



The ALMA Science Archive

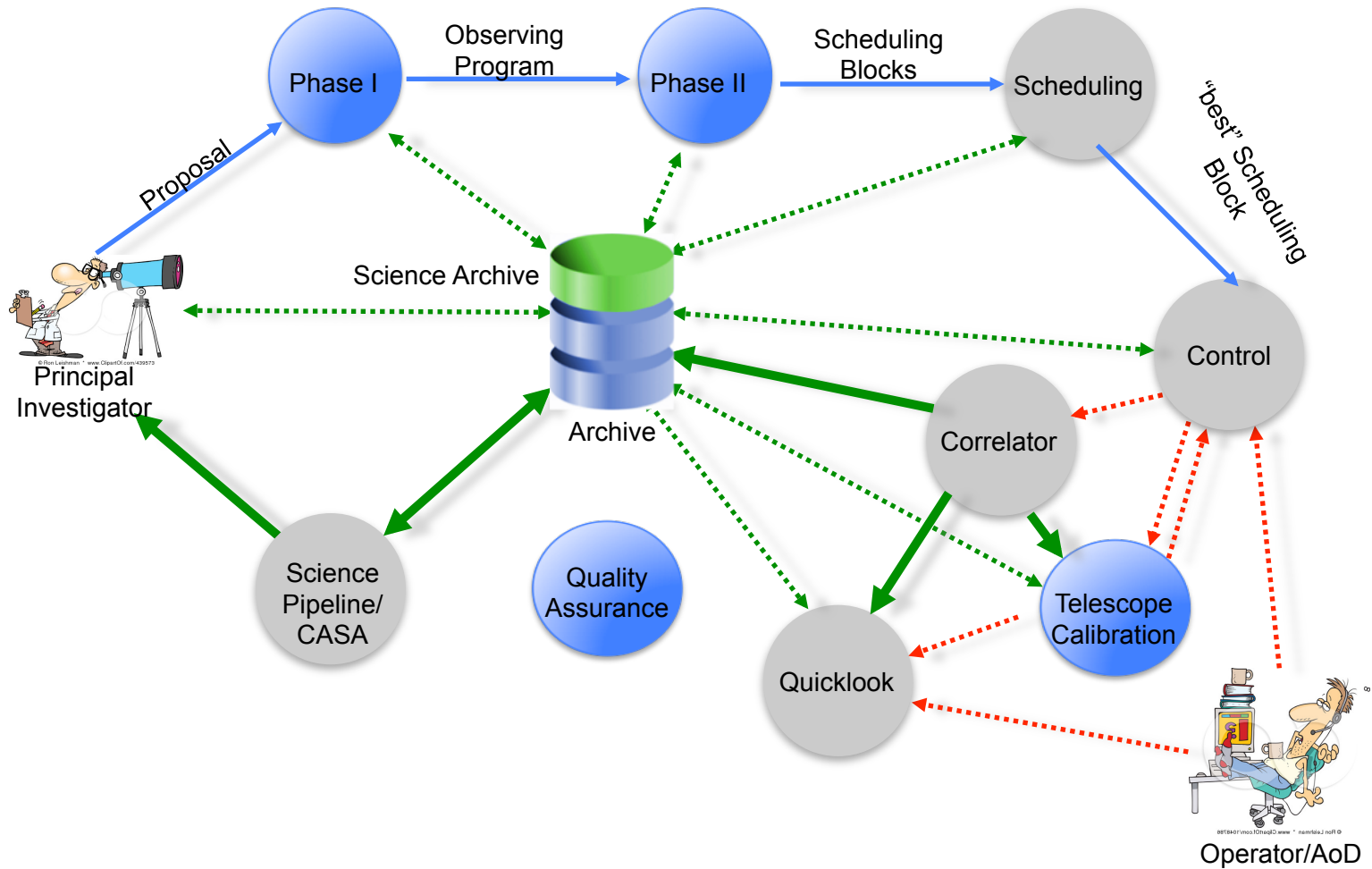
Felix Stoehr
Subsystem Scientist



ALMA

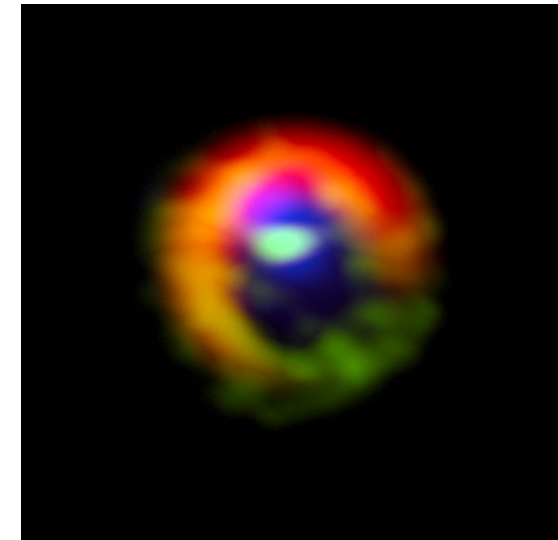
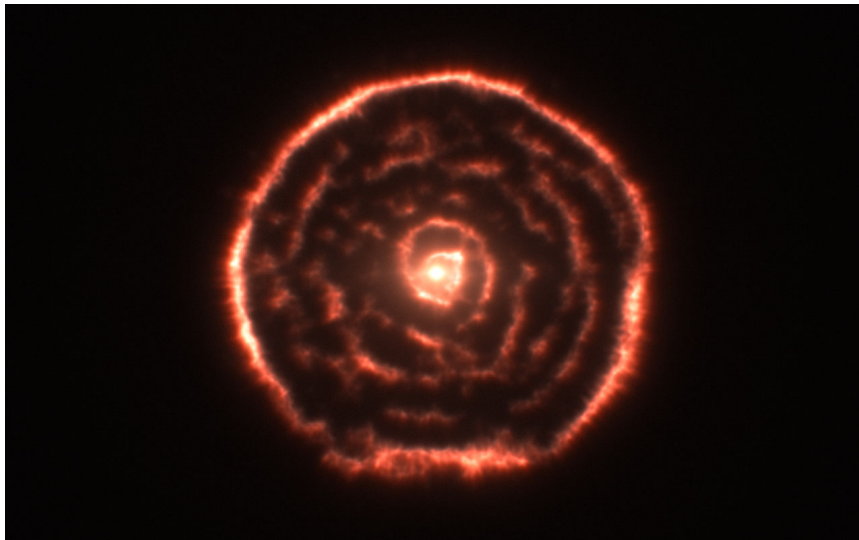
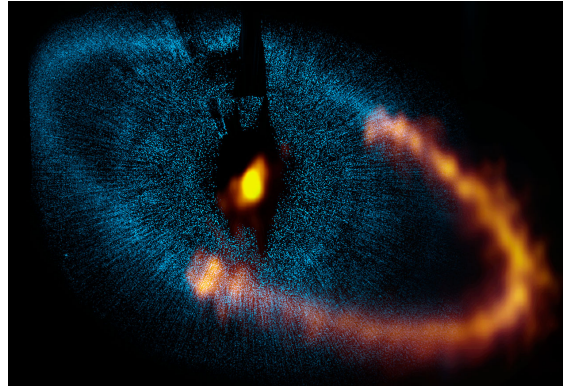
- ◆ 66 antennas at 5000m elevation in the Atacama Desert
 - ◆ Interferometry with baselines up to 16km
 - ◆ Wavelength range from 3mm to 400 μ m (84 to 702 GHz)
 - ◆ Average data rate in full operations: 200TB/yr=6.6Mbytes/s
 - ◆ Built by ESO, NRAO, NAOJ in cooperation with Chile
-
- ◆ Credits: the whole ALMA Archive team
 - ◆ Acknowledgement: long discussions with the (former) HST teams at CADC and ESO

The ALMA Work Flow



◆ Slide courtesy Erich Schmidt

ALMA Science



Credits: ALMA (ESO/NAOJ/NRAO). Visible light image: the NASA/ESA Hubble Space Telescope, ALMA (ESO/NAOJ/NRAO)/M. Maercker et al., A.C. Boley et al., ALMA (ESO/NAOJ/NRAO), T.A. Rector (University of Alaska Anchorage).
Visible-light image: ESO

