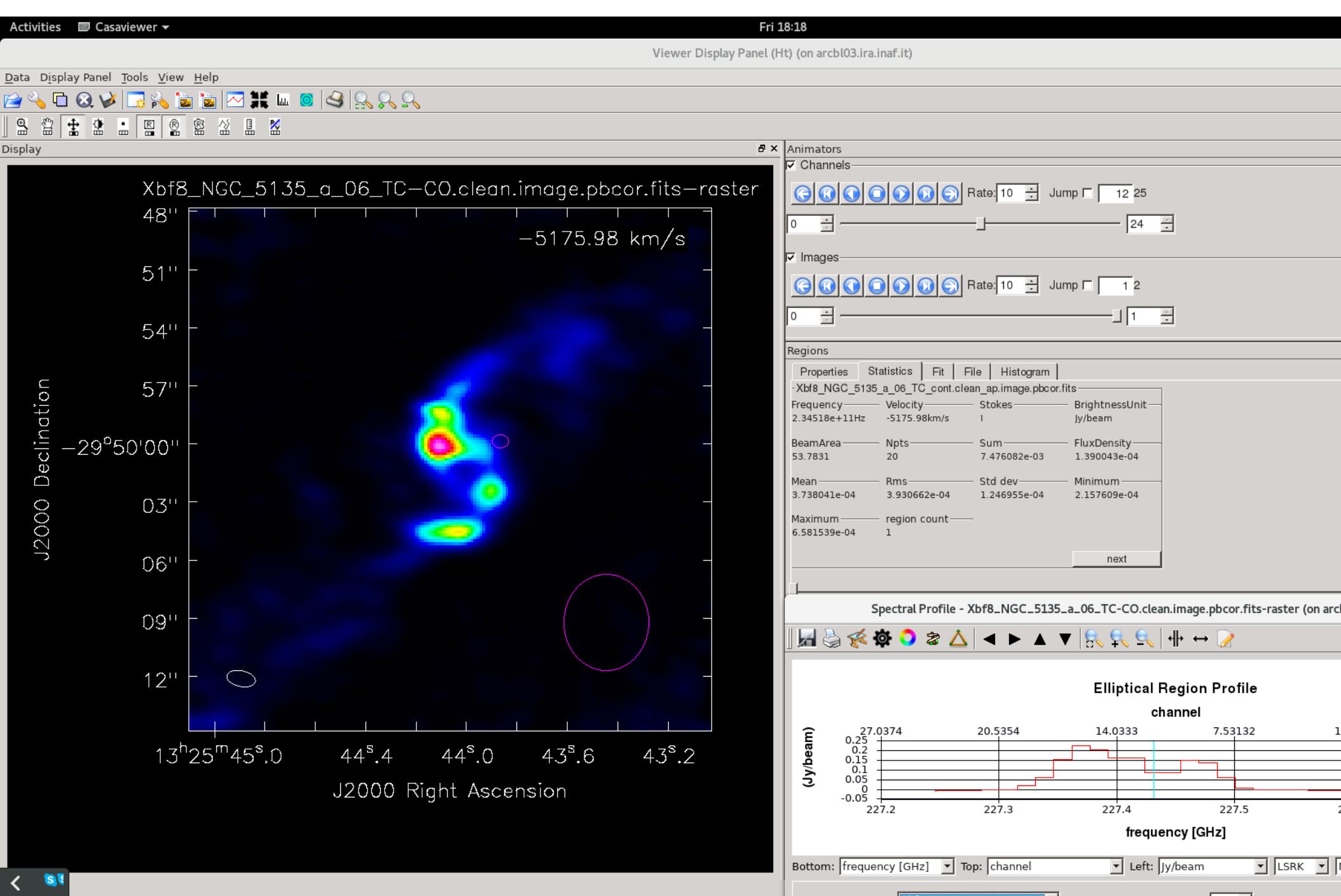
Software available on ARC cluster could be listed here

Nodes

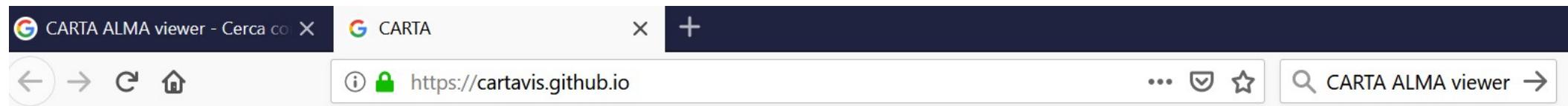
Software package	setup command
CASA	casapy-setup
Miriad	miriad-setup
aips	aips-setup
analysis utils	analysisUtils-setup
analytic infall	analytic_infall-setup
astron	astron-setup
Coyote library	coyote-setup
fits Viewer	fv-setup
GCC Compiler	gcc-setup
Gildas	gildas-setup
Healpix	healpix-setup
IDL	idl-setup
heasoft	heasoft-setup
QA2	qa2-setup
Ratran	ratran-setup
Starlink	starlink-setup

Name	RAM	CPU	Cores	Clock	Data Net	Work Disk	Scratch Disk	scheduler	groups	
arcbl01	32G	Intel Xeon E5-2637 v3	2 x 4/8	3500	10GbE	1TB	1TB	N	*	
arcbl02	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	reserved
arcbl03	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	reserved
arcbl04	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	
arcbl05	8G	AMD Opteron 2352	8	2100	10GbE	1T		Y	*	
arcbl06	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	
arcbl07	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	
arcbl08	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	
arcbl09	8G	AMD Opteron 2352	8	2100	10GbE	34G		Y	*	
arcbl10	32G	Intel Xeon E5-2637 v3	2 x 4/8	3500	10GbE	1T	1T	N	arc-staff, arc-vlbi	
arcbl11	8G	AMD Opteron 2352	8	2100	10GbE	34G		N	arc-staff, arc-vlbi	
arcbl12	16G	AMD Opteron 2352	8	2100	10GbE	34G		N	arc-staff, arc-vlbi, arc-f2f	
arcbl13	16G	AMD Opteron 2387	4	2800	10GbE	136G		N	arc-staff, arc-vlbi, arc-f2f	
arcbl17	64G	AMD Ryzen 7 1800X	8/16	3600	1GbE	3,5TB		N	arc-staff, arc-vlbi	
arcbl18	64G	Intel Xeon E3-1275 v6	4/8	3800	10GbE	22T	57G	N	arc-staff, arc-vlbi	
arcbl19	64G	Intel Xeon E3-1275 v6	4/8	3800	10GbE	11T	57G	N	arc-staff, arc-vlbi, arc-f2f	
arcbl20	256G	Intel Xeon E5-1650 v4	6/12	3600	10GbE	11T	65G	N	arc-staff, arc-vlbi	
arcbl21	64G	Intel Xeon E3-1275 v6	4/8	3800	10GbE	11T	57G	N	arc-staff, arc-vlbi, arc-f2f	

The CASA viewer



The CARTA viewer



CARTA

Cube Analysis and Rendering Tool for
Astronomy

The CARTA team

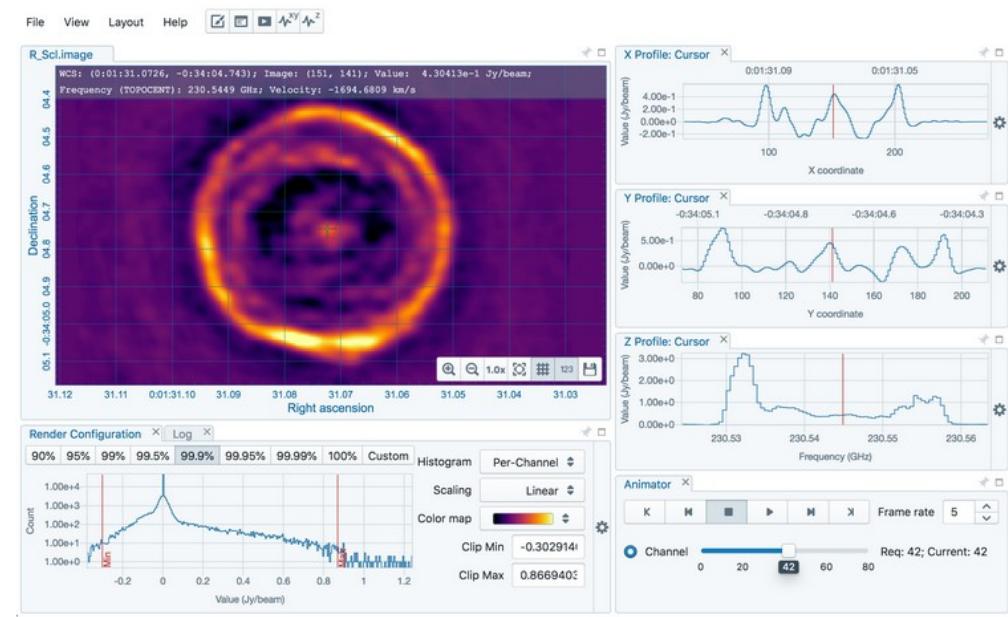
CARTA is developed by
[Academia Sinica Institute of Astronomy
and Astrophysics \(ASIAA\)](#)
[Inter-University Institute for Data Intensive
Astronomy \(IDIA\)](#)
[National Radio Astronomy Observatory
\(NRAO\)](#)
[Department of Physics, University of
Alberta](#)



Inter-University Institute
for Data Intensive Astronomy

CARTA

CARTA is the Cube Analysis and Rendering Tool for Astronomy, a new image visualization and analysis tool designed for the [ALMA](#), the [VLA](#), and the [SKA pathfinders](#). As the image size increases drastically with modern telescopes in recent years, viewing a image with a local image viewer or with a remote image viewer via the ssh protocol becomes less efficient. The mission of CARTA is to provide usability and scalability for the future by utilizing modern web technologies and computing parallelization.



