ALMA Data –
what to expect after your observations are made

Presenter: Sarah Wood

Authors: Sarah Wood, Nathan Brunetti

NRAO / NA ALMA Science Center

Atacama Large Millimeter/submillimeter Array
Expanded Very Large Array
Robert C. Byrd Green Bank Telescope
Very Long Baseline Array

The Condensed Version

- Data are delivered after passing Quality Assurance (QA) testing.
- Download data from *Archive Query* and *Request Handler* tools in ALMA Science Portal
- Delivered data include:
  - Calibration tables and diagnostics
  - Preliminary images, *PI should plan to re-image*.

Goals of Quality Assurance (QA) Process

- Ensure reliable final data product
  - Desired sensitivity
  - Desired resolution
- Ensure data is free of calibration and imaging artifacts
- Errors in PI-supplied parameters are outside scope
  - Incorrect source coordinates
  - Inadequate frequency specification
  - Inadequate sensitivity limits

See ALMA Technical Handbook for details.

During Observations – QA0

- Monitoring of calibrations and overall performance
- Rapidly-varying parameters
  - Scheduling block timescales or shorter
  - Atmospheric effects
  - Antenna issues
  - Front-end issues (sideband ratios, antenna temp.)
  - Connectivity issues (System temp., delay measurements…)
  - Back-end issues (total power levels, delays…)
- Tolerances for each are explicitly laid out
  - Fewer than 26 antennas in 12-m array
  - Bandpass calibrator too weak

See ALMA Technical Handbook for details.

Between Observations – QA I

- Regularly Measured Array Performance
- Slowly Varying Parameters
  - Timescales of a week or longer
- Array Calibrations
  - Baseline measurements
  - Delays
- Antenna Calibrations
  - All-sky pointing
  - Focus curves
  - Beam patterns etc.
- Source Calibrations
  - Solar-system and quasar flux monitoring

After Observations – QA2

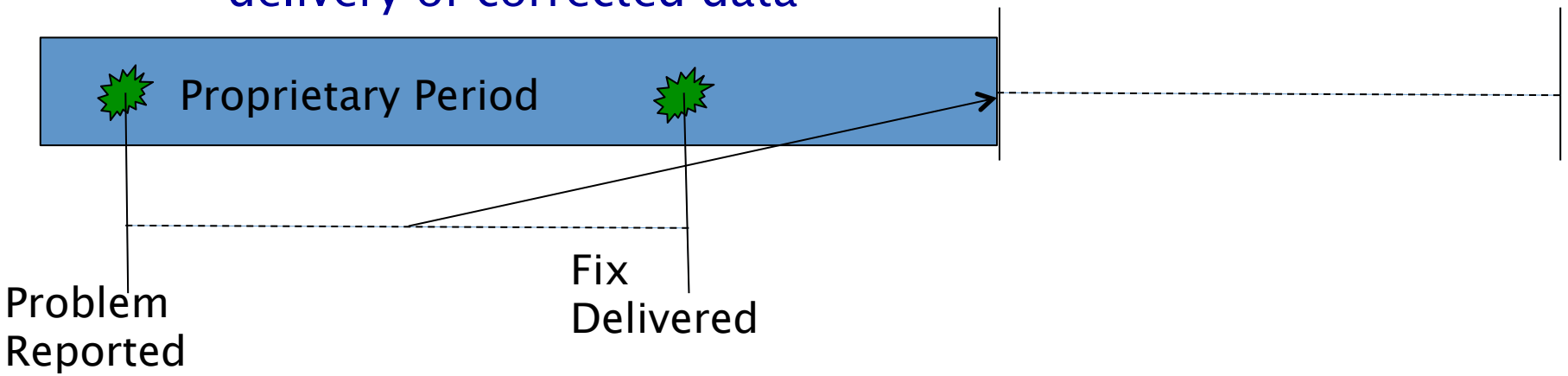
- Calibration by pipeline or analyst
  - Extreme T_{sys}
  - Water vapor measurements and corrections
  - Absolute flux scale certainty
  - Noisy or extreme bandpass corrections
  - Proper phase transfer cadence followed
  - Noisy phase and amplitude corrections
- Image data and compare with requested science goals
  - Signal-to-noise (within ~10% of goal)
  - Resolution
  - Artifacts (e.g. sidelobe levels or striping)

After Observations – QA3

- Post-reduction evaluation of **delivered** data products
- Triggered by PI reporting possible underlying problems with:
  - Data
  - Observing procedure
  - Calibration
- Likely results in fix being implemented and products re-ingested into ALMA archive
- May include extension of proprietary period

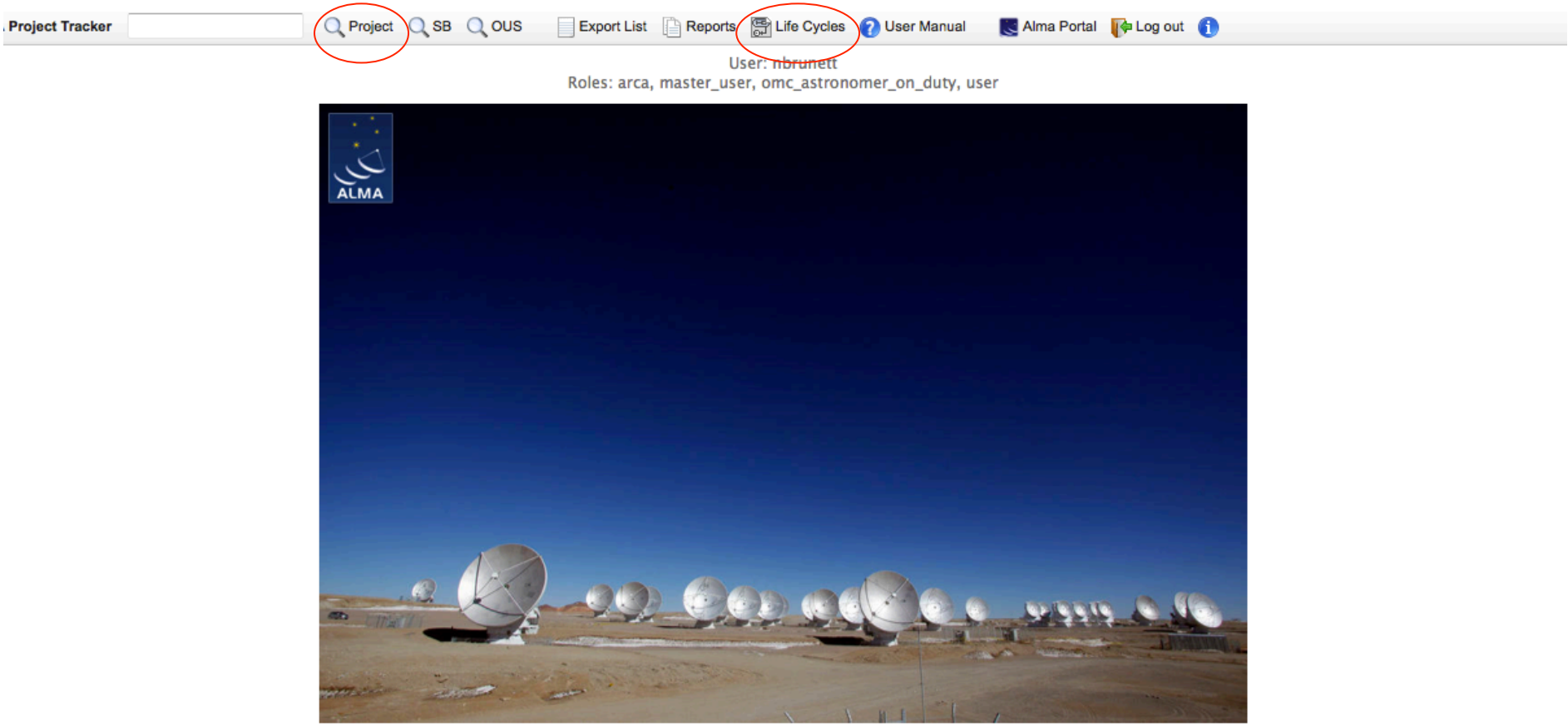
QA3

- May Affect Proprietary Period
  - *If reported within 2 months* of delivery:
    - Period extended by interval from original delivery to delivery of corrected data

