

The ALMA Science Archive: Implementation

Juande Santander-Vela, Matthias Bauhofer, Holger Meuss, Felix Stoehr, Alisdair Manning European Southern Observatory, ALMA Archive Subsystem



The Atacama Large Millimeter Array (ALMA) radio interferometer has started Early Science observations, providing a copious stream of new, high-quality astronomical datasets of unprecedented sensitivity and resolution. We present here how the ALMA Science Archive (ASA) is being implemented, leveraging existing Virtual Observatory (VO) technologies and software packages, together with Web 2.0 techniques, to provide scientists with an easy to use, multi-parameter discovery tool for ALMA data, integrated with the VO.

ASA Architecture

The ASA implementation follows the requirements spelled out by S. Etoka et al. in 2007 (1), and more recently by F. Stoehr et al., (see 2 in this poster exhibition).

The highlights of the implemented architecture (see Figure 1) are:

- a harvesting process to pass data from the ALMA Fronted Archive (AFA) to the ASA tables;
- a simple database structure (see 2 for details) supporting aggregation of data by Fields, Spectral Windows, Projects, and product Dependency for elaborated data products;
- reuse of community-maintained software, such as the openCADC stack (3), and the VOView (4), to support the Table Access Protocol (5);
- and a Spring MVC-based web graphical interface that works as VO client of the web application services;