The Japanese Virtual Observatory

<http://jvo.nao.ac.jp/portal/alma.do>

Link to the products of the ALMA Science Archive (no reprocessing done!)
 Updated simultaneously to the archive. Fully searchable.

ALMA FITS Archive

Using the data for publication

The following statement should be included in the acknowledgment of papers using the ALMA datasets obtained from the JVO portal:

"This paper makes use of the following ALMA data: ADS/JAO.ALMA#<Project code>. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) , NAOJ (Japan), and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ."

You can find the project code (e.g. 2011.0.01234.S) on the dataset info page where you download the data.

Please also include the following sentence on the title page as a footnote to the title or in the acknowledgment of the paper.

"[Part of] the data are retrieved from the JVO portal (<http://jvo.nao.ac.jp/portal>) operated by the NAOJ"

- 2019-02-07: **FITS WebQL v3 will end** at the end of March 2019.
- 2018-10-17: **FITS WebQL v4 was released**: New feature '[FITS Cube Slicer](#)' is available.

Search Results						
Target Name	Simbad Name	Project Code	Coords	Frequency	Advanced	Download
Nombre of Projects found : 1982 Clear all the filters						
#	Project Code	# of Data	Title	Category	Last Update	
1	2017.1.00555.S	59	Nitrogen chemistry in Eta Carinae	ISM and star formation	2019-02-25	
2	2017.1.01500.S	7	The diffuse molecular component in the nuclear bulge of the Milky Way	ISM and star formation	2019-02-25	
3	2016.1.01346.S	2161	Galactic Census of All Massive Starless Cores within 5 kpc	ISM and star formation	2019-02-25	
4	2017.1.00270.S	24	Dust vs. CO: Do both trace molecular gas emission in high-redshift galaxies?	Galaxy evolution	2019-02-25	
5	2017.1.00886.L	16	100,000 Molecular Clouds Across the Main Sequence: GMCs as the Drivers of Galaxy Evolution	Galaxy evolution	2019-02-25	
6	2017.1.00888.S	22	A puzzling outflow in a QSO a z=2.4	Active galaxies	2019-02-25	
7	2017.1.00629.S	190	Constraining the cold accretion onto the most massive Black Holes	Active galaxies	2019-02-25	
8	2016.1.00671.S	5	Revisiting the star formation efficiency of low-metallicity starburst galaxy NGC5253	Active galaxies	2019-02-24	
9	2016.1.00119.S	9	Probing the distribution of H2CO and warm SO2 toward the O-rich CSE of IKTau	Stars and stellar evolution	2019-02-24	
10	2017.1.01722.S	922	Spectral scans for high CO transition lines in z>6.5 QSO candidates selected from PanSTARRS1 & WISE	Active galaxies	2019-02-24	

The Japanese Virtual Observatory

<http://jvo.nao.ac.jp/portal/alma.do>

ALMA Science Portal at ESO X ALMA Archive : Target Info X +

Back Forward Home Address: jvo.nao.ac.jp/portal/alma/archive.do?action=target.info&orderBy=&order ... Star Search Cerca

JVO Ver.2 | Top | Search | VOServices | Subaru | ALMA | JVOSpace Login

p11 ver.181211 News I am a guest

=> Location: [Top Page](#) > [ALMA](#) > [FITS Archive](#) > [Target Info](#)

ALMA FITS Archive : Target Info

Target Name : NGC5135

Filter by Frequency

Show all the data including calibration (*.flux.fits, *.pb.fits, target=J#####[+-]####), duplicated, and deprecated data.

Number of data per page : 20

Ordered by dataset_id (desc)

Total number : 4

* WebQLv3 will end at the end of March in 2019.

#	dataset id	project code	<input type="checkbox"/> all	Download all the checked data	image	spect	ra/dec (J2000)	Cube size (XxYxF) ?	Image size (arcmin2)	pixel scale, beam size (arcsec)	band	freq. range (GHz)
1	ALMA01081334	2013.1.00524.S	<input type="checkbox"/>	Download WebQLv4 WebQLv3*			13h25m43.9 -29d50m01	600 x600 x1 x1	0.33 x0.33	0.033, 0.170 x0.143	Band9	677.243 -- 684.737
2	ALMA01081333	2013.1.00524.S	<input type="checkbox"/>	Download WebQLv4 WebQLv3*			13h25m43.9 -29d50m01	600 x600 x1 x1	0.33 x0.33	0.033, 0.170 x0.143	Band9	677.243 -- 684.737

The ALMA Data Mining Toolkit

<http://admit.astro.umd.edu/admit/>

The ALMA Data Mining Toolkit

The ALMA Data Mining Toolkit (ADMIT) is a value-added Python software package which integrates with the ALMA archive and CASA to provide scientists with quick access to traditional science data products such as moment maps, as well as with new innovative tools for exploring data cubes and their many derived products. The goals of the package are to:

- make the scientific value of ALMA data more immediate to all users
- create an analysis infrastructure that allows users to build new tools
- provide new types of tools for mining the science in ALMA data
- increase the scientific value of the rich data archive that ALMA is creating
- re-execute and explore the robustness of the initial pipeline results.

ADMIT is funded as an ALMA Development Project, and was introduced as version 1.0 in May 2016, with a final delivery to NRAO in October 2016 as version 1.1, but under continued development in git.

Set of scriptable tools that could be customized to analyze fits images.

Example:
Continuum subtraction script

```
#!/usr/bin/env casarun
# set up admit in the casa environment
import admit
# Master project.
p = admit.Project('you-name-choice.admit', dataserver=True)
# Flow tasks.
t0 = p.addtask(admit.Ingest_AT(file='your-image-cube-name.fits'))
t1 = p.addtask(admit.CubeStats_AT(ppp=True), [t0])
t2 = p.addtask(admit.CubeSum_AT(numsigma=5.0, sigma=99.0), [t0, t1])
t3 = p.addtask(admit.CubeSpectrum_AT(), [t0, t2])
t4 = p.addtask(admit.LineSegment_AT(csub=[0, 0], minchan=4, maxgap=6, numsigma=5.0), [t1, t3])
t5 = p.addtask(admit.ContinuumSub_AT(fitorder=1, pad=60), [t0, t4])
t6 = p.addtask(admit.CubeStats_AT(ppp=True), [t5])
t7 = p.addtask(admit.CubeSpectrum_AT(), [t5, t6])
t8 = p.addtask(admit.Moment_AT(mom0clip=2.0, numsigma=[3.0]), [t5, t6])
t9 = p.addtask(admit.LineID_AT(csub=[0, 0], minchan=4, maxgap=6, numsigma=5.0), [t6, t7])
t10 = p.addtask(admit.LineCube_AT(pad=40), [t5, t9])
t11 = p.addtask(admit.Moment_AT(mom0clip=2.0, moments=[0, 1, 2]), [t10, t6])
t11 = p.addtask(admit.CubeSpectrum_AT(), [t10, t11])
p.run()
```

The ALMA FITS Keywords Filler project Scope

Liuzzo et al. subm

- Focus on *FITS headers*
- they should ***collect all the information*** that an archive user might want to search for in order to quickly select a dataset, compare several datasets or discard useless ones
- From FITS Standards documents, we identify a ***list of kwds could benefit a typical archival user***
- ***Ideally: complete FITS header in the Archive***

FITS keywords

Two main categories:

1. data acquisition and reduction kwds

(info from ms)

RA_TARG, DEC_TARG, UVRANGE, BAND, MINPRLBL, MAXPRBL, MAXANGSCALE, PADLIST, NANT12M, NANT7M, NANTPP,
MINELT12, MINELT7, MINELTP, DATE-OBS, DATE-END, MJD-OBS, MJD-AVG

2. Image description kwds

(info from images)

BNDRES, CHNRMS, SPATRES, STOKES, DATAMAX, DATAMIN, **DYNRANGE**, BNDCTR, BNDWID, **FLUXTOT**

For all kwds we provide a clear definition and we provide
missing FITS kwds

ALMA Keywords Filler project (Liuzzo et al. 2019)