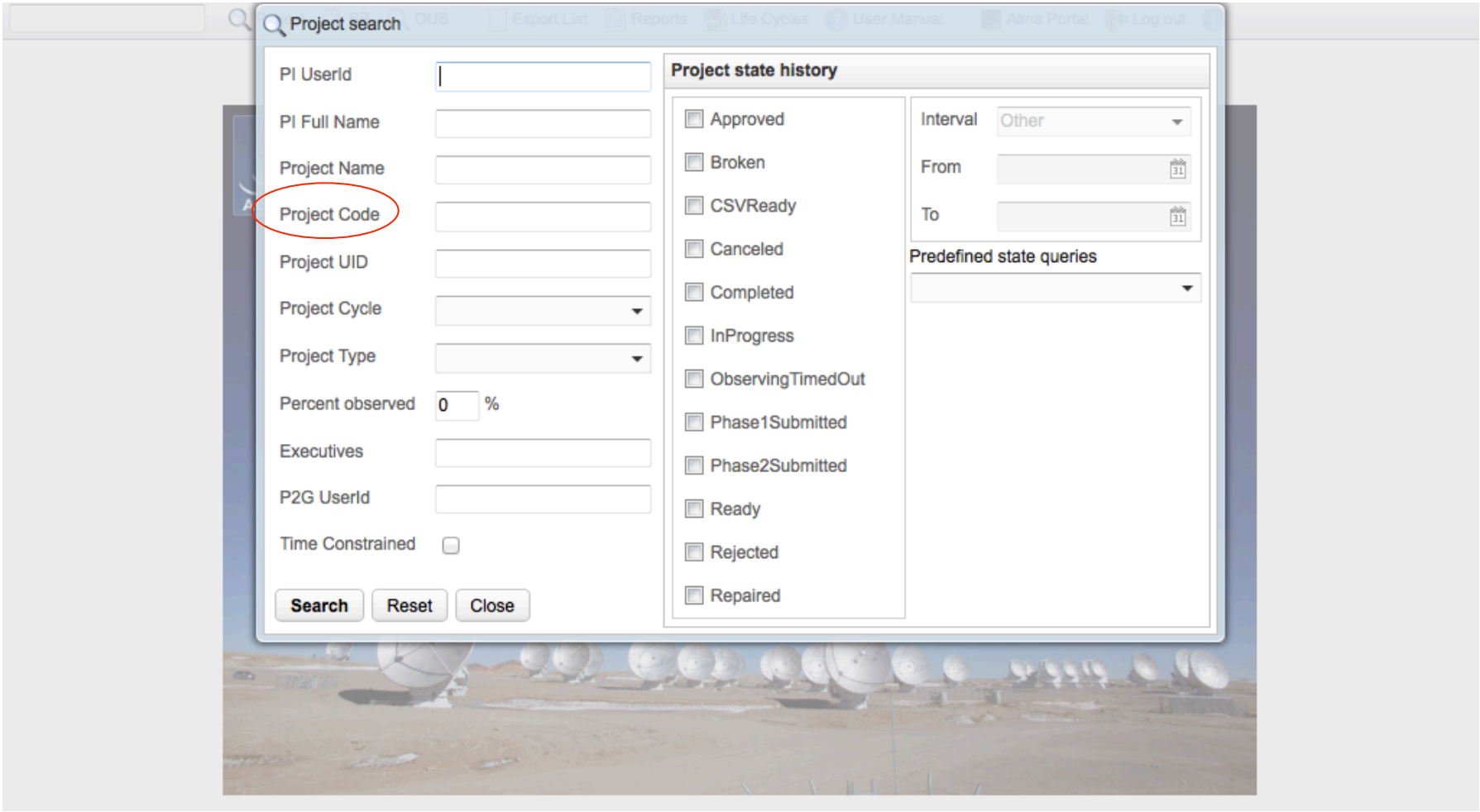


  - Else:
    - Period extended until delivery of corrected data

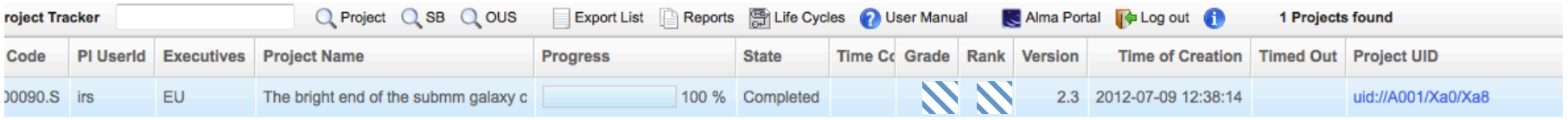
Monitor Project Status:
Project Tracker

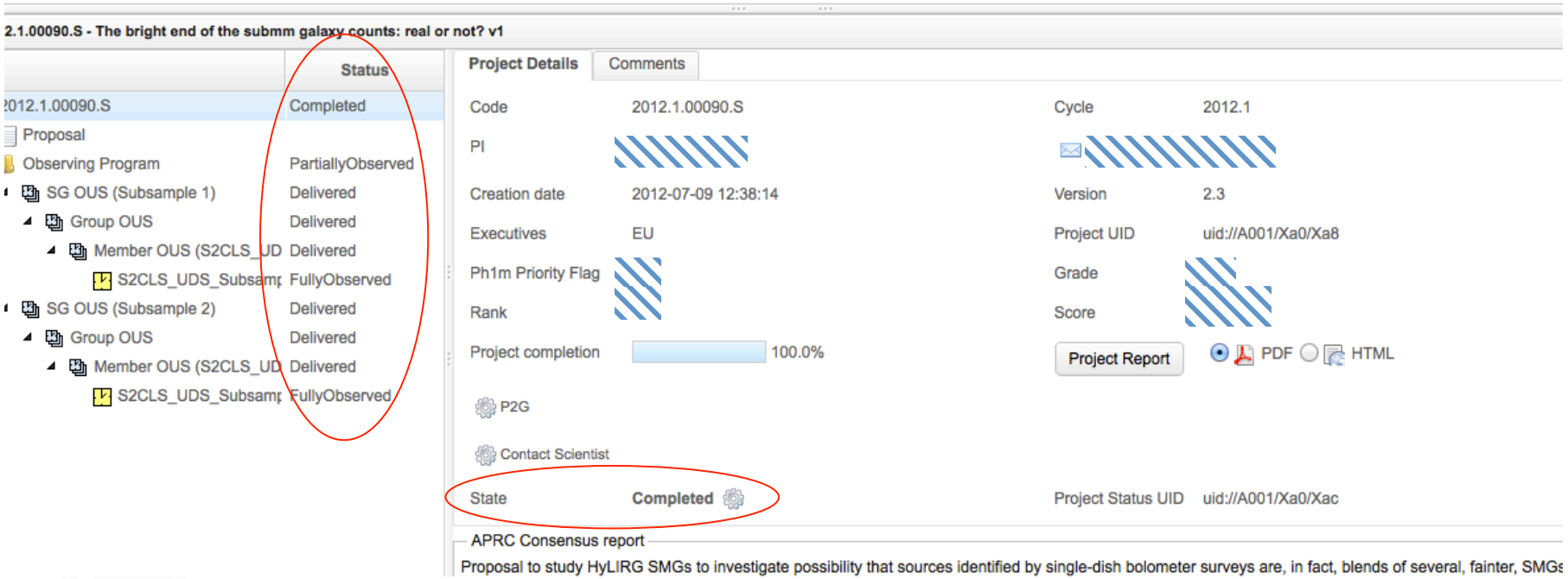


Project Tracker

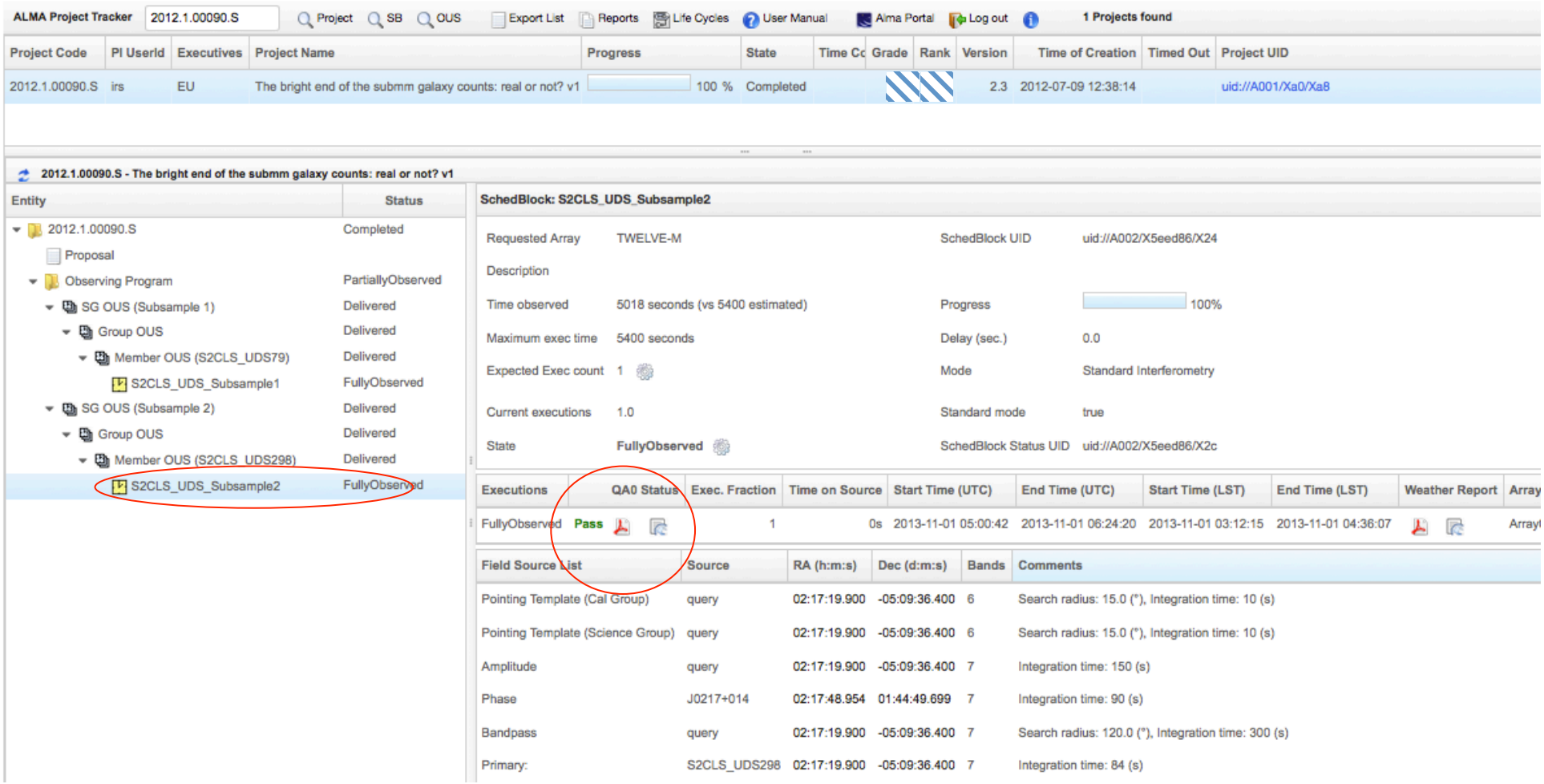


Project Tracker





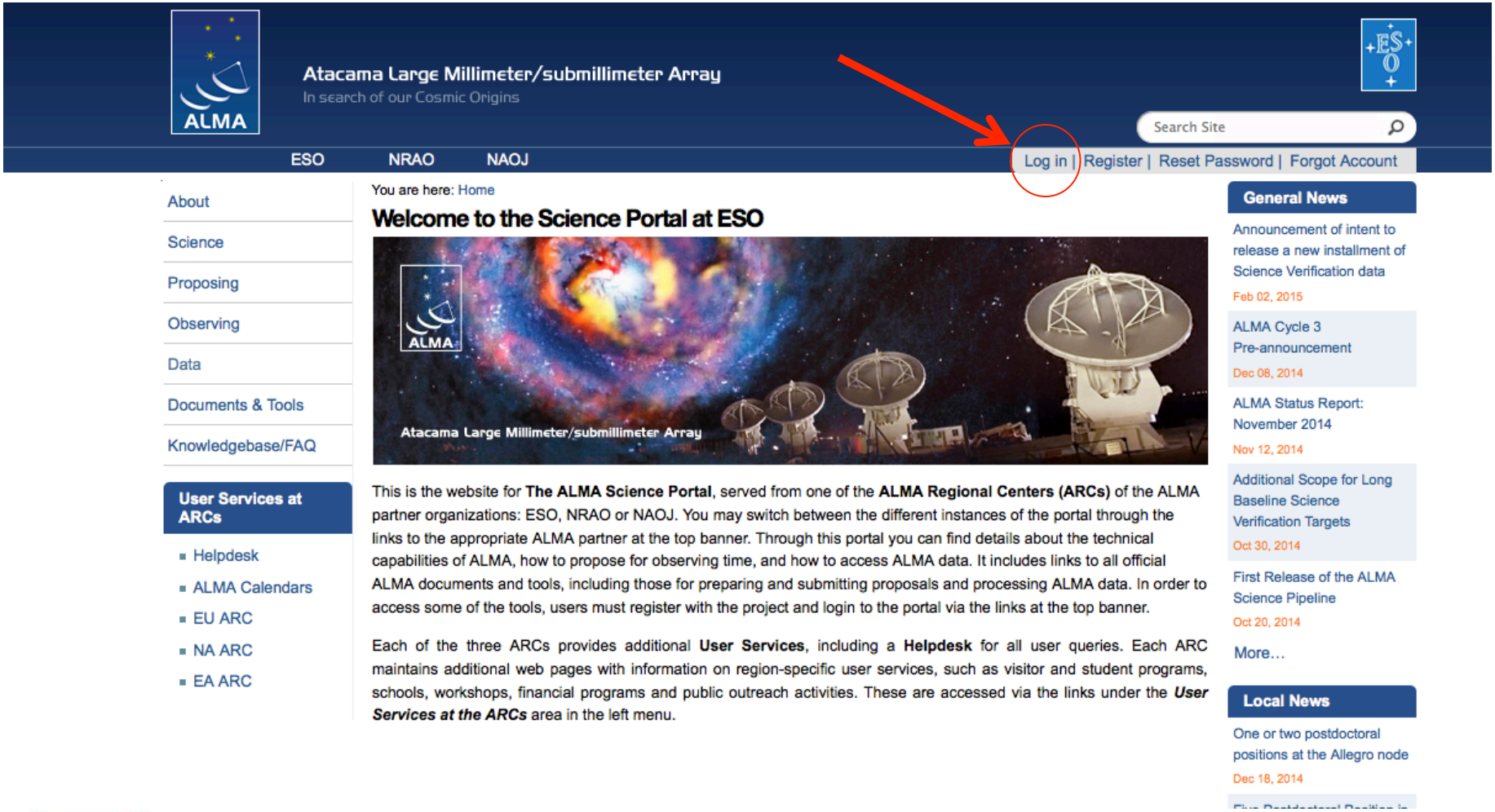