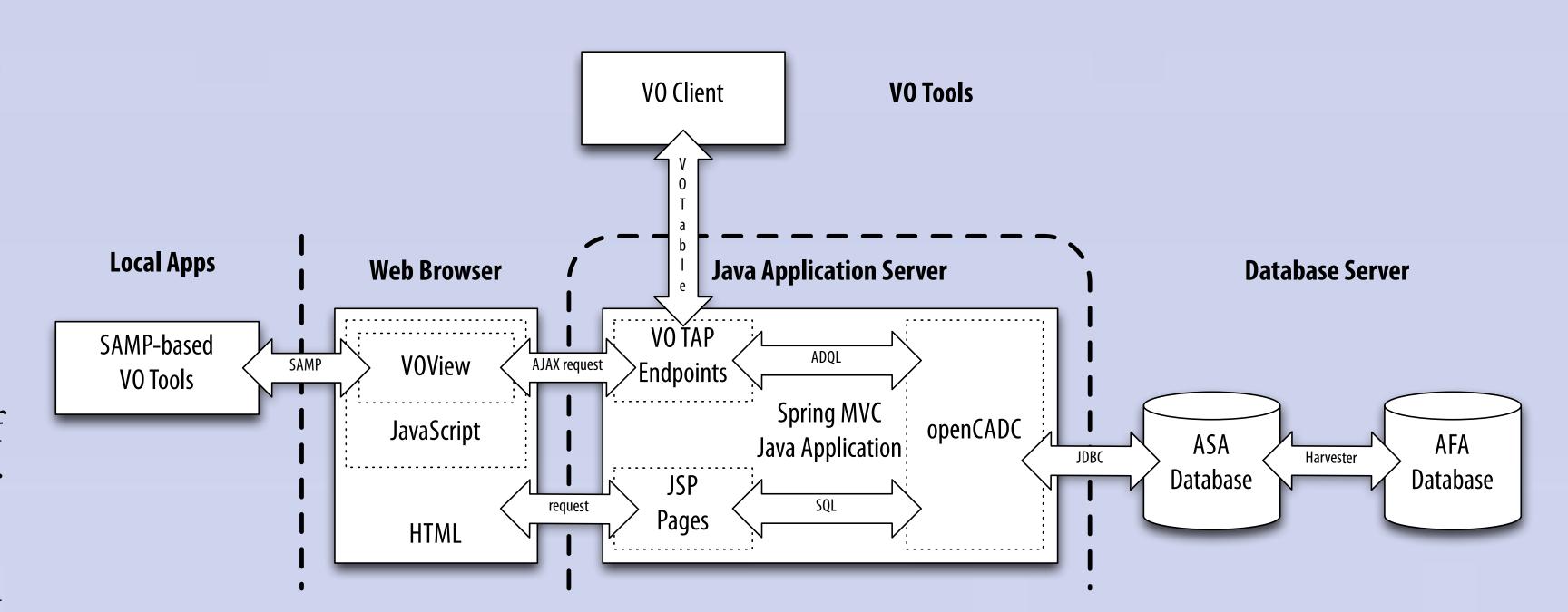


Figure 1: The layered view of our current ASA implementation shows the harvesting process feeding the ASA database from the ALMA Front-end Database, the Spring-based Java application providing GUI and endpoint services, the web-browser with embedded VOView, and different clients of the system.

Query Interface

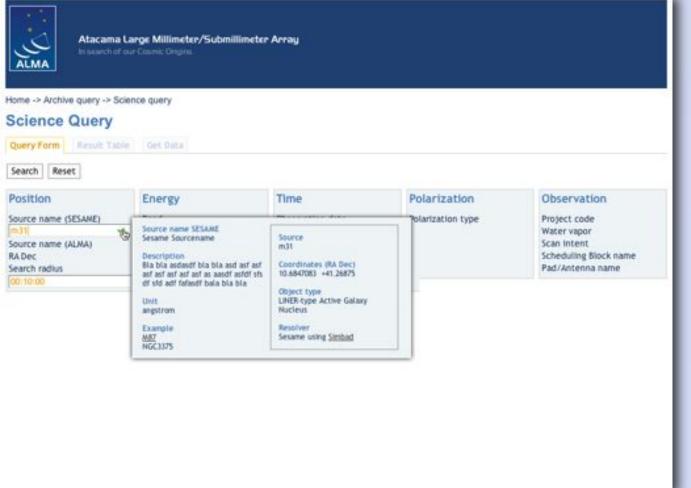


Figure 2: Above, the search interface showing IVOA characterisation-oriented axes, and dynamic name resolution via Sesame using JavaScript. To the right, the search results in VOView.