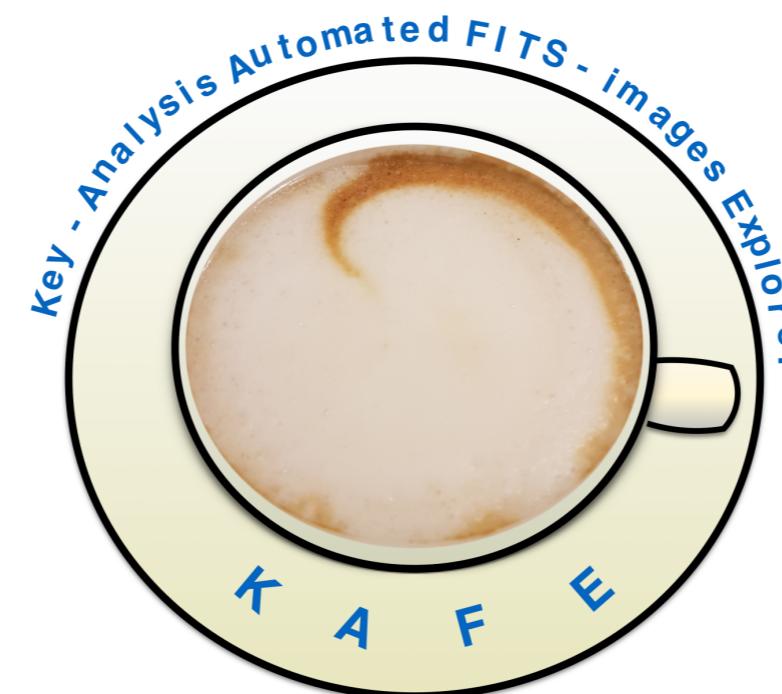
- Python code implemented in a CASA task

```
-----> inp()
# FITSkwd :: FITS keywords
imName          =      ''          # Input image name
kwdlist         =      ''          # Select kwds: '' ==> all kwds will be
                                # calculated
outfile         =      ''          # Name of disk file to write output,
                                # ''==>to terminal
include          =      ''          # Include the calculated kwds in the
                                # FITS header, True or False. Default
                                # is ''==> False
```

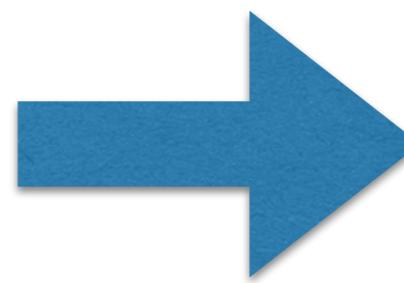
- It returns a dictionary of kwds that could be included in metadata
- The FITS code is working on different types of data,
NOT ONLY ALMA e.g. VLA, VLBA, EVN

KAFE: automated FITS image analysis + visualisation

*Burkutean et al., J. Astron.
Telesc. Instrum. Syst. 4(2), 028001 (2018)*



user-generated/archival
FITS-images



“a few clicks”

image analysis
for ALMA, JVLA, PdB etc.

The KAFE project

send this file: Choose Files No file chosen

filters

POS RANGE	0	CNTRFREQ RANGE	0	FREQRES RANGE	0
ANGRES RANGE	0	CHANRMS RANGE	0	FLUX TOTAL RANGE	0

requested keywords

ALL	<input type="checkbox"/>
RA_centre	<input type="checkbox"/>
DEC_centre	<input type="checkbox"/>
SPATRES	<input type="checkbox"/>
BNDCTR	<input type="checkbox"/>
BNDRES	<input type="checkbox"/>
BNDWID	<input type="checkbox"/>
CHANRMS	<input type="checkbox"/>
DYNRANGE	<input type="checkbox"/>
FLUXTOT	<input type="checkbox"/>
DATAMAX	<input type="checkbox"/>
DATAMIN	<input type="checkbox"/>
STOKES	<input type="checkbox"/>

spectrum analysis options

ALL	<input type="checkbox"/>
3D view	<input type="checkbox"/>
continuum subtraction	<input type="checkbox"/>
Channel gallery	<input type="checkbox"/>
Spectrum_3D_mask	<input type="checkbox"/>
Spectrum inner quarter	<input type="checkbox"/>
Spectrum around max	<input type="checkbox"/>
Spectral gallery	<input type="checkbox"/>
3D posvel	<input type="checkbox"/>
moments	<input type="checkbox"/>
PosVel along maj/min axis	<input type="checkbox"/>
Spectral fit	<input type="checkbox"/>
Cube morph	<input type="checkbox"/>

further analysis options

ALL (except LC,3colour)	<input type="checkbox"/>
Source detection	<input type="checkbox"/>
Source detection SNR layer	<input type="checkbox"/>
radial average	<input type="checkbox"/>
Image cuts	<input type="checkbox"/>
power spectrum	<input type="checkbox"/>
Polarization maps	<input type="checkbox"/>
Light curve	<input type="checkbox"/>
3-colour image	<input type="checkbox"/>

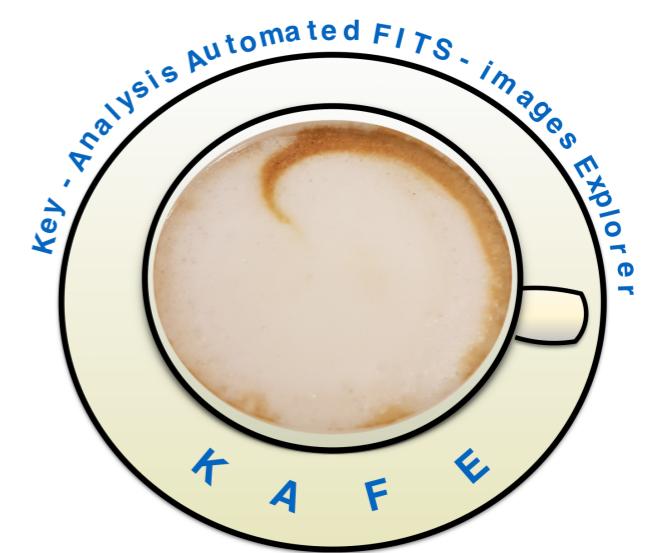
catalog selection

HDF	<input type="checkbox"/>	ATHDFSOID	<input type="checkbox"/>	HUDF	<input type="checkbox"/>	Chandra DFS	<input type="checkbox"/>
Chandra DFN	<input type="checkbox"/>	COSMOS Chandra bright src	<input type="checkbox"/>	COSMOS VLA deep	<input type="checkbox"/>	FERMILAC	<input type="checkbox"/>
FRICAT	<input type="checkbox"/>	FRIICAT	<input type="checkbox"/>	BzCAT	<input type="checkbox"/>	SPTSZSPSC	<input type="checkbox"/>

cross-match query and output specifications

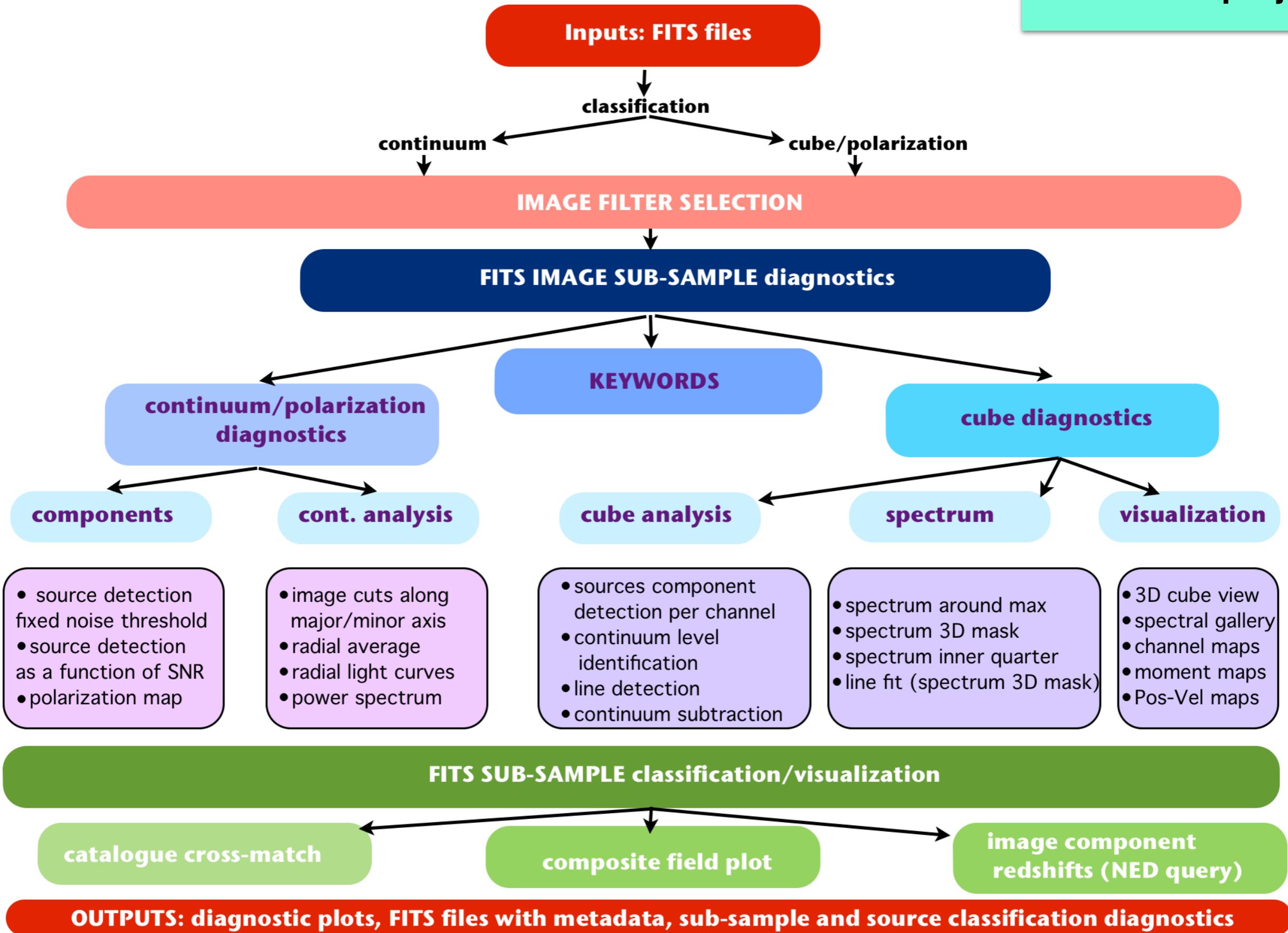
catalogue cross-match	<input type="checkbox"/>	NED redshift catalogue cross-match	<input type="checkbox"/>	NED photometry cross-match	<input type="checkbox"/>	composite field (FOV) plot	<input type="checkbox"/>
RA-redshift pie plot	<input type="checkbox"/>	Mollweide all-sky sample plot	<input type="checkbox"/>	insert KAFE keywords into FITS header	<input type="checkbox"/>	*.png *.txt *.dat products only	<input type="checkbox"/>