Project Tracker



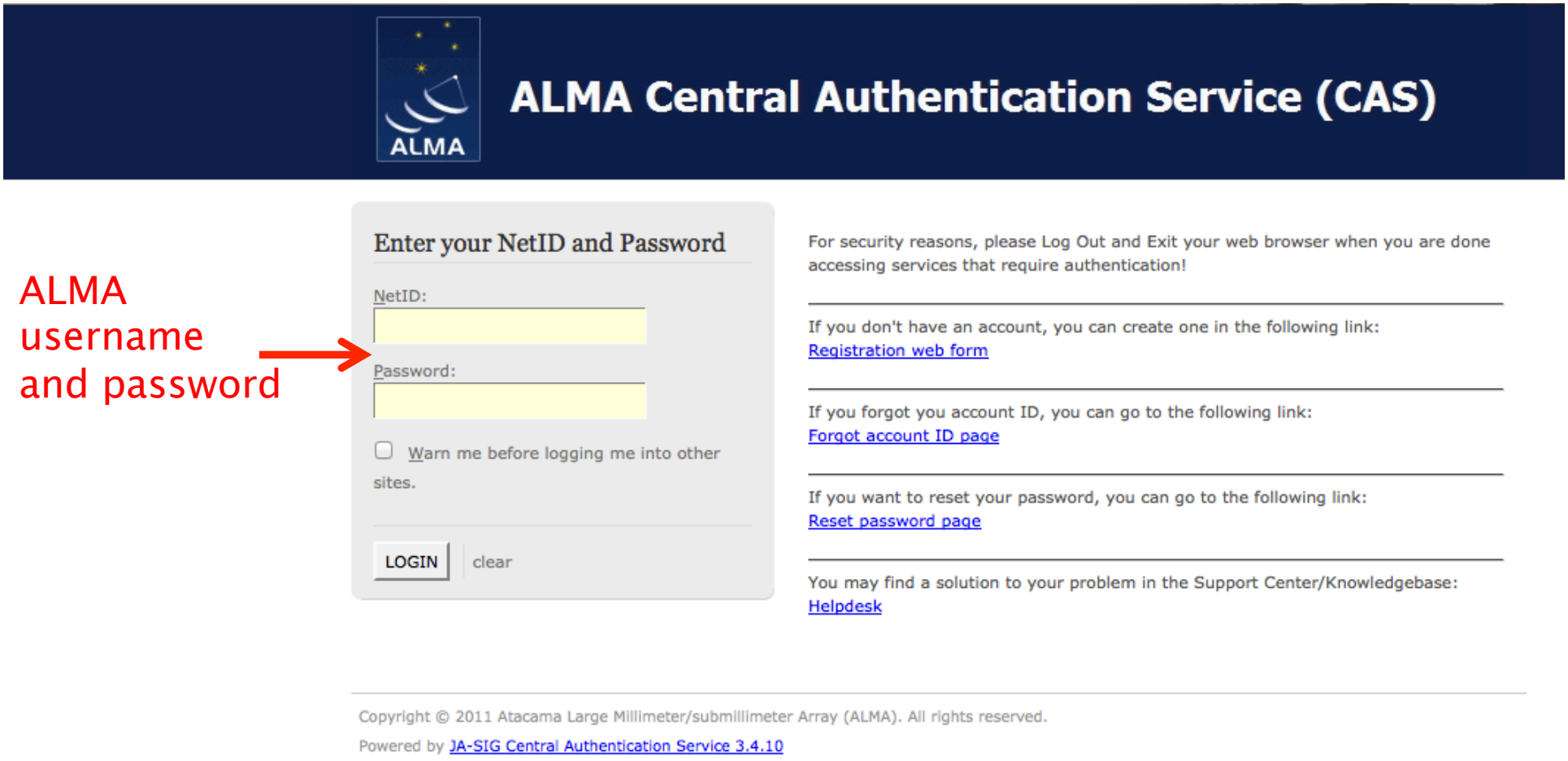
Monitor Project Status: Optional emails

- Subscribe to email notification for updates on changes to project status
  - …
  - Phase2Submitted
  - Running
  - Partially Observed
  - Fully Observed
  - Pipeline Processing
  - …
- With or without optional emails, PIs always receive notification when new data are available