_		e Country	imeter/Submit	Salata Cara					
ALMA									
ome -> Archive query ->	- Serie	nini mar							
cience Query		13.00							
Query Form Sysuit 1	Labor								
Download data									
suits 1-10 of 21 (21 Defen	- Dilyer	ing) sorter	to sale, and		2m (H 18	results per page	47000 11100		
■ mdm_old		-	dec source.	name (64, start	prost_code		and the second of	Int. time	wi_motifies = +
(Artis)		-	Name Time	-	2010	Surrier!	Total	-	Name and Address of the Owner, where
C HETHERTONIA	styr i	NOT SAIT	mires more	os pro-co.	pollomas	198,96229407	\$1.250-40.05.2473.0(U (M. 2793.4) U (104.250000, VIII 550000)	115,241	2704.305
- servettresen			-	9773E 25-01-30 36-1-01-	Street & Street or	Ob. 90257407	[81.255148.35.247336] U (M. E19624) U (164.150000, VM. 550000)		17814-000
☐ vac-read_reader			SAME UNIVERSITY	2011-01-	Man Y manual as	U6.900F0F	(0.255+0.35.24/136) U (04.239/24) U (144.550000, 146.550000)		1701A-1015
				9/702/25/05 88	*				
- NETTHEL TERMS				2011 OF 1012 STORES		18.9037437	\$11250-918275-0038 in the TANON in Line 200000 ART 200000		17816-1975
- wat 1 wat 2 17 186 40 61 7				91102:25:01:36	Not more	18.92541	[81.150-94.85.24736] U [84.179534] U [184.50000, 96.50000]	319.406	27816,1015
Unit restaurance				the LAK SCHOL WIN	D 1010 T-000-014	116.90231437	let 122-e/tz 34,280 n (let 33424) n Gertaccco vac zosmil	311.462	27816-1015
U METHODOLOGICA	NOTE I	0:06:40.00	E00543 1005	Per 2010 200 00	Name of Street	136.95257427	\$1150-M1679-0388 in Bert 1603-6 in East 20000 Law 200008	315.563	27916.1015
C WITHOUT TRANS	11737 B	5:06:40.99	stocks for	440 DITTOLOGICAL AND	2010.3.000 EW	UA NECESTALE	[0.1255] + 0.35, 24704] + [0.17924] + [194.25000, 96.50000]	315.374	27916,1013
C HEARING TORONS	utar a	E10:46.31	\$101-02.0 EDM	M2 BYT 61-	2001.100min	138,9029,907	[83.250:44,85.24738] U [84.27924] U [184.75000,740.50000]	315.663	270%-2015
□ war ment rote each	KATE	e1821.38	morine an	2011-01-	2010 3 00010 N	tia suriter	the party of the foreign to the property of the party of	Trial Service	CORNE SONS
4000, old sylts 1-10 of 21 (21 before			dec source, to endre, and	SURVINO N	project_code		4 mm 111 mm	int time	WUMBERSON A 4
econ, and replic 1-46 of 21 (21 before columns		7.00		SUPPLIES TO SE	project_code	ref_frequency	Help No car	int_time	and the state of t
econ, and rayto 1-46 of 21 (21 before followers	r Niter	mgi sartés		SUPPLIES TO SE	project_code	ref_frequency	Help Help to car. -dis column headers to change sort are -drag column headers to recorder.	int_time	and the second
audin, old rujos 1-10 of 21 (21 better ofurmes hard column order 1-128 mon, util	New	ing) sorter	Constitution US of the ASSM 11	name obc_start	project_code	ref_frequency	Help this car: -dis column headers to change sort on -disp column headers to change sort on -disp column headers to recover -tips to filter: - see - and - for clipping it.e., -di.	101, \$100 007	and the second
exists, add replace 1-10 of 21 (21 better columns forest column order same adds, add	No.	ing) sorted char[1] dauble	Constitution US of the Acide or Right Accommon of	econing the field.	project_code	ref_frequency	Help Help to car: - circ column headers to change sort or - drag column headers to change sort or - drag column headers to recorder - type to filliar: - see - and - for cloping it.e., vii see -, to give a range it.e. 1.0.3 - see " se wildcard it.e., 2016.9 "S	not, those	and the second
endin, old replace 1-10 of 21 (21 better columns force soon, pel a leci course, name	No.	char[] dauble shar[]	Consciption US of the ACON 11 Right Accessor of the ALAN Terpit name	ectaining this Field. The field pointing. Theid pointing.	Stoke (M) (B)	ref_frequency	Help this car: -dis column headers to change sort and -drag column headers to monder -type to filler: - see - and - for cliquing it.e., -di see -, to give a range file, 1.8.3	not, those	and the second
exists, and replace 1-rip of 21 (21 betters replace order replace order replace replace order replace	No.	ing) sorter char[1] dauble dauble	Consciption URS of the ACEN II Declination of the ALAS Torget name Start time of the 5	name obc_start otalining titls field. the field pointing. field pointing.	State (SE SE	ref_frequency	Help Help to car: - circ column headers to change sort or - drag column headers to change sort or - drag column headers to recorder - type to filliar: - see - and - for cloping it.e., vii see -, to give a range it.e. 1.0.3 - see " se wildcard it.e., 2016.9 "S	not, those	and the second