process



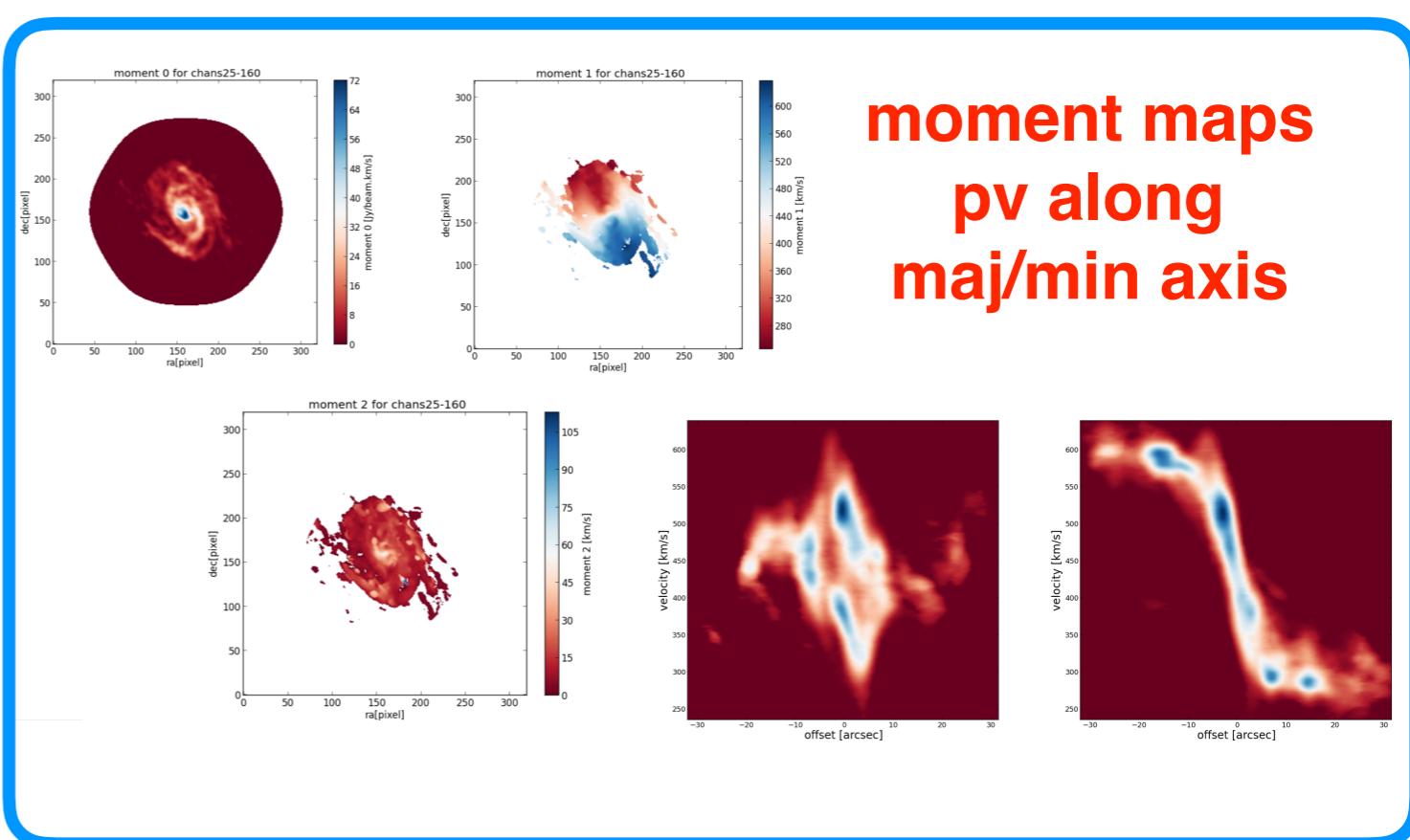
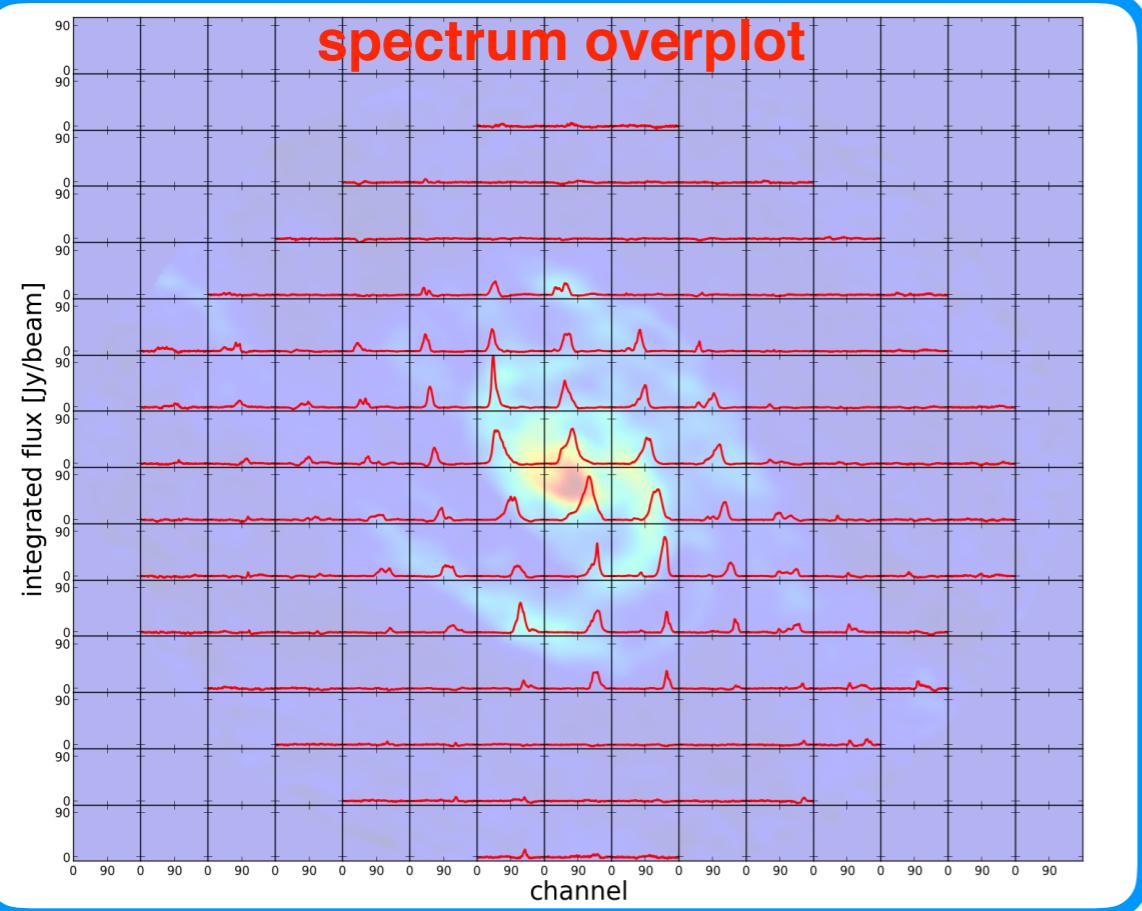
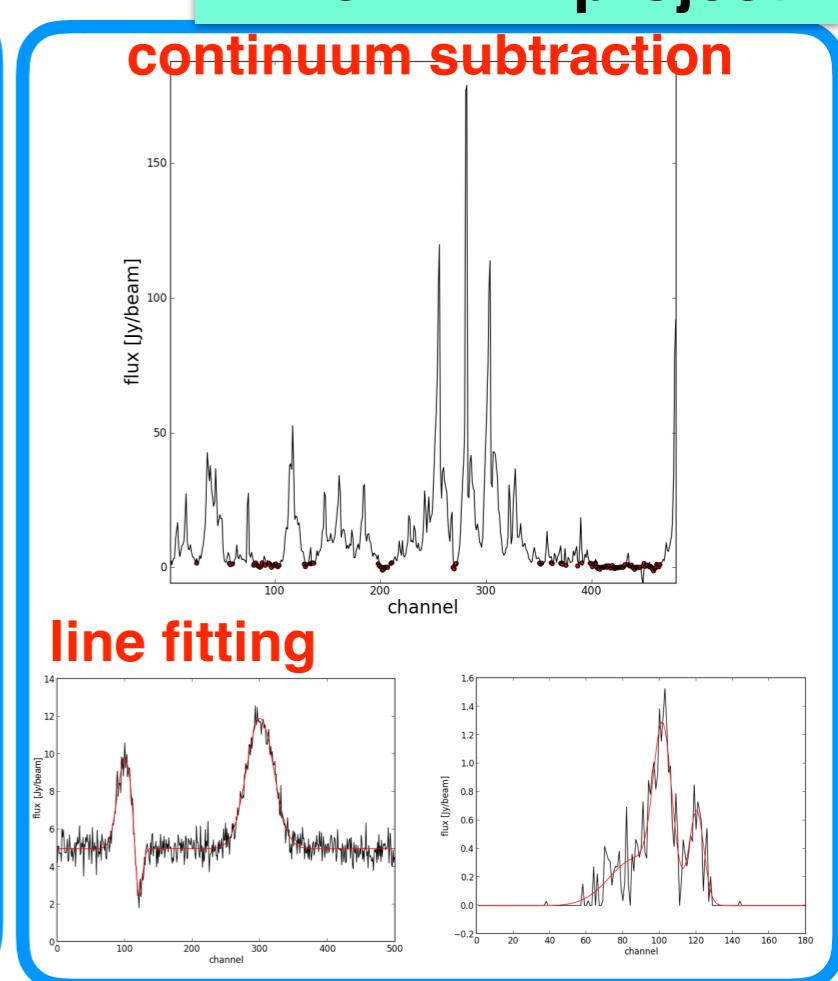
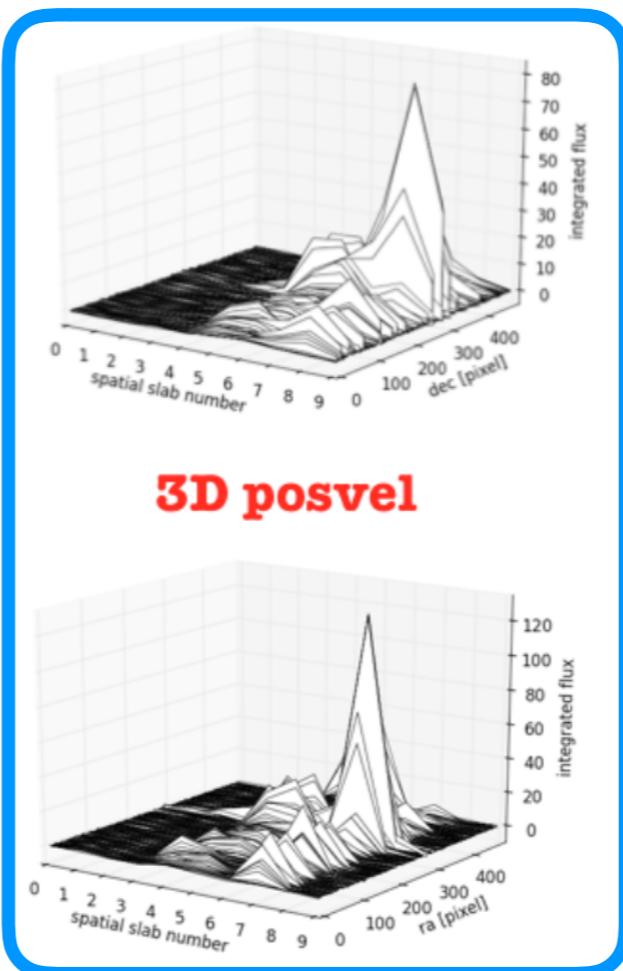
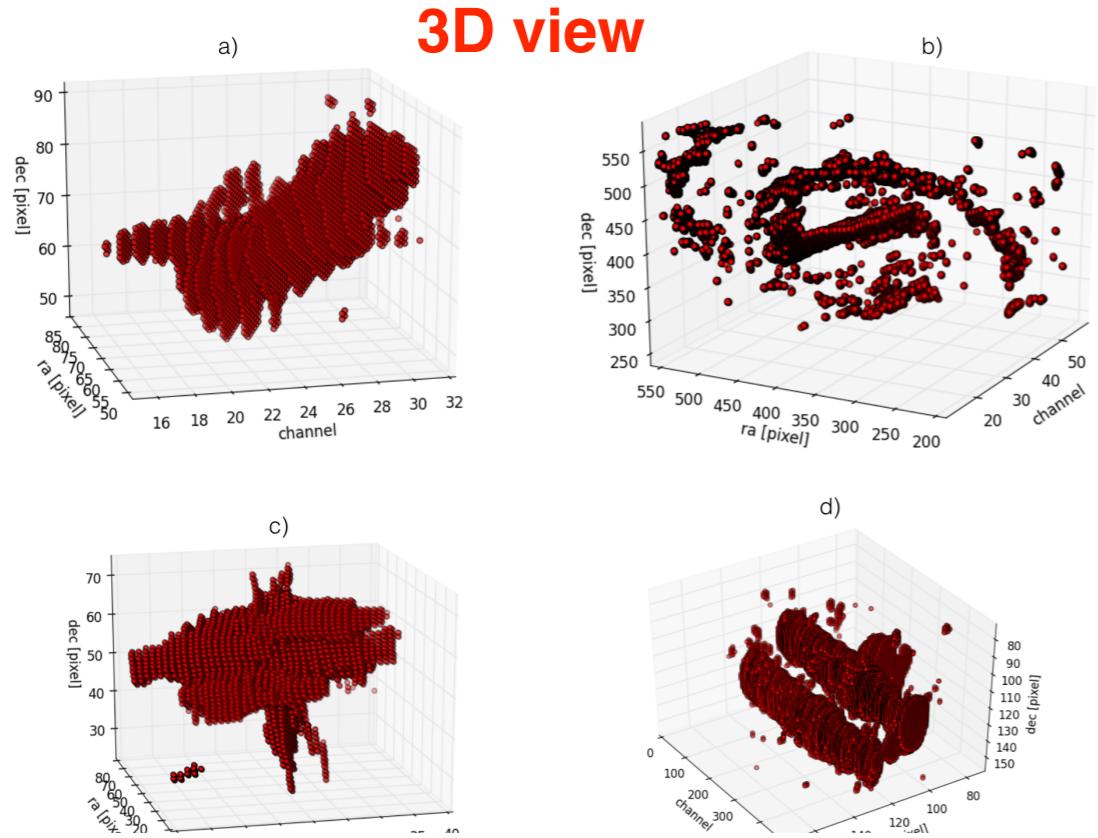
AIMS:

- provide advanced image analysis diagnostic plots in the spatial, spectral and temporal domain for user input FITS images
- provide catalogue cross-matching
- minimal user input required (just tick the boxes) - the image computations and the required parameter settings are fully automated



The KAFE project

continuum subtraction



further analysis options

- ALL (except LC,3colour)
- Source detection
- Source detection SNR layer
- radial average
- Image cuts
- power spectrum
- Polarization maps
- Light curve
- 3-colour image