Optional emails



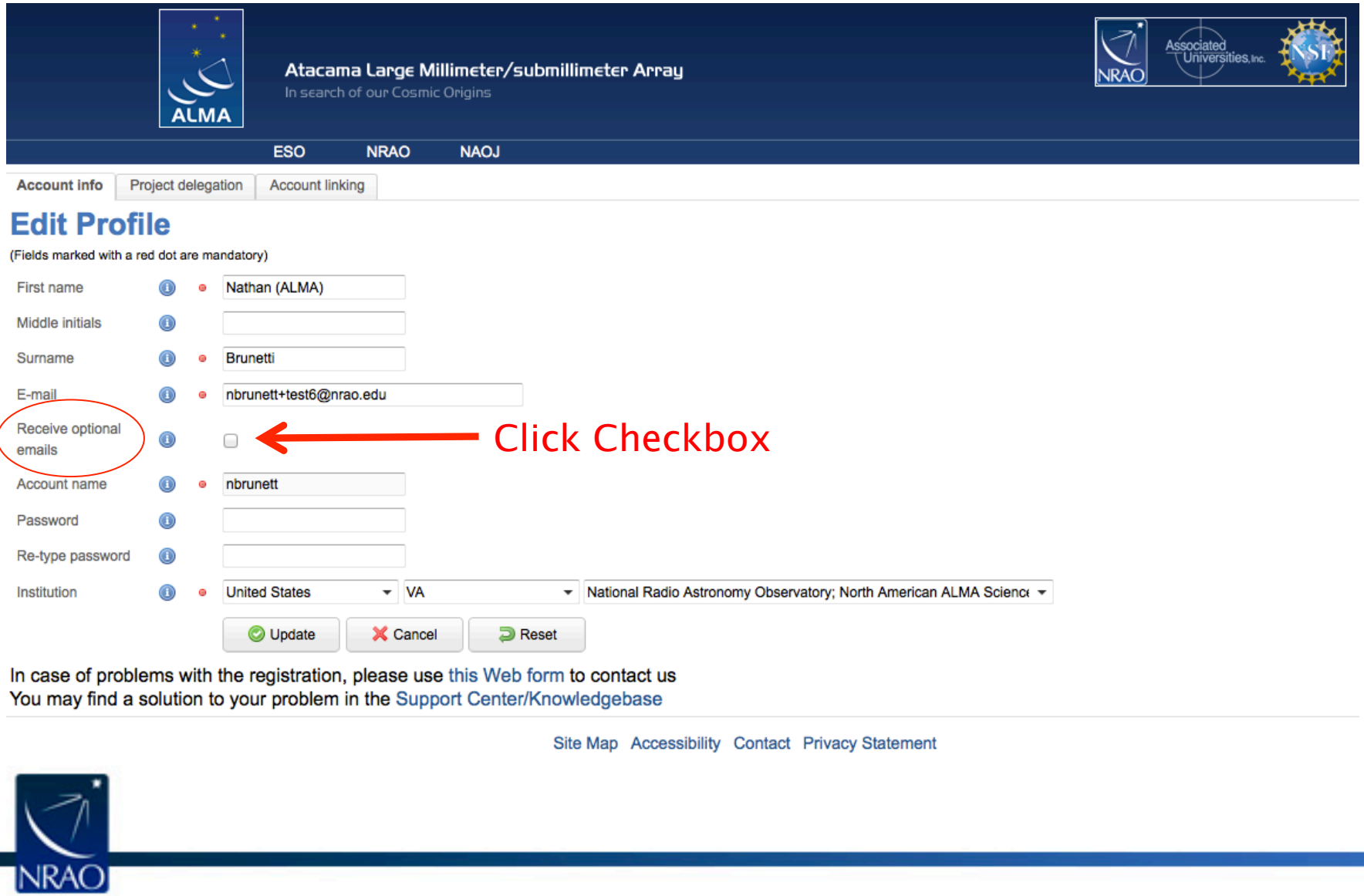
Optional emails



ALMA username and password

Optional emails

Optional emails



Data Delivery Email

- Sent when Member Observation Unit Set (MOUS) passes QA2 and data is copied to PI's Regional Center
- Triggers Start of Proprietary Period
  - Usually 12 months
- Only Sent to PI