exists, and replace 1-rip of 21 (21 betters columns) force and public a force force, name force, code	One of the last	Greater charge c	Decorporal US of the ASSM II Right Assemble of the ALAN Target name Start time of the 5 Project code, is to Observed guired.	ectaining this field. The field pointing, field pointing. Tell integration on the or form YYYYEARROWS, C. Interescop thequescy or	State (St.) S State (St.) S State (St.) S State (St.) SAA, where o the say	reti. Respuesto	Help this car: - cite column headers to change sort on - drag column headers to monoir - type to titler: - see - and - for clipping it.e., -(2, - see -) a wildcard it.e., 2016,6-75 - 1 to regular condition (i.e., 1000)	not, those	and the second
andin, did routs 1-10 of 21 (21 better foliumns forest column order color, util a tes ourse, name deu, start, project, code et, frequency por, windows	One of the last	charry double charry double charry double charry	Constitution United the Added to Right Accorder of Declination of the ALAN Terpet name Start time of the 5 Project code, in the Observed starteds Union of Intervals Union of Intervals Union of Intervals	ettaining titls Field. The field powers, field powers, field intergration on the is form YYYYXANOON C. Writerings the joint frequency or techning the joint frequency or	Stoke (38 - 15 Stoke (38 - 15 Stoke (38 - 15 Stoke (38 - 15 Stoke (38 - 15) Stoke (38 - 15) Stoke (38 - 15) Stoke (38 - 15)	reti. Respuesto	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
andin, did routs 1-10 of 21 (21 better foliumns foliumns form foliumns	day day	Control of the contro	Constitution URD of the ACEN II Right Accented to Author Terpit name Start time of the 5 Arquet code, is to Observed guareds Union of intervals Aggregated integr	ectaining this field. The field pointing, field pointing. Tell integration on the or form YYYYEARROWS, C. Interescop thequescy or	State 28 15 State	results per page	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
andin, did routs 1-10 of 21 (21 better foliumns foliumns form foliumns	day day	Colors of the charge of the ch	Decorporal US of the ASSM II Right Assemble of the ALAN Terpet name Start time of the 5 Project code, is to Observed duneds Union of intervals upon of intervals upon of intervals persisted winding Estimated winding Estimated winding Estimated winding	estaining this field. The field pointing, field pointing, field pointing, and interpretate on the or form VYYT shoops, C, reference thequescy of defining the joint field residuation from all the on.	State 28 15 State	results per page	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
audit, did rayits 1-10 of 21 (21 better of unteres for unteres see course order see course name se course name se plant reject pools et frequency pool production	Ong One	Control of the contro	Consciption (ACT the ACEN II Rept Accentes of the ALAS Terget name Start time of the 5 Angest code, is to Observed guards Union of intervals Estimated frequent Estimated frequent	ettaking titls Field. The field powers, field powers, field powers, field powers, a form YYYYXANOON, ordinary the paint frequency or the field frequency or sensition from all the or.	State 28 8	results per page	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
audit, did sylta 1-16 of 21 (21 better ofurmes folumes folumes folumes folumes folumes folume folumes	Ong One	Control of the charge of the c	Consequent US of the ACEA II Right Accession of the ALAM Target name Start time of the 5 Project code, is to Observed disneeds Union of intervals Union of intervals Union of intervals Estimated integral	estaining this field. The field pointing, field pointing, field pointing. In this integration on the interested frequency of the good frequency of defining the paint frequency or continue for the field president on the me. Dogs to strenge cy resolution from all interests within the painting frequency or conduction from all interests within the painting filling to code as a significant painting filling filling to code as a significant painting filling filli	State (18 18 18 18 18 18 18 18 18 18 18 18 18 1	results per page emotig all	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the state of t
action, old sylts 1-10 of 21 (21 better of unteres foliamens f	Ong One	College Charge Cha Charge Charge Charge Charge Charge Charge Charge Cha Charge	Consciption (Idl) of the ACEA of Right Accented to Destination of the ALAS Terget name Start time of the 5 Project code, is to Observed stated). Unique of statedown, Aggregated integra Estimated velocity frequency resultati Estimated frequent median values of a Marke of the Schel ACEA containing the	estaining titls Field. The facilit powering, field powering, field powering, field powering, freid powering, freid powering, freid powering, freid powering, freid powering, freid management of freid powering freid management freid management freid management freid management freid management freid, freid, gypricable to the field	State 28 8	emote per page emote per page emote all a tree	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