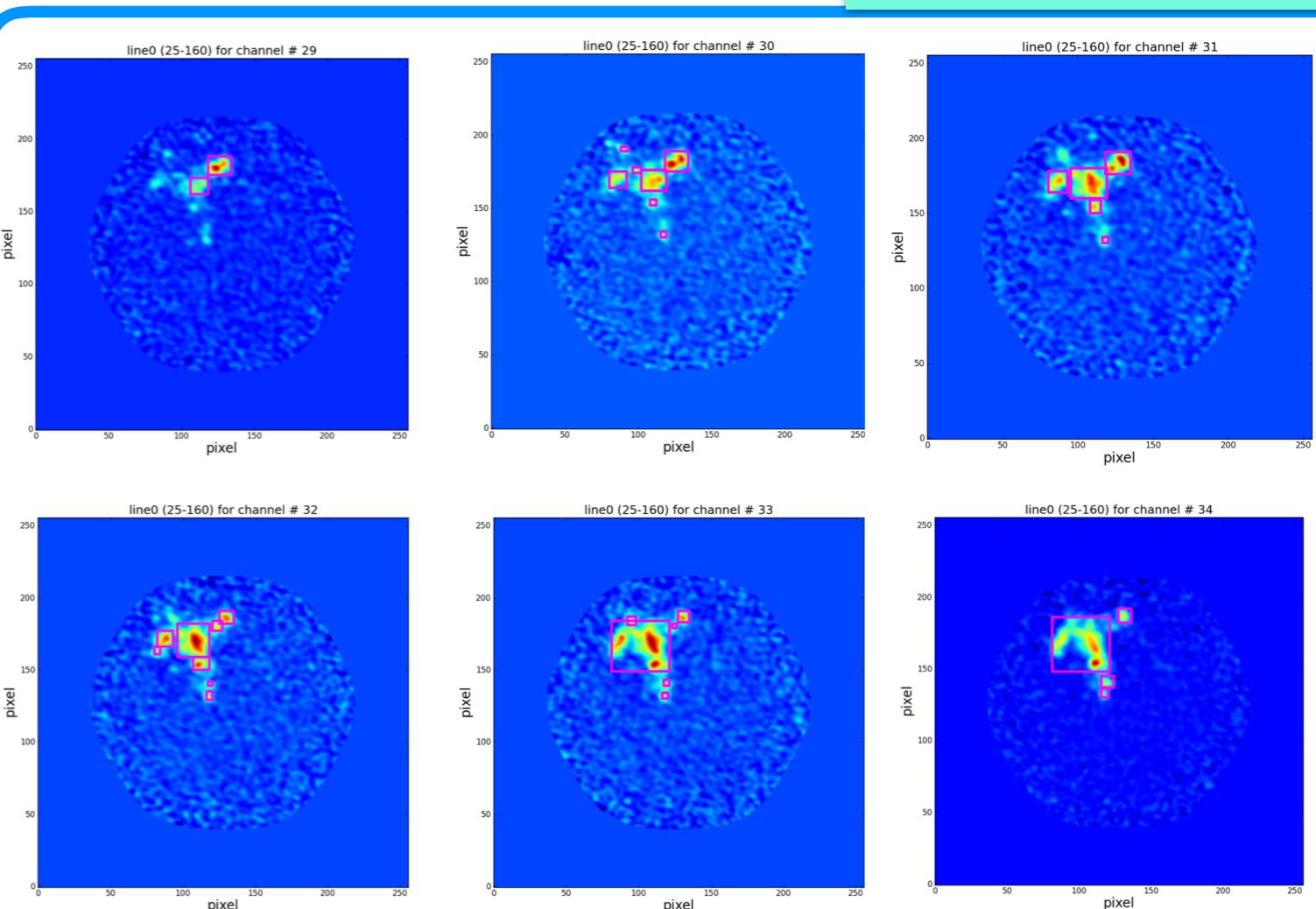
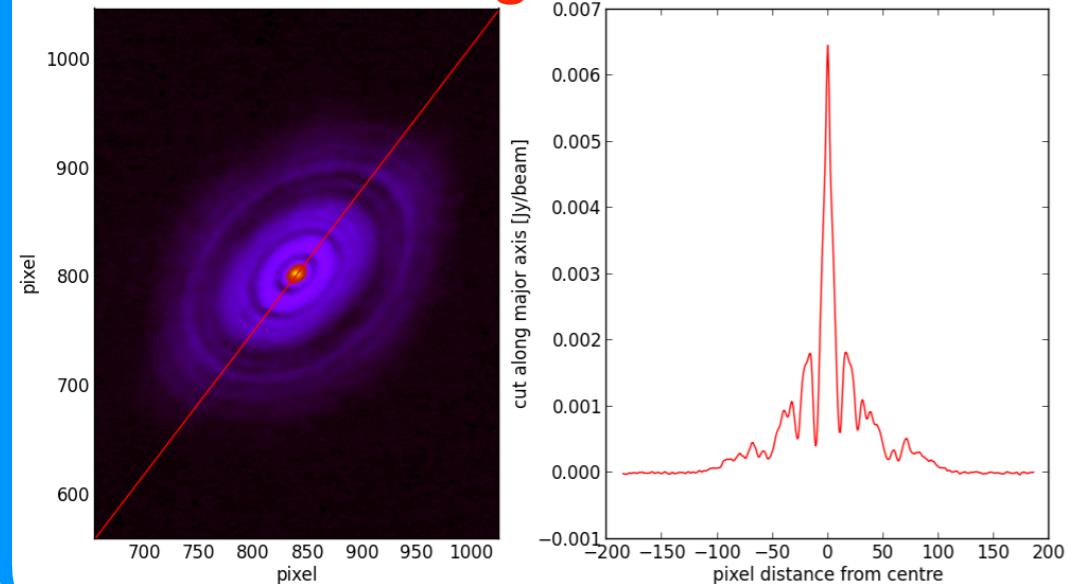
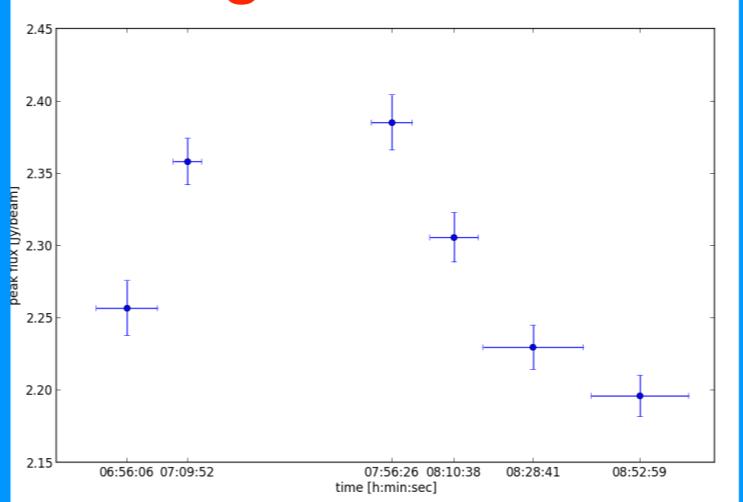