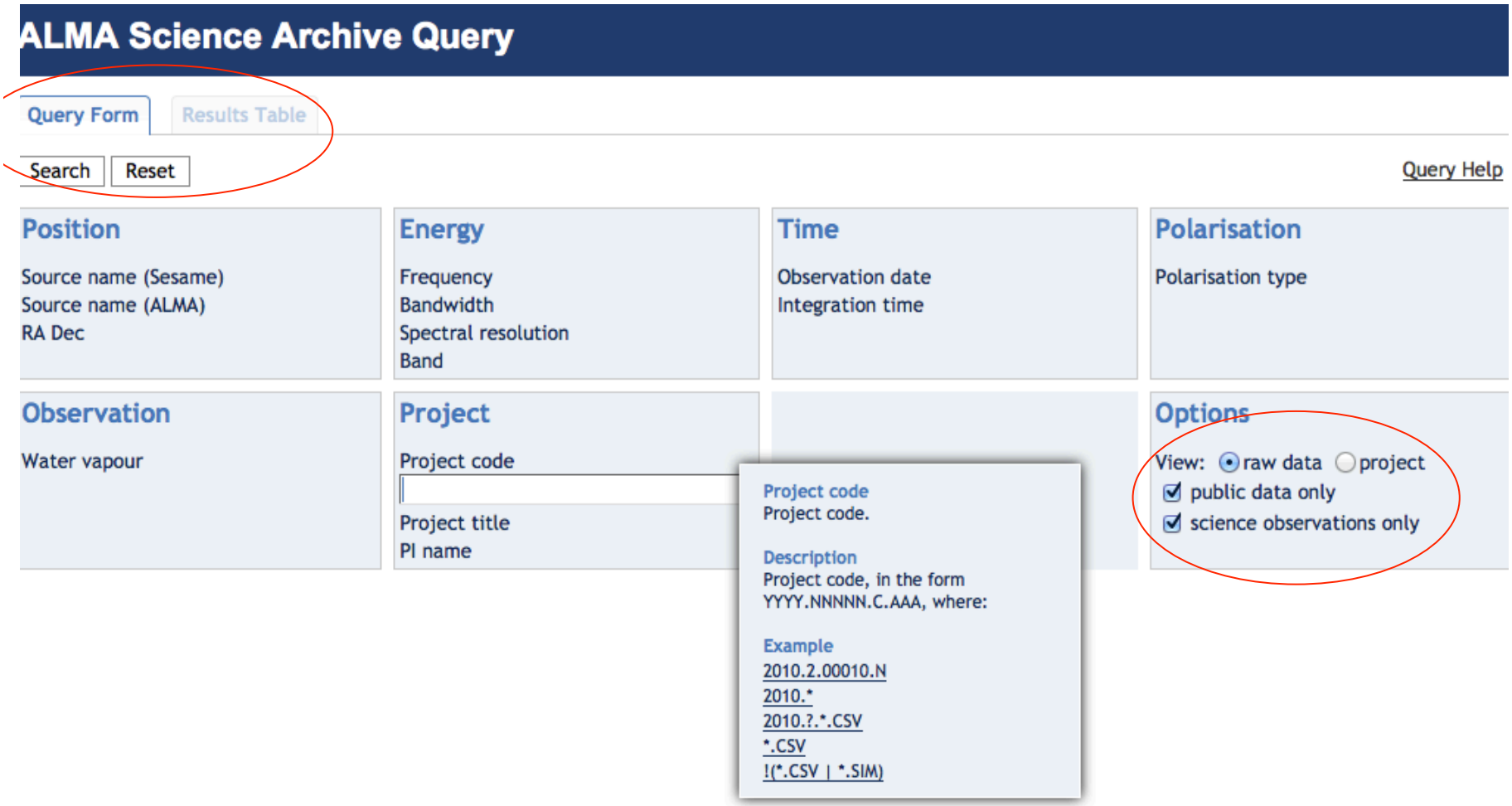
Data Delivery Email

- Included Metadata:
  - MOUS ID, Scheduling Block (SB) name, project title
- Included Instructions:
  - Downloading data
  - Delegating access for registered ALMA users
- Included Descriptions:
  - Proprietary period
- Included Links:
  - Archive query for MOUS package
  - Fully-calibrated MSs (NA ARC only)
  - CASA download and mailing lists

Data Delivery Email

- Included Warnings:
  - Early science data cannot be guaranteed to meet expectations of full scientific operations
- Publication Requirements:
  - ALMA acknowledgement
  - ARC specific acknowledgement
- Additional Support:
  - Funded face-to-face reduction visits to your home ARC
  - Contact info for ARC PR personnel