andio, did rojos 1-10 of 21 (21 berton ofurmos ferent column orden sociological a del course , name est, ptart project , code ef , frequency porc , anne est, ptart roject , code ef , frequency porc , anne est, ptart roject , code ef , frequency porc , anne est, ptart roject , code ef , frequency porc , anne con , anne con , anne con , anne con , anne	ong one	Collection (char[1] double char[1]	Consciption (Idl) of the ACEA of Right Accented to Destination of the ALAS Terget name Start time of the 5 Project code, is to Observed stated). Unique of statedown, Aggregated integra Estimated velocity frequency resultati Estimated frequent median values of a Marke of the Schel ACEA containing the	estaining this Field. The field pointing, field pointing, field pointing. Interence frequency of fething the joint frequency of fething the joint frequency from time for the flend recolution from all the field. Spring frequency frequen	State 28 8	emote per page emote per page emote all a tree	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
and	ong one	Colors charty ch	Consequent USE of the ASSM II Right Assemble of the ALAN Terpit name Start time of the 5 Project code, is to Observed disneds Union of astervisis, spectial winds Union of astervisis, spectial winds Union of astervisis, spectial winds Estemated brings Estemated	estaining this field. The field pointing, field the point field field pointing field field freeduction from all the field. Integrate to the field integrate to the field applicable to the field applicable to the field	State 28 8	emote per page emote per page emote all a tree	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
and to did replica 1-10 of 21 (21 betters foliumnis forest column order school and	ong one	Colleges Charge Char	Consciption (Idl) of the ACEA II Right Accented in Destination of the ALAS Terget name Start time of the 5 Project code, is th Observed guareds Union of intervals Estimated integer Estimated integer Estimated integer Estimated integer Union of the ACEA conceiver than Tallal Estimated Estimated procipts table. ALRIS receives than Tallal Estimated Estimated procipts	estaining this Field. The facilit powering, field powering, field, field, field water vigous from field field water vigous from field.	State 200 (2) State 200 (2) State 200 (2) AAA, whore on the sky wendow describes described windows to spectral windows somplists for execution the sky.	emote per page emote per page among all to trans	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
and	ong one	Colors charty ch	Consciputor URS of the ASSA to Regist Assessmen of Deviloation of the ALAN Terpot name Start time of the 5 Project code, in th Observed disserts, Lincon of intervals spectial windows, Aggregated integs Estimated veriality Integrated integs Estimated reviality Integrated integrated Disserted integrated Disserted veriality Integrated integrated Disserted integrated	estaining this Field. The facilit powering, field powering, field, field, field water vigous from field field water vigous from field.	State Still	emote per page emote per page among all to trans	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
and	ong one	College charry c	Consciputor Units of the Acides or Regist Acceptation of Destination of the ALAS Terget name Start time of the 5 Project code, to the Conserved disserted Union of intervals Union of intervals Union of intervals Union of intervals Estimated integer Estimated integer Estimated integer Estimated integer Estimated integer Estimated integer Estimated pre-class tradia. ALAS receives than Total bandwells. Estimated mumber mention value of cit Total countered in con-	ectaining this Field. The faild pointing, field pointing, field pointing. Tell interpretate on the or form YYYYANDOON, C. Interested the paint frequency of ordining the joint frequency of ordining the joint frequency or enablation from all the or. Date to therefore cy renolution from all the or. Jacobs both to the field applicable to the field of charmets from all the ordining the control of the ordining the ordining the control of the ordining the o	State 200 (2) State 200 (2) State 200 (2) AAA, whore on the sky wendown described windown to spectral windown to spectral windown to the spectral wind	emong all those one good the g	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