Science Archive Rationale

- ◆ As the success of ALMA is measured by **the science output of the community**, the goal is to render the user-experience for PIs and archival researchers as perfect as possible
- ◆ Creating a Science Archive that allows archival researchers to discover and retrieve the data they can use, helps **maximise the scientific return** of ALMA
- ◆ Great **return-for-investment** ratio

Science Archive Design

- ◆ speak the **language of the scientists** (query by **physical concepts**)
- ◆ provide only **relevant** information, keep hurdles **small**
- ◆ help users to **quickly decide** whether or not the data returned in a given search is relevant for them

- ◆ **complete, correct, consistent**, homogenized
- ◆ **fast** responses
- ◆ **programmatic** access to metadata and data

- ◆ one small set of **optimized database tables**
- ◆ use **existing standards** and **technologies**
- ◆ **iterative** development and evolution

Science Archive Policies

- ◆ All public data and metadata is accessible **anonymously**
- ◆ There will be **no private metadata** in the ASA
- ◆ PIs are normal archive users with more access rights
- ◆ Authors are **required** to acknowledge the use of ALMA data using a standard data tag
- ◆ Metadata will be available when the **first data has been ingested** into the archive. Titles, PI/Col names, abstracts, project codes directly after Science Assessment.



Current version (almascience.org/aq)



Atacama Large Millimeter/Submillimeter Array
In search of our Cosmic Origins

You are here: [Home](#) > [ALMA Data](#) > Archive Query

ALMA Science Archive Query

Query Form Result Table

Search Reset

[Query Help](#)

Position Source name (Sesame) <input type="text" value="R scl"/> ✓ Source name (ALMA) RA Dec Search radius <input type="text" value="0:10:00"/>	Energy Source name (Sesame) Source name to be resolved with Sesame. Description We use the external resolver Sesame to obtain coordinates for the source. Sesame queries the Simbad, NED and VizieR services. Example M87 NGC3375	Time Source V* R Scl Coordinates (RA Dec) 01:26:58.09 -32:32:35.4 Object type sr* (Semi-regular pulsating Star) Resolver Sesame using Simbad	Polarisation Polarisation type
Observation Water vapour Scan intent <input type="text" value="Observe target"/>			Options Results view <input checked="" type="radio"/> raw data <input type="radio"/> project Release status <input checked="" type="radio"/> public data <input type="radio"/> all data

- ◆ Still very limited
- ◆ Only raw data searches possible



Current version



Atacama Large Millimeter/Submillimeter Array
In search of our Cosmic Origins

You are here: [Home](#) > [ALMA Data](#) > Archive Query

ALMA Science Archive Query

Query Form

Result Table

Submit download request

Showing 5541 rows (5541 before filtering)

<input type="checkbox"/>	project_code	SOURCE_NAME
Filter:	<input type="text"/>	<input type="text"/>
<input type="checkbox"/>	2011.0.00101.S	B0007+106; J0010+109
<input type="checkbox"/>	2011.0.00101.S	B0007+106; J0010+109
<input type="checkbox"/>	2011.0.00101.S	Callisto
<input type="checkbox"/>	2011.0.00101.S	Callisto
<input checked="" type="checkbox"/>	2011.0.00101.S	GRB021004
<input checked="" type="checkbox"/>	2011.0.00101.S	GRB021004
<input checked="" type="checkbox"/>	2011.0.00101.S	J2253+161; 3c454.3
<input type="checkbox"/>	2011.0.00101.S	J2253+161; 3c454.3
<input type="checkbox"/>	2011.0.00131.S	3c454.3
<input type="checkbox"/>	2011.0.00131.S	3c454.3

Add/remove displayed columns

Drag & drop columns above or below the red bar, move the red bar itself or click on the checkboxes.

Reorder columns

Drag & drop the columns or drag & drop the column headers directly in the results table.