Find data in archive:
Archive Query



Archive Query

Query Form | Results Table

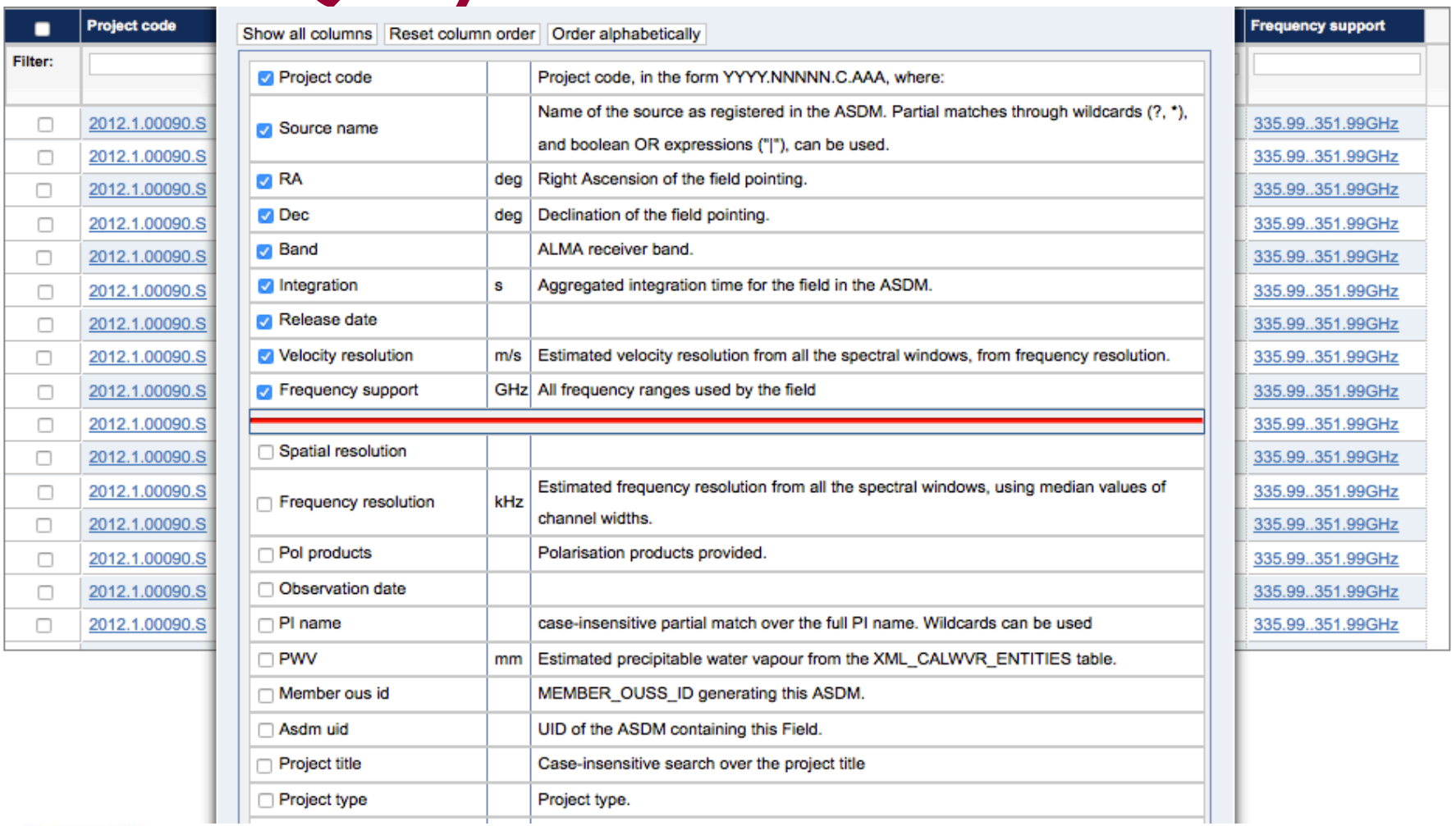
Submit download request

Results Bookmark | Export Table | Results Help

Showing 30 rows (30 before filtering). More columns

	Project code	Source name	RA	Dec	Band	Integration	*Release date* ▲	Velocity resolution	Frequency support
Filter:								m/s	
☑	2012.1.00090.S	S2CLS_UDS110	02:18:48.44	-05:18:05.0	7	9.326	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☑	2012.1.00090.S	S2CLS_UDS156	02:18:24.23	-05:22:53.4	7	8.836	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☑	2012.1.00090.S	S2CLS_UDS160	02:18:23.86	-05:11:36.2	7	8.842	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS168	02:18:20.34	-05:31:41.6	7	8.843	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☑	2012.1.00090.S	S2CLS_UDS199	02:18:07.38	-04:44:11.7	7	8.812	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS204	02:18:03.01	-05:28:39.8	7	8.873	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS216	02:17:56.80	-04:52:39.6	7	8.82	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS252	02:17:37.79	-05:20:10.2	7	8.827	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS286	02:17:25.76	-05:25:36.5	7	9.657	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS292	02:17:21.85	-05:19:03.3	7	8.815	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS298	02:17:19.90	-05:09:36.4	7	9.55	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS334	02:17:02.81	-04:57:24.9	7	8.856	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS408	02:16:22.59	-05:11:06.0	7	8.819	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS421	02:16:17.62	-05:09:02.0	7	8.803	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz
☐	2012.1.00090.S	S2CLS_UDS47	02:19:24.97	-05:09:19.9	7	8.785	2014-11-07T09:35:00.000	27236.96	336.00..351.99GHz

Archive Query: more columns



	Project code
Filter:	
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S
☐	2012.1.00090.S

Show all columns | Reset column order | Order alphabetically

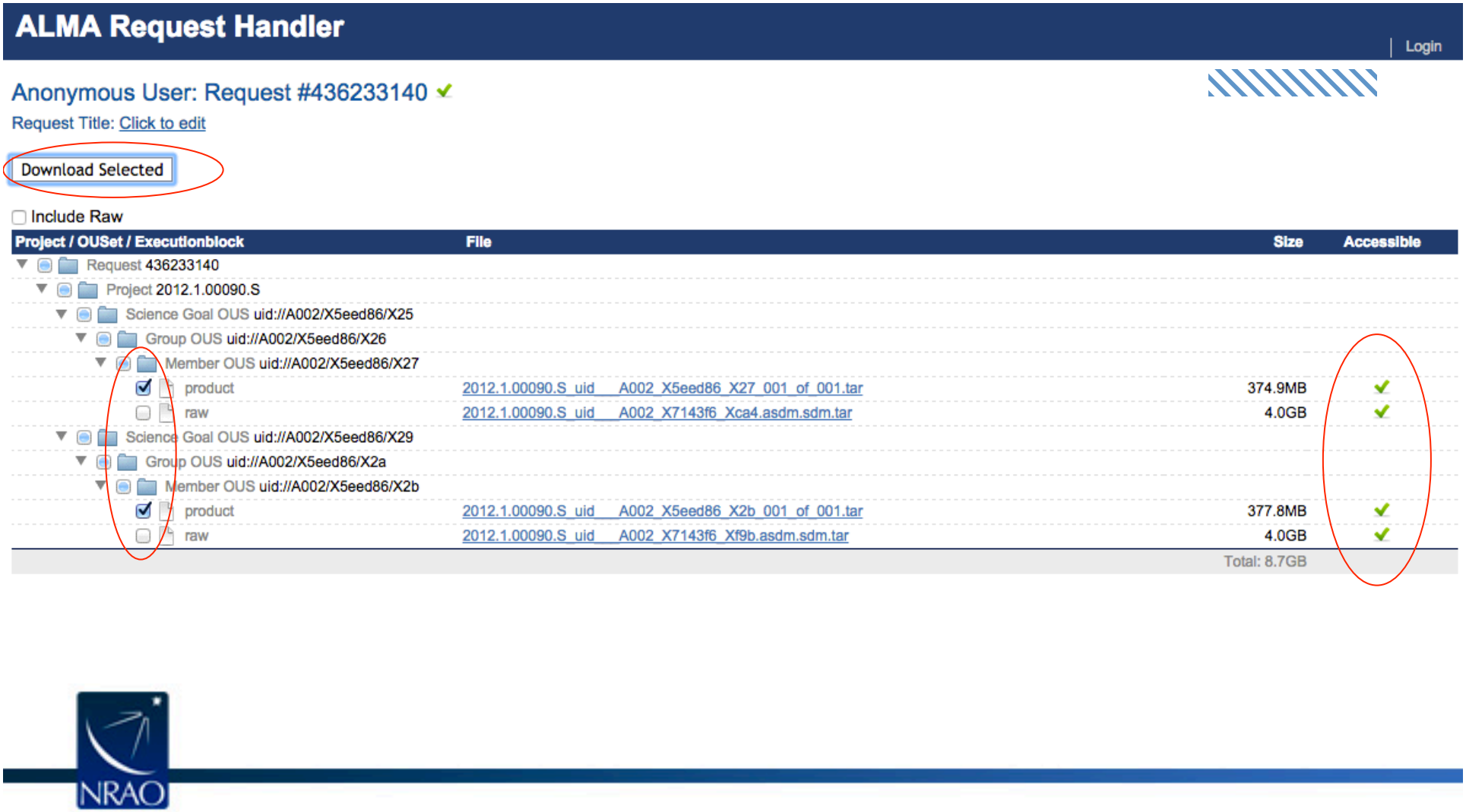
☑ Project code		Project code, in the form YYYY.NNNNN.C.AAA, where:
☑ Source name		Name of the source as registered in the ASDM. Partial matches through wildcards (?, *), and boolean OR expressions ("|"), can be used.
☑ RA	deg	Right Ascension of the field pointing.
☑ Dec	deg	Declination of the field pointing.
☑ Band		ALMA receiver band.
☑ Integration	s	Aggregated integration time for the field in the ASDM.
☑ Release date		
☑ Velocity resolution	m/s	Estimated velocity resolution from all the spectral windows, from frequency resolution.
☑ Frequency support	GHz	All frequency ranges used by the field
☐ Spatial resolution		
☐ Frequency resolution	kHz	Estimated frequency resolution from all the spectral windows, using median values of channel widths.
☐ Pol products		Polarisation products provided.
☐ Observation date		
☐ PI name		case-insensitive partial match over the full PI name. Wildcards can be used
☐ PWV	mm	Estimated precipitable water vapour from the XML_CALWVR_ENTITIES table.
☐ Member ous id		MEMBER_OUSS_ID generating this ASDM.
☐ Asdm uid		UID of the ASDM containing this Field.
☐ Project title		Case-insensitive search over the project title
☐ Project type		Project type.

Frequency support
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz
335.99..351.99GHz

Downloading the data:
Request Handler

- All data downloaded as tar files
- Large data sets may be broken into several pieces
  - Name is `[project_code]_[OUS_ID]_m_of_n.tar`
  - Raw data packaged as one tar file per execution block (EB)
    - name is [project_code]_[EB_ID].asdm.sdm.tar
- Cannot directly download individual data products yet…
  - FITS images
  - Diagnostic plots, etc.

Request Handler



Request Handler

From do-not-reply@nrao.edu

Subject **ALMA Archive at NRAO: Request 223292105**

Reply to [redacted]

To [redacted]

Reply | Reply All | Forward | Archive | Junk | Delete

1:24 PM

Other Actions

Dear [redacted],

Thank you for using the ALMA archive.

Your data selection (4.3GB) is available from this link

https://almascience.nrao.edu/rh/requests/nbrunett/223292105

We hope they meet your expectations and will lead to a successful completion of your scientific program.

Publications making use of these data must include the following statement in the acknowledgment:

"This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00090.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ."

Please submit your requests for help, for a visit to the ARC, or to report any problems discovered in your data through the ALMA Helpdesk at https://help.almascience.org.