and	ong one	Collection (chart) deader chart)	Consciputor USS of the ASSA II Regist Assessmen of Deviloation of the ALAN Terpot name Start time of the 5 Project code, in th Observed dissends. Unique of intervals spectial integral Estimated resiscing Regionary translut Determined resiscing Regionary translut Stammated frequent meetian values of o Name of the Scheel ASSA containing of Use of scan integral Stammated procipit table. ALAN received that Total bundwells. Stammated number meetian value of ci Total number of at Stam represented in 1005-DVM for area and 2505, respects	ectaining this Field. The faild pointing, freid pointing, freid pointing. Tel interpretate on the or form YYYYANNOON C. Interestor frequency or ordining the joint frequency or or resolution from all the or. Date for therepre cy resolution from all the or. James to therepre cy resolution from all the or. of character valued as a till field. applicable to the field applicable to the field of character valued in or the feel. of character valued in or or the feel. of character valued in or the feel. from a contributing to or the feel.	State 200 (2) State 200 (2) State 200 (2) AAA, whore on the sky wendown described windown to spectral windown to spectral windown to the spectral wind	emong all those one good the g	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
andis, did replica 1-10 of 21 (21 better distrates fener column order second, parte des, prosecution fener, production free, produc	One o	Colors Charty	Consciputor Utili of the Acide (s) Right Accenter of Declination of the ALAN Terget name Start time of the 5 Project code, in the Chaerwell statement Union of intervals Estimated smitchs Estimated smitchs Estimated smitchs Estimated smitchs Estimated smitchs Estimated proclipt trabia. ALIAN receiver ban Total bundwists. Estimated proclipt trabia. ALIAN receiver ban Total bundwists. Estimated member median value of ci Total number and and 2505, respectit for batter produc and 2505, respectit for batter produc stamming special in	estaining this Field. The field pointing, field pointing. Interesce the question on the field orderence the question of the production from all the field resolution from all the field. Interescent the field resolution from all the field. Interescent widther, in the field estate videous to the field estate videous from all the field estate videous from all the field. In of charmonic from all the week manufacture or the field estate videous from all the second contributing to provide a contributing to provide a contributing to provide a contribution of the field estate the fi	State 200 (2) State 200 (2) State 200 (2) AAA, where in the sky windows to spectral wi	emong all smoong all smoong the scountries scountries scountries scountries scountries scountries	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second
eroin, ald eroins 1-16 of 21 (21 before Columns News column order	one one one one	Colors Charty	Consciption (Idl) of the ACEA to Right Accented to Destination of the ALAS Terget name Start time of the 5 Project code, is to Observed guareds Liver of steered Colorination intervision (pectical introduces, Aggregated integra Estimated procipit rable. ALAS conceived that Tatal Survivolts, Estimated procipit rable. ALAS conceived that Estimated guardened in 1505-DVSh for ante and 2505, respecto Marination product wiscemum special or from minimum special or from minimum special or from minimum special or from minimum special or	estaining this Field. The faild pointing, field pointing, field pointing. It interpretate on the control of they stan time for the field residution from all the on. Dags for sharper or resolution from all of harmal station in the on. Jones for sharper or resolution from all of harmal station or on field. applicable to the field applicable to the field of chemical sharper of field. Income controlluting to or of field ones and or of field ones and ones o	State Still	emote per page emote per page emote di total tot	Help Tax car: -discontent headers to change sort or -disc column headers to change sort or -disc column headers to monder -type to filter: - see - and - for change is a181 - see - and - for change is a181 - see ' as wildcard is a., 2016.5.75 - I to negate condition (i.e., 1970) Tablie Parameters	not, those	and the second