Show all columns | Reset column order | Order alphabetically

<input checked="" type="checkbox"/>	project_code	Project code, in the form YYYY.NNNNN.C.AAA, where:
<input checked="" type="checkbox"/>	SOURCE_NAME	Name of the source as registered in the ASDM. Partial matches through wildcards (? , *), and boolean OR expressions (" "), can be used.
<input checked="" type="checkbox"/>	RA	Right Ascension of the field pointing. (Default unit 'deg')
<input checked="" type="checkbox"/>	DEC	Declination of the field pointing. (Default unit 'deg')
<input checked="" type="checkbox"/>	BAND	ALMA receiver band.
<input checked="" type="checkbox"/>	integration	Aggregated integration time for the field in the ASDM. (Default unit 's')
<input checked="" type="checkbox"/>	RELEASE_DATE	RELEASE_DATE
<input checked="" type="checkbox"/>	vel_resolution	Estimated velocity resolution from all the spectral windows, from frequency resolution. (Default unit 'm/s')
<hr style="border: 2px solid red;"/>		
<input type="checkbox"/>	freq_resolution	Estimated frequency resolution from all the spectral windows, using median values of channel widths. (Default unit 'kHz')
<input type="checkbox"/>	POL_PRODUCTS	Polarisation products provided.
<input type="checkbox"/>	start_date	to_char(asa_science.start_date, 'YYYY-MM-DD HH24:MI:SS')
<input type="checkbox"/>	PI_NAME	case-insensitive partial match over the full PI name. Wildcards can be used
<input type="checkbox"/>	PWV	Estimated precipitable water vapour from the XML_CALWVR_ENTITIES table. (Default unit 'mm')

More columns

resolution

58.00

58.00

58.00

58.00

58.00

58.00

58.00

58.00

58.00


76

76

- ◆ Similar to HST interface (now at ESA!)
- ◆ Code reuse from CADC (VO technology)



Current version



Atacama Large Millimeter
In search of our Cosmic Origins

Request Handler

Archive Requests Req #36,504,981

Request #36504981 by Anonymous
[ALMA](#)

Requested Projects / OUSets / Execution

UNIX/Linux Shell script if you prefer command li

Select	Project / OUSet / Executio
<input checked="" type="checkbox"/>	<input type="checkbox"/> uid__A002_X2cbea7_X7
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	

Atacama Large Millimeter/Submillimeter Array
In search of our Cosmic Origins

Filename	Status	Progress
2011.0.00131.S_2011-12-06_002_of_028.tar	Downloading	0%
2011.0.00131.S_2011-12-06_003_of_028.tar	Completed	100% - 361.4MB of 361.4MB, 17.6GB/s
2011.0.00131.S_2011-12-06_004_of_028.tar	Downloading	0%
2011.0.00131.S_2011-12-06_005_of_028.tar	Completed	100% - 420.0MB of 420.0MB, 21.2MB/s
2011.0.00131.S_2011-12-06_006_of_028.tar	Downloading	6% - 525.7MB of 8.6GB
2011.0.00131.S_2011-12-06_007_of_028.tar	Downloading	91% - 264.3MB of 290.1MB
2011.0.00131.S_2011-12-06_008_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_009_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_010_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_011_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_012_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_013_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_014_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_015_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_016_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_017_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_018_of_028.tar	Queued	0%
2011.0.00131.S_2011-12-06_019_of_028.tar	Queued	0%

14:25:23 2011.0.00131.S_2011-12-06_007_of_028.tar Queued ----> 14:25:55 Downloading

14:25:35 2011.0.00131.S_2011-12-06_005_of_028.tar Downloading ----> 14:25:55 Completed 420.0MB 21.2MB/s

Speed: 37.8MB/s | Completed 3 of 28 files, failed 0 | 4.6GB of 108.1GB Conc. Downloads 1.5 4

[2011.0.00131.S_2011-12-06_009_of_028.tar](#)