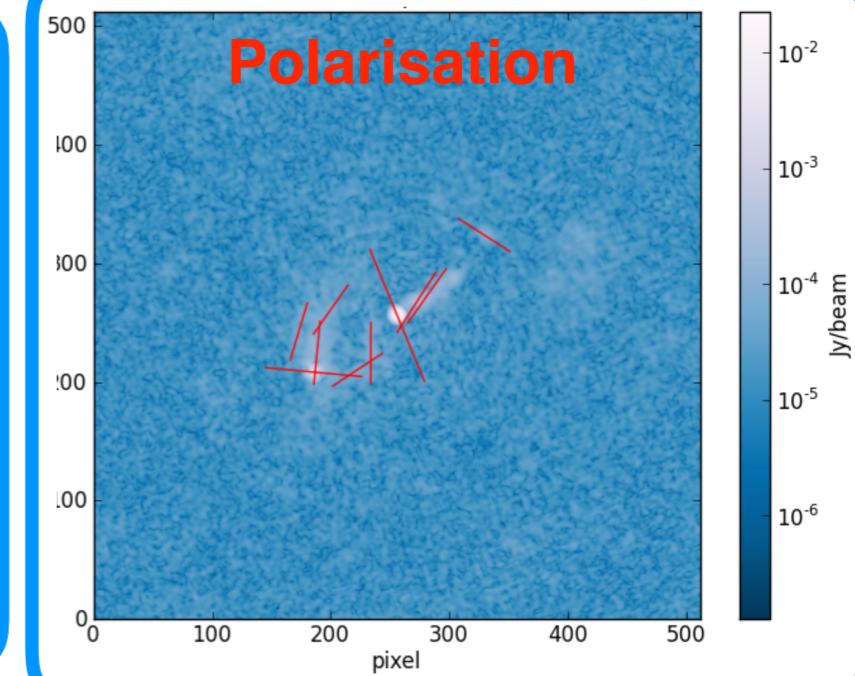
image cuts



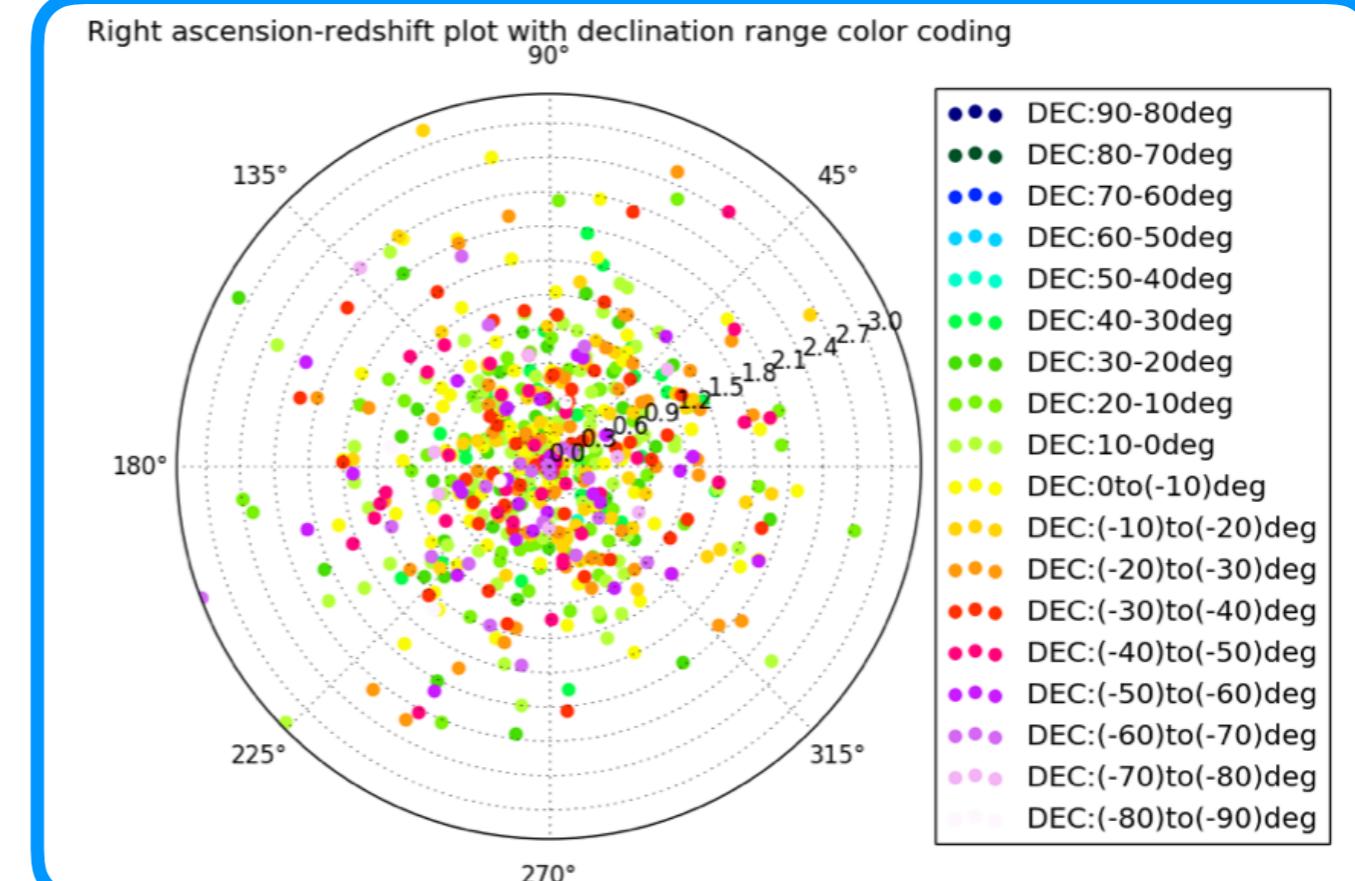
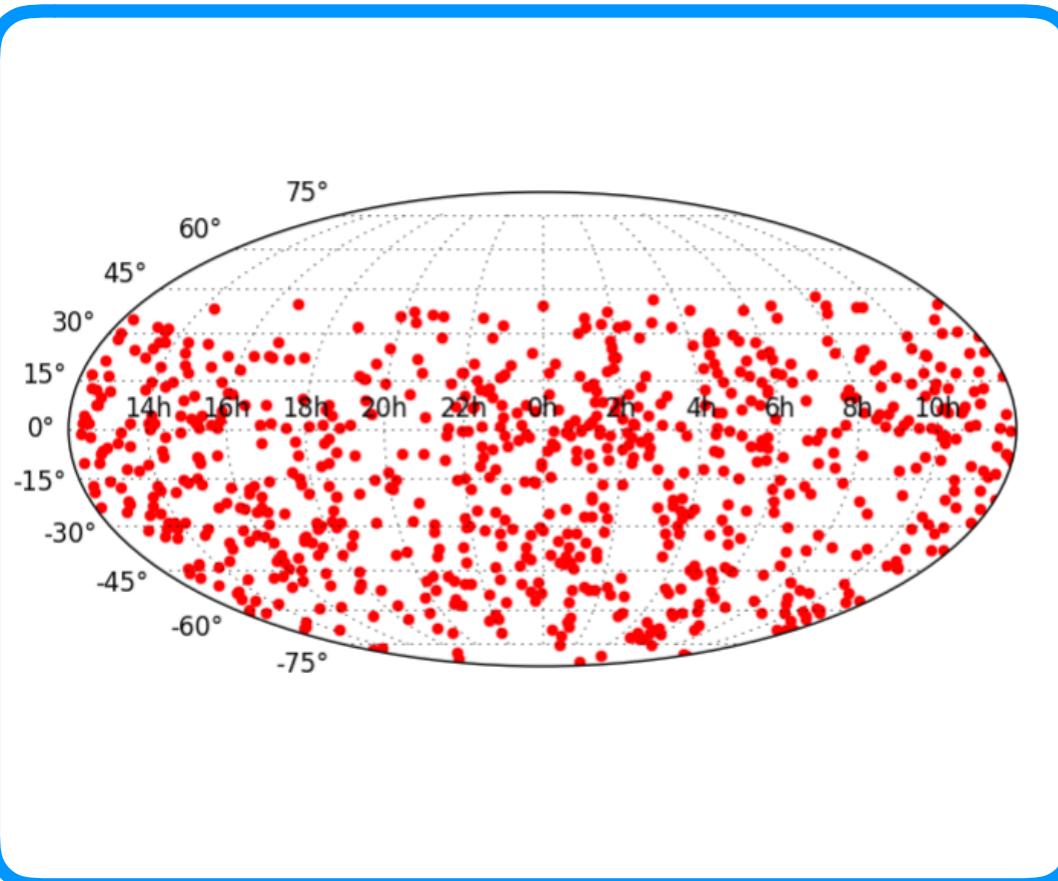
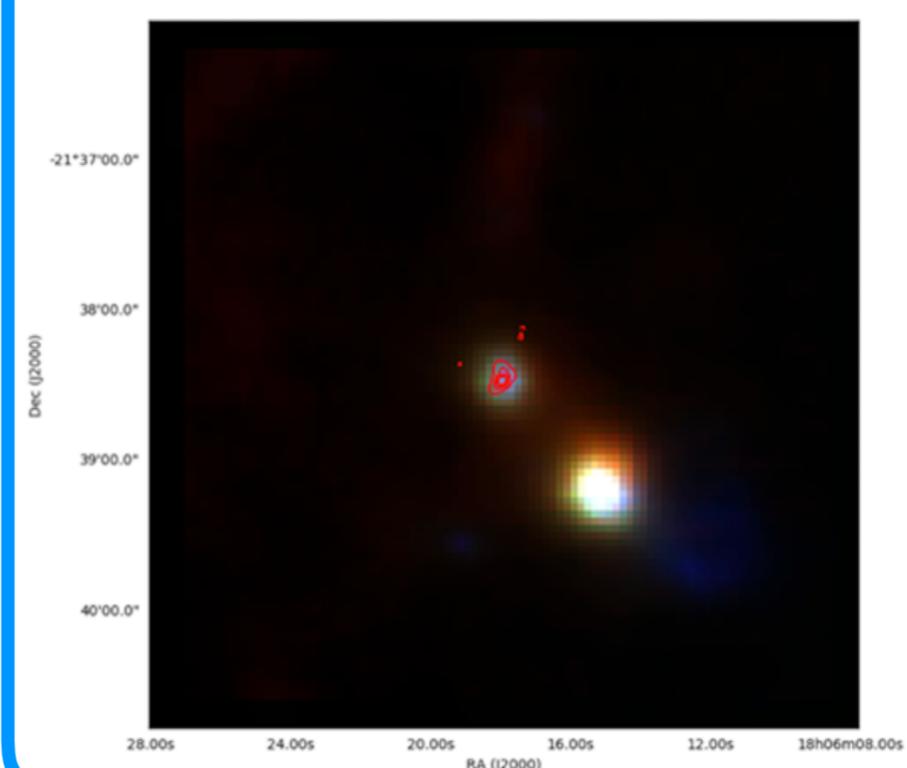
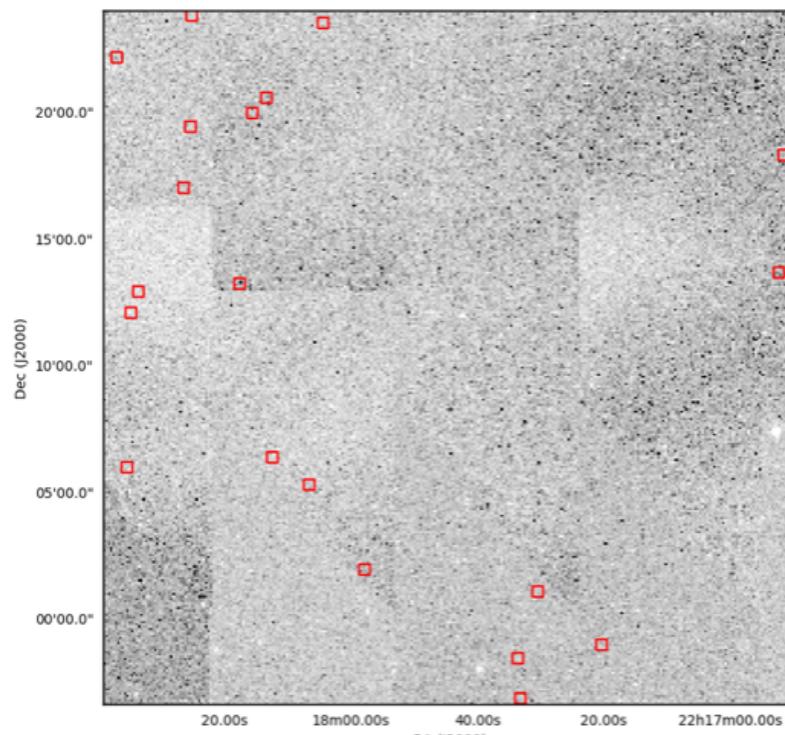
Light curves



