

THE ADDITIONAL REPRESENTATIVE IMAGES FOR LEGACY (ARI-L) DEVELOPMENT PROJECT PI: M. MASSARDI





...the ALMA Science Archive will become the primary source for an increasing number of publications. The ability to efficiently mine the archive contents is therefore vital for the community and ALMA's future...



#### Mining efficiency affects the science exploitation efficiency

## THE ALMA SCIENCE ARCHIVE STATISTICS



# THE CURRENT STATUS OF THE ALMA SCIENCE ARCHIVE FROM A MINER PERSPECTIVE



- more than 1800 project observed
  - and archived in Cycle 0-4
- data are already calibrated
- not all the possibilities have been exploited
- archive will keep growing



- induction/skill (luckily there are the ARC nodes)
- to run scripts CASA is needed (old versions ...)

- images available are made for QA2

= inhomogeneous

= not easily comparable

= incomplete (<10 % in Cy 0-4,

- archive development and new visualization tools (CARTA, VO; KAFE...) need images

## SCIENTIFIC EXPLOITATION OF THE ASA



Statistical analysis requiring many datasets are impossibly time comsuming (see the timeline of J0635-7516, calibrator In hundreds of projects)

Images in the archive are incomplete and Not representative of the data content (see the comparison of published and archived spectra for a pre-stellar object)

(Massardi et al. 2019 - ALMAmemo 614)



### THE ASA TOOLS NEED IMAGES: ASA DEVELOPMENTS AND INTERFACES



SPW 3: 328.32..330.20GHz, 976.56kHz, XX YY



### THE ASA TOOLS NEED IMAGES: JVO

#### ALMA FITS Archive : Target Info

#### Target Name : NGC5135

#### Filter by Frequency

 $\Box$  Show all the data including calibration (\*.flux.fits, \*.pb.fits, target=J####[+-]####), duplicated,  $\epsilon$ 

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	□ all	Download all the checked data	image	spect
1	ALMA01081334	2013.1.00524.S		Download WebQLv4		
2	ALMA01081333	2013.1.00524.S		Download WebQLv4	i i	
•		00404005040	_	Download		

## THE ASA TOOLS NEED IMAGES: KAFE & IMAGE ANALYSIS TOOLS



(Burkutean et al. 2018 See Burkutean's talk)



Please write to <u>kafe@ira.inaf.it</u> for access information to the web interface.



# THE ARI-L DEVELOPMENT PROJECT

The "Additional Representative Images for Legacy (ARI-L) in the ASA" project will produce and ingest into the ASA a set of additional image products representative of the whole data content for more than 70% of the observing projects in cycles 2-4 that can be processed through the ALMA Imaging Pipeline, to complement the QA2-generated images. In addition calibrated MS of the processed dataset will be released too.



- constitute a complete set of imaging products highly relevant for all science-cases and enhance the possibilities of exploitation of archival data
- facilitate the archive access and the data usage for science purposes also to non-expert data-miners,
- provide a homogeneous view of all data to compare datasets and to do a more conscious download selection,
- allow to more profitably link the archive to several tools of visualization and analysis (e.g. VO, CARTA, ADMIT, KAFE, ...),
- allow to generate previews similar to those planned for the following cycles.



## THE ARI-L DEVELOPMENT PROJECT : A FEW NUMBERS

Cy 2-3-4, excluding TP, VLBI, Solar, Polarization, QA3 for mosaicking issue

Cycle	2	3	4	Total
MOUSs	1085	1719	672	3476
Total raw data size [TB]	22	66	50	138
Median size per MOUS [GB]	11.2	16.8	20.3	14.7

**3** yrs (Kick-off Meeting on the 7<sup>th</sup> of June)

Collaboration including: INAF-IRA-It ARC, IA2, JBCA-UMAN-UK ARC, ESO



# THE ARI-L DEVELOPMENT PROJECT : WORKFLOW



The ARI-L processing will be run at ESO and at the IT-ARC site

A QA procedure will certify the products before ingesting them into the ASA.