Login

Select All Download Selected

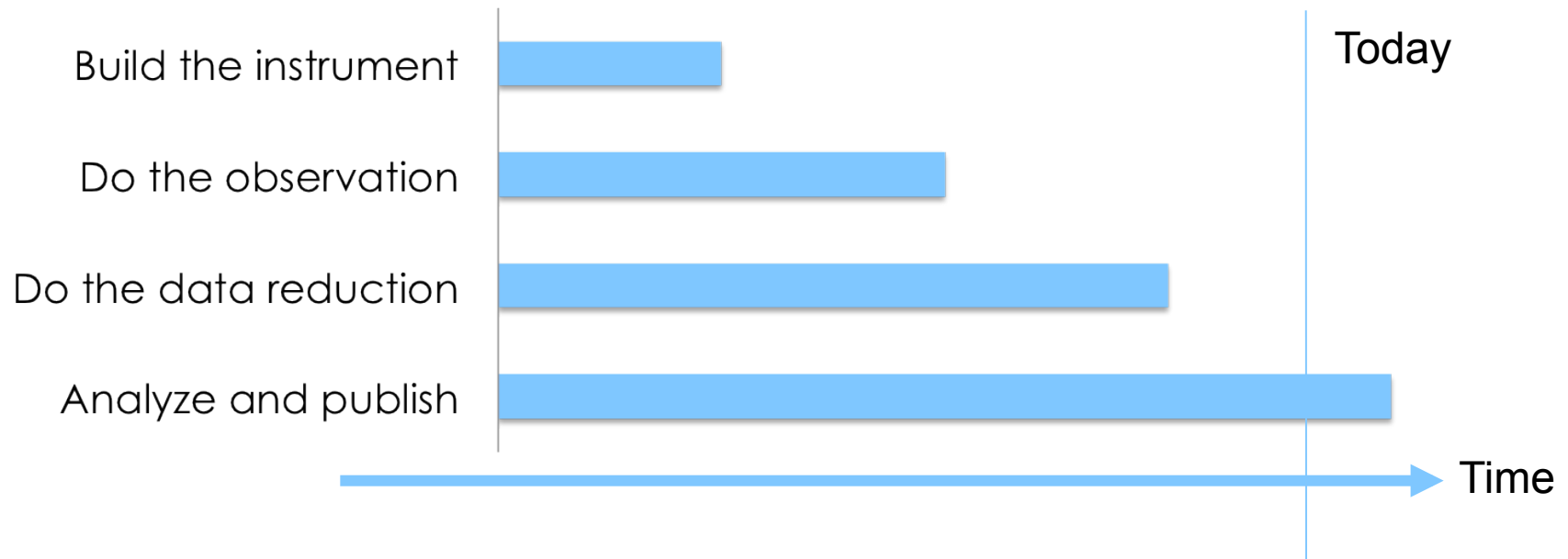
Size	Access
60.0KB	✓
8.4GB	✓
361.4MB	✓
6.8GB	✓
420.0MB	✓
8.6GB	✓
290.1MB	✓
7.4GB	✓
496.6MB	✓

◆ ESO's SAF and ALMA use the same Request Handler



Evolution: Division of Labour

◆ What do PIs do?



- ◆ As telescopes will become more complex and deliver more data, telescopes will move to deliver **science-grade** products
- ◆ The importance of the ALMA **pipeline** for the success of ALMA can not be overemphasized

Evolution of Astronomy

- ◆ ALMA will produce about the same amount of data in one year as ESO has produced in its first 50 years
- ◆ LSST, LOFAR, SKA, PanStars, Euclid, Gaia, ELTs
- ◆ T. Tyson: Astronomy is **transformed** from being a data-starved science to one where data is overabundant
- ◆ Multi-wavelength science: **less time** per wavelength regime
- ◆ **Astronomers do not scale**: bytes/astronomer grow exponentially
- ◆ Now: astronomers compete for observing time
Future: observatories will **compete for astronomers** to work with their data

Conclusion

- ◆ The goal of ALMA is to help the scientists wherever possible from proposal preparation over the **science-grade data products** and data-reduction to archival research and to work towards a great end-to-end **user-experience** (by “**turning around**” and looking at the whole workflow from a user’s point of view)
- ◆ Part of this effort is to construct a powerful Science Archive which helps **maximising the scientific return** of the observatory