Polarisation



The KAFE project



send this file: No file chosen

filters

POS RANGE ANGRES RANGE	0 <input type="text"/>	CNTRFREQ RANGE CHANRMS RANGE	0 <input type="text"/>	FREQRES RANGE FLUX TOTAL RANGE	0 <input type="text"/>
---------------------------	------------------------	---------------------------------	------------------------	-----------------------------------	------------------------

requested keywords

ALL	<input type="checkbox"/>
RA_centre	<input type="checkbox"/>
DEC_centre	<input type="checkbox"/>
SPATRES	<input type="checkbox"/>
BNDCTR	<input type="checkbox"/>
BNDRES	<input type="checkbox"/>
BNDWID	<input type="checkbox"/>
CHANRMS	<input type="checkbox"/>
DYNRANGE	<input type="checkbox"/>
FLUXTOT	<input type="checkbox"/>
DATAMAX	<input type="checkbox"/>
DATAMIN	<input type="checkbox"/>
STOKES	<input type="checkbox"/>

spectrum analysis options

ALL	<input type="checkbox"/>
3D view	<input type="checkbox"/>
continuum subtraction	<input type="checkbox"/>
Channel gallery	<input type="checkbox"/>
Spectrum_3D_mask	<input type="checkbox"/>
Spectrum inner quarter	<input type="checkbox"/>
Spectrum around max	<input type="checkbox"/>
Spectral gallery	<input type="checkbox"/>
3D posvel	<input type="checkbox"/>
moments	<input type="checkbox"/>
PosVel along maj/min axis	<input type="checkbox"/>
Spectral fit	<input type="checkbox"/>
Cube morph	<input type="checkbox"/>

further analysis options

ALL (except LC,3colour)	<input type="checkbox"/>
Source detection	<input type="checkbox"/>
Source detection SNR layer	<input type="checkbox"/>
radial average	<input type="checkbox"/>
Image cuts	<input type="checkbox"/>
power spectrum	<input type="checkbox"/>
Polarization maps	<input type="checkbox"/>
Light curve	<input type="checkbox"/>
3-colour image	<input type="checkbox"/>

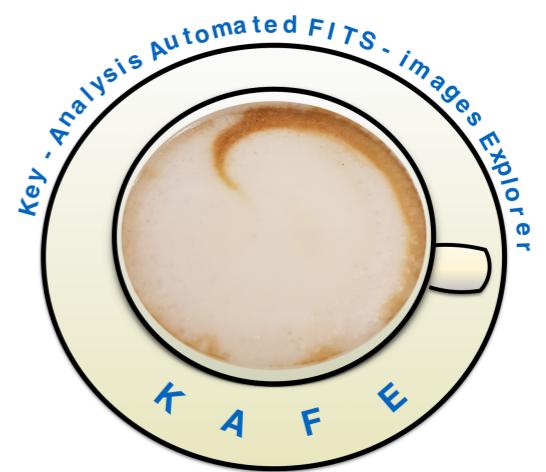
catalog selection

HDF	<input type="checkbox"/>	ATHDFSOID	<input type="checkbox"/>	HUDF	<input type="checkbox"/>	Chandra DFS	<input type="checkbox"/>
Chandra DFN	<input type="checkbox"/>	COSMOS Chandra bright src	<input type="checkbox"/>	COSMOS VLA deep	<input type="checkbox"/>	FERMILAC	<input type="checkbox"/>
FRIICAT	<input type="checkbox"/>	FRIICAT	<input type="checkbox"/>	BzCAT	<input type="checkbox"/>	SPTSZSPSC	<input type="checkbox"/>

cross-match query and output specifications

catalogue cross-match	<input type="checkbox"/>	NED redshift catalogue cross-match	<input type="checkbox"/>	NED photometry cross-match	<input type="checkbox"/>	composite field (FOV) plot	<input type="checkbox"/>
RA-redshift pie plot	<input type="checkbox"/>	Mollweide all-sky sample plot	<input type="checkbox"/>	insert KAFE keywords into FITS header	<input type="checkbox"/>	*.png *.bt *.dat products only	<input type="checkbox"/>

process



Please write to
kafe@ira.inaf.it
 for access information
 to the web interface.

TOAST: Telescope Observational Archive Sample Tool

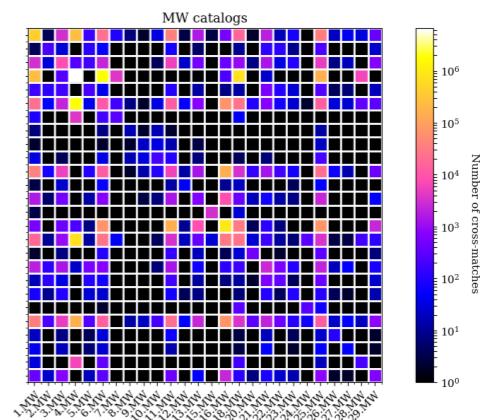
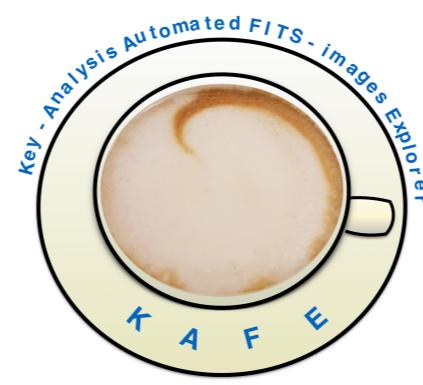
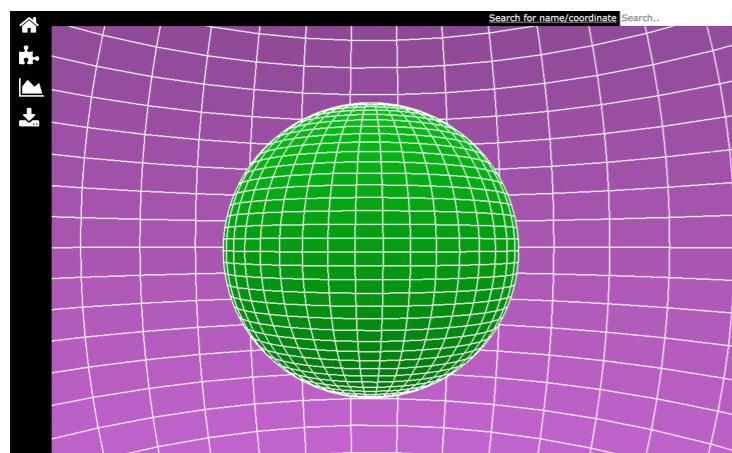
Burkutean in prep.



TOAST: Telescope Observational Archive Sample Tool

The TOAST project

Burkutean in prep.



Big DATA
visualisation

Big DATA
analysis

Big DATA
database statistics

TOAST: Telescope Observational Archive Sample Tool

The TOAST project

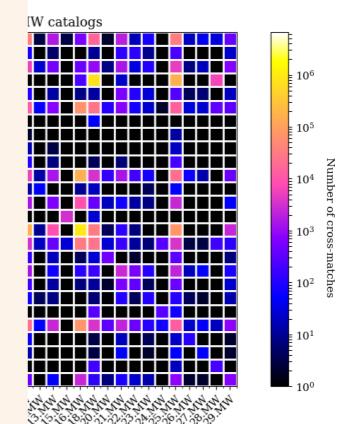
Burkutean in prep.



<https://www.freepik.com/free-photos-vectors/design>

designed by freepik.com

Big DATA
visualisation



DATA
base statistics