Figure 2 shows two tabs of the search interface. The initial tab (left) supports searching on Position, Energy, Time, Polarisation, and Observation provenance parameters, and uses JavaScript and AJAX to allow for in-page searching, and automatic coordinates' resolution for object names. The result tab shows data found in the VOView, and thus allows for interactive filtering, searching, column reordering, hiding, and sorting. Search operators such as ! (NOT), * (wildcards) and . . (numeric ranges) are supported.

We are enhancing the VOView so that it can interact with SAMP-based applications (see 6), allowing interactivity between local desktop applications and our query interface.

Harvesting

Until the ALMA pipeline heuristics are fully defined, only the raw data coming from observations will be available. ASDMs (Viallefond, see 7) are *harvested* from the ALMA Front-end Archive (AFA), where they were created and stored by the ALMA Correlator subsystem. For details on how the data reaches the AFA, see 8.

The most delicate part of the algorithm (see Figure 3), once non-observational data have been rejected, is the *folding*, or dimensionality reduction, of the data. During folding, data from all other ASDM tables are compiled at the Field level.

In particular:

- integration time from Scans is attributed to observed Fields;
- weather and water vapour information is averaged;
- spectral information is aggregated per Field via interval operations, computing characteristic channel numbers, bandwidth, etc.;
- Antennas and Station pairs are created, baseline vectors calculated, and then reprojected to estimate spatial resolution and field of view;
- and other quantities are estimated.

In total, more than 70 parameters are calculated per Field, and available to be both searched and retrieved.

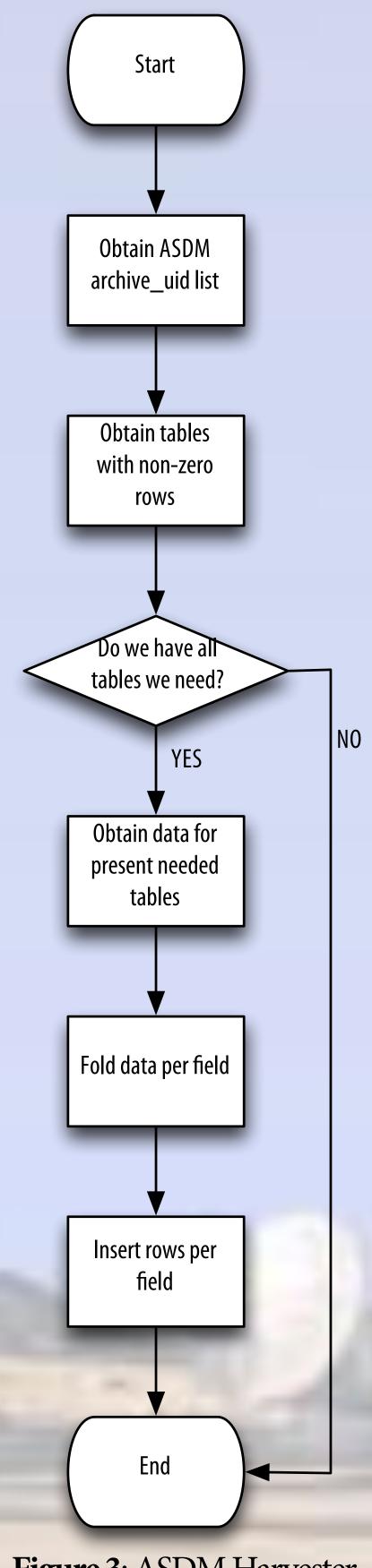


Figure 3: ASDM Harvester

References

- 1. S. Etoka, G. Fuller, and A. Wicenec, "ALMA Science Archive Requirements," Draft COMP-70.50.00.00- 004-A-GEN, Atacama Large Millimeter Array, April 2007.
- 2. F. Stoehr, et al., "The ALMA Science Archive: Design", ADASS XXI, P147
- 3. Patrick Dowler et al., openCADC: http://code.google.com/p/opencadc/
- 4. Dean Hinshaw, VOView: http://code.google.com/p/voview/

- 5. P. Dowler, G. Rixon, and D. Tody, "Table Access Protocol," IVOA Recommendation 1.0, March 2010.
- 6. M. B. Taylor, T. Boch, M. Fitzpatrick, A. Allan, L. Paioro, J. Taylor, and D. Tody, "Simple Application Messaging Protocol," IVOA Recommendation 1.11, April 2009.
- 7. F. Viallefond, "The Alma Science Data Model," ADASS XV, vol. 351, ASP Conference Series, 2006
- 8. A. Manning, et al., "ALMA Archive: Data to the Scientists", ADASS XXI, **P089**