# THE ARI-L DEVELOPMENT PROJECT : QUALITY ASSURANCE

3-levels QA:

- check that the ARI-L code run was performed correctly
- perform a weblog review of the imaging pipeline products
- extract the portion of the data corresponding to what is available as QA2 products, smooth them to the same resolutions and compare rms, fluxes and dynamic ranges

Experiment Name	Quality Summary	Compare	Notes	Discard	Accept
2013.1.01342.S_uidA001_X147_X5b	<u>Task Summary</u>	Compare FITS	Comment experiment	Blacklist	ОК
2015.1.00942.S_uidA001_X2fe_X680	Task Summary	Compare FITS	Comment experiment	Blacklist	OK
2015.1.00942.S_uidA001_X2fe_X68c	Task Summary	Compare FITS	Comment experiment	Blacklist	OK
2015.1.00942.S_uidA001_X2fe_X690	Task Summary	Compare FITS	Comment experiment	Blacklist	OK

😸 🏫 Home By Topic By Task

Project Code N/A

#### Task Summaries

Task		QA Score		Duration
1. hifa_importdata: Register measurement sets with the pipeline	1/1 have HISTORY		0.50	0:03:44
2. hif_mstransform: Create science target MS			1.00	0:05:02
3. hifa_flagtargets: ALMA Target flagging			1.00	0:00:27
4. hifa_imageprecheck: ImagePreCheck			1.00	0:13:45
5. hif_checkproductsize: Check product size			1.00	0:00:53
6. hif_makeimlist: Set-up parameters for target per-spw continuum imaging			1.00	0:00:14
7. hif_findcont: Detect continuum frequency ranges			1.00	0:25:43
8. hif_uvcontfit: UV continuum fitting			1.00	0:03:38
9. hif_uvcontsub: UV continuum subtraction			1.00	0:04:41



rms: -6.013%	dyn: 5.399%	flux: 6.223%	
flux: 1.812			flux: 1.600
min: -0.055			min: -0.054
max: 1.323			max: 1.340
dyn: 20.941			dyn: 18.796
rms: 0.063			rms: 0.071



## THE ARI-L DEVELOPMENT PROJECT : PRODUCTS IN THE ASA

The ARI-L products will not modify the current deliverables, but will be included as "externally contributed products". The project will produce a cube and a continuum image for each source (including continuum images for calibrators) for at least 70% of the 3476 MOUS in cycle 2-4.



ALMA Request Handler			Login
Anonymous User: Request #409589486	×		
Request Title: Click to edit			
Download Selected			
🖉 readme 🖉 product 🗆 raw 💭 raw (semipass) 💭 exter	nal		
Project / OUSet / Executionblock	File	Size	Accessible
🔻 🖲 🚞 Request <b>409589486</b>			
🔻 📄 🚞 Project 2016.1.00164.S			
🥑 🛅 readme	<u>2016.1.00164.S.readme.txt</u>		
🔻 📄 🚞 Science Goal OUS uid://A001/X87a/X9ee			
🔻 🖲 🚞 Group OUS uid://A001/X87a/X9ef			
🔻 📄 🚞 Member OUS uid://A001/X87a/X9f0			
► SB M83_a_03_TM1			
🗹 📄 product	2016.1.00164.S_uidA001_X87a_X9f0_001_of_001.tar	10.5GB	<b>×</b>
🗋 🕒 raw	2016.1.00164.S_uidA002_Xbc0724_X7456.asdm.sdm.tar	59.2GB	✓
📄 🕒 raw	2016.1.00164.S_uidA002_Xbc22e9_X1535.asdm.sdm.tar	49.6GB	<b></b>
🔻 📄 🚞 Member OUS uid://A001/X87a/X9f2			
SB M83_a_03_7M			
🕨 🗹 📄 product	2016.1.00164.5_uid	419.1MB	✓
🕞 💾 raw	2016.1.00164.S_uidA002_Xb91513_X6ea2.asdm.sdm.tar	1.4GB	<b>×</b>
Externally contributed products			
🕨 📄 💾 external (ARI-L) 👩	2016.1.00164.S_uid A002_Xb91513_X76eb_ariel_001_of_001.tar	10.9GB	✓
		Total: 122 5GB	

#### 7<sup>th</sup> of June 2019 – kick-off

Dec 2019 - First 100 MOUs delivered, documentation published

June 2019 – 2022 Processing



Products will be added to the archive soon after quality assessment.

A website is under construction and will report the status of the project: **STAY TUNED!!!!** 

#### The ARI-L project

Additional Representative Images for Legacy in the ALMA Science Archive

#### • THE PROJECT

The Additional Representative Images for Legacy (ARI-L) project aims to increase the legacy value of the ALMA Science Archive (ASA) by producing and ingesting into the ASA a uniform set of full data cubes and continuum images covering at least 70% of the data from Cycles 2-4.

#### THE PRODUCTS

Follow the status of the MOUS processing and the project periodic reports.

Read more

#### THE MILESTONES

Follow the periodic reports and documents

**Read more** 

#### **CONTACT US**

The ARI-L group is available for questions related to the ARI-L products.

Read more

=

**Read more** 

O