Best regards,

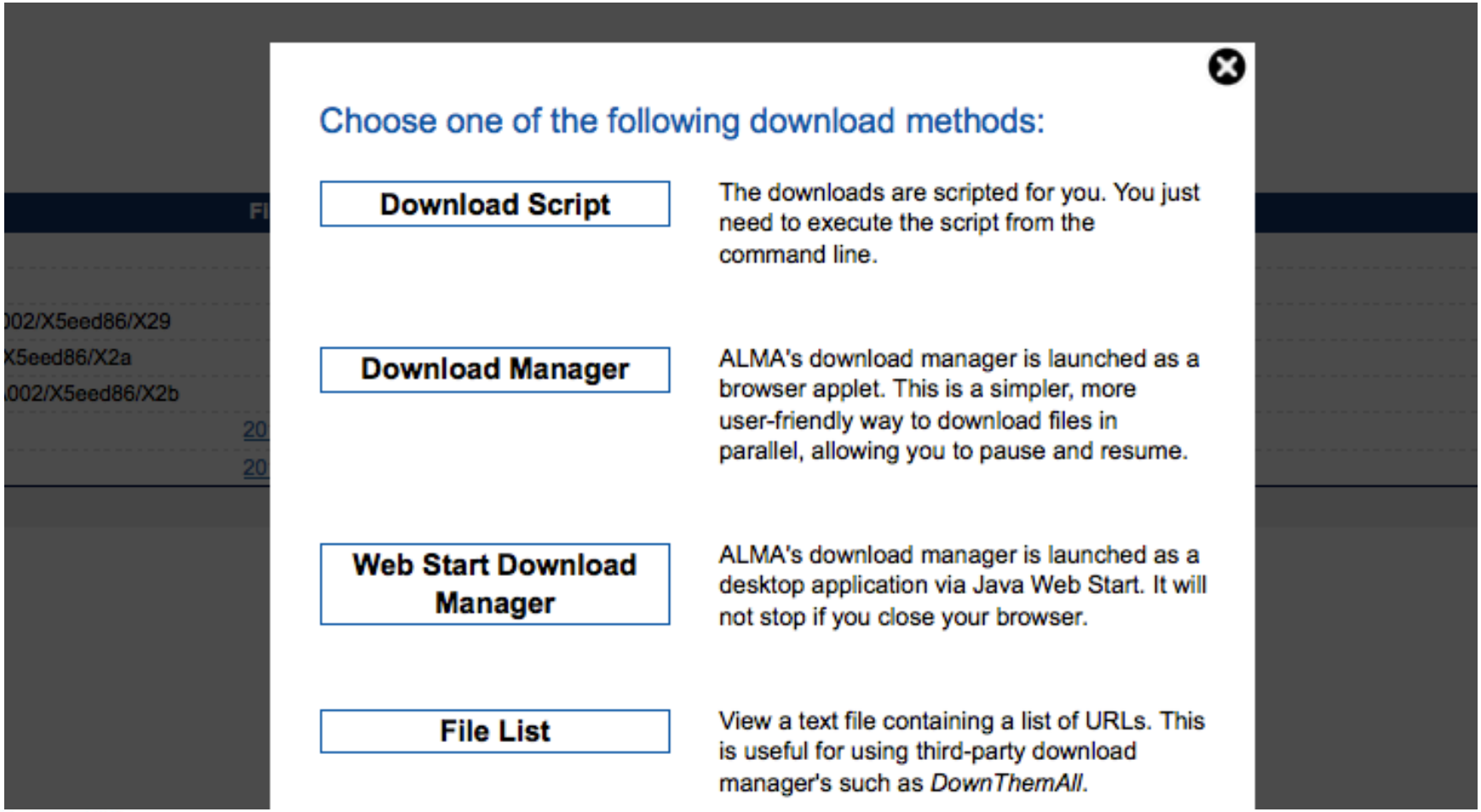
The North American ALMA Archive at the NAASC

Summary:
Files available: 2 (4.3GB)
Files under proprietary period: 0 (-)
Files not available: 0 (-)

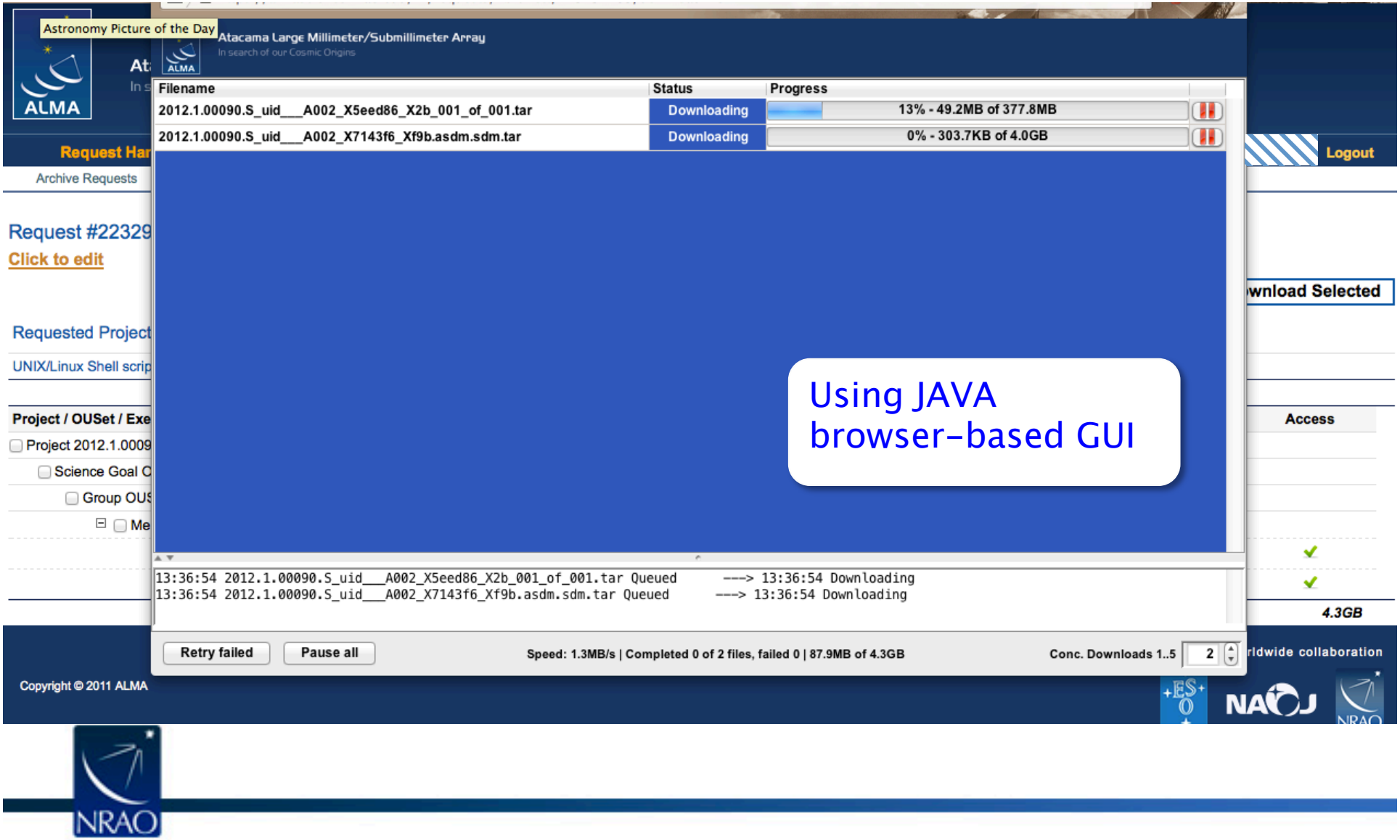
Details:
Files available:
- 2012.1.00090.S_uid___A002_X5eed86_X2b_001_of_001.tar : AUTHORIZED
- 2012.1.00090.S_uid___A002_X7143f6_Xf9b.asdm.sdm.tar : AUTHORIZED

Files under proprietary period:

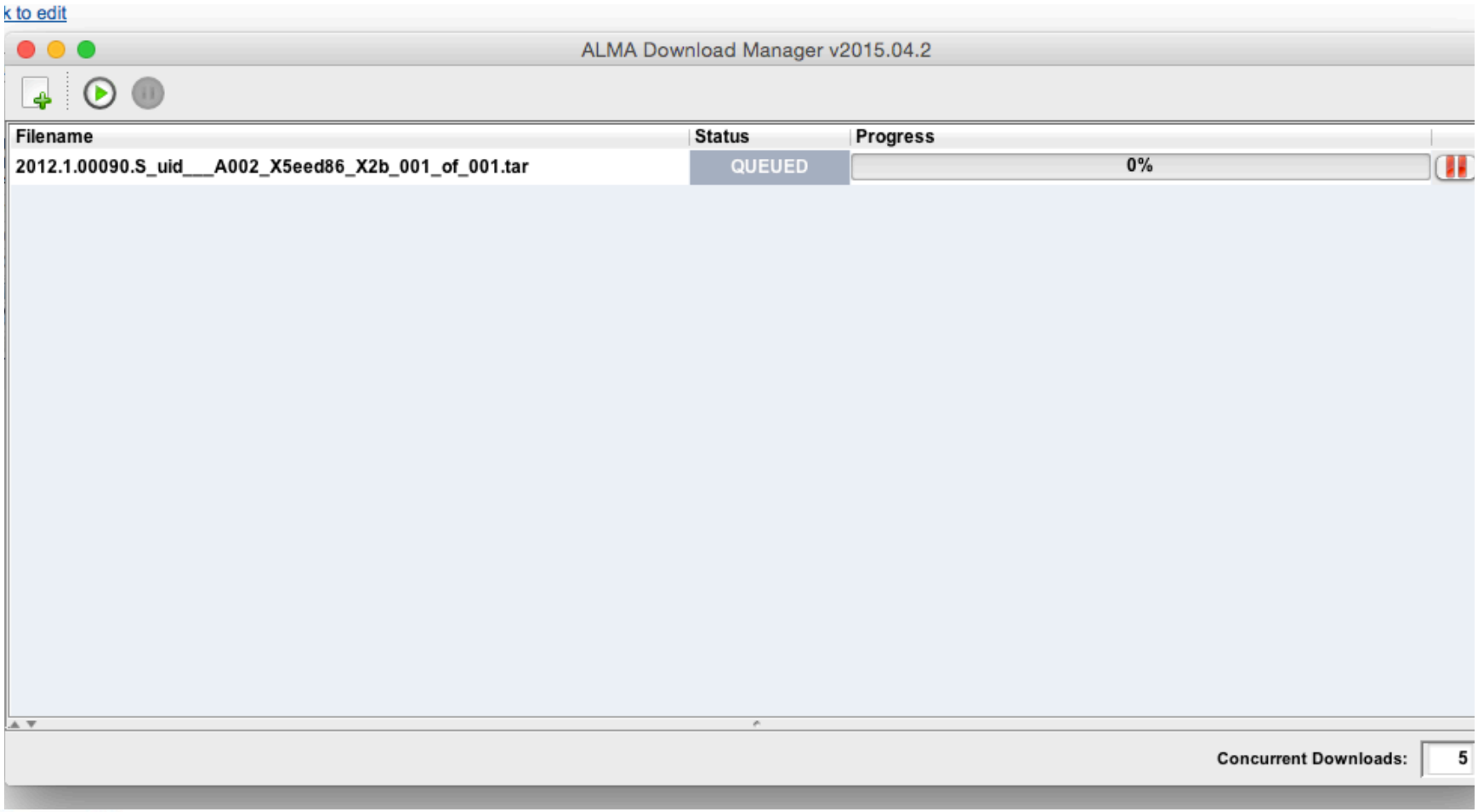
Request Handler: Download options



Request Handler: Download Manager



Request Handler: Web start



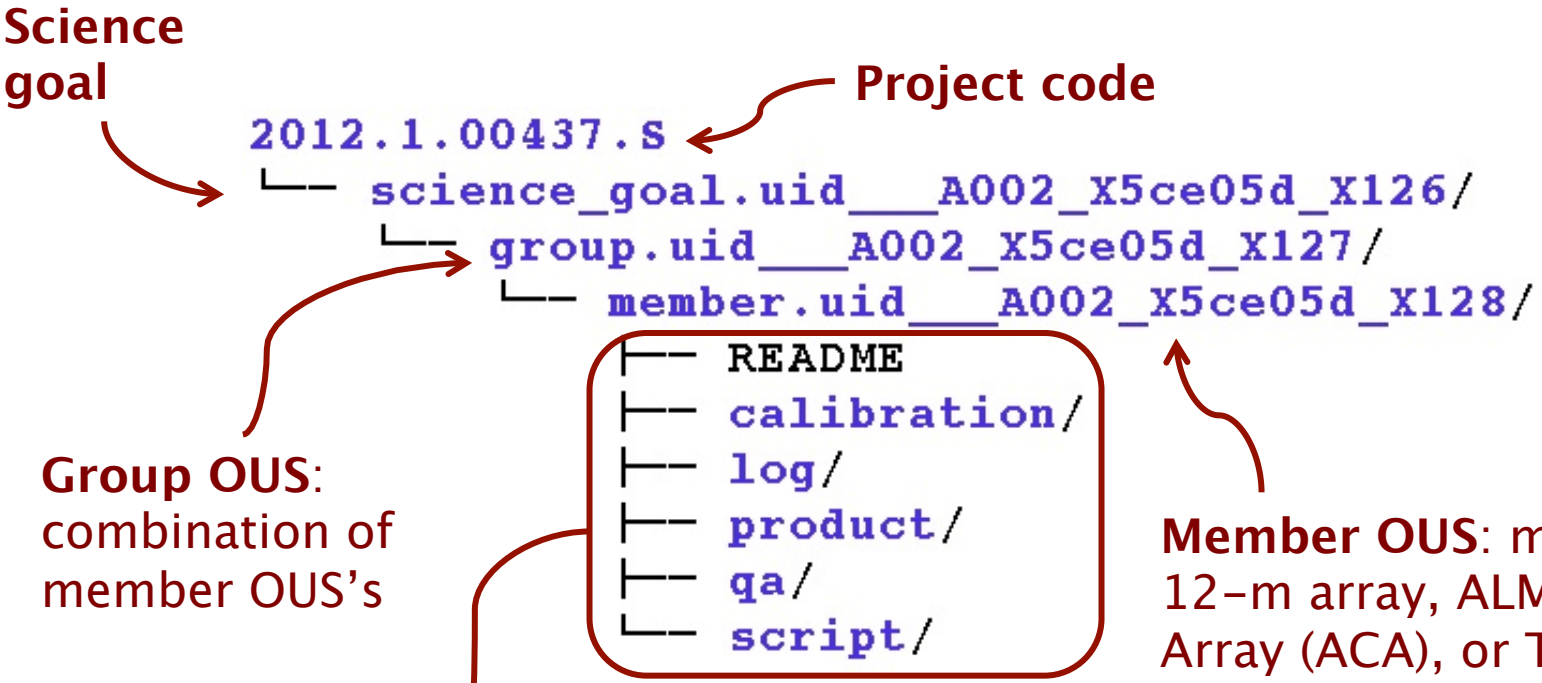
Request Handler: script

Using auto-generated shell script (wget)

```
#!/bin/bash
#Please use the current script to download the whole content of request
223732763

echo "Please provide a password"
read -s PASSWORD

wget   --auth-no-challenge  --no-check-certificate  --http-user="nbrunett" --
http-password=$PASSWORD https://almascience.nrao.edu/dataPortal/api/requests/
nbrunett/223732763/ALMA/2012.1.00090.S_uid___A002_X5eed86_X2b_001_of_001.tar/
2012.1.00090.S_uid___A002_X5eed86_X2b_001_of_001.tar
.
.
.
```

QA2 Data Products Package: the processed data

After un-tarring the processed data we have a directory tree:



QA2 Data Products Package: the README file

Contains a summary of the QA2 results: achieved resolution and image noise RMS.

Shown here is an example README for a Cycle 2 project. (Blurred information protected by proprietary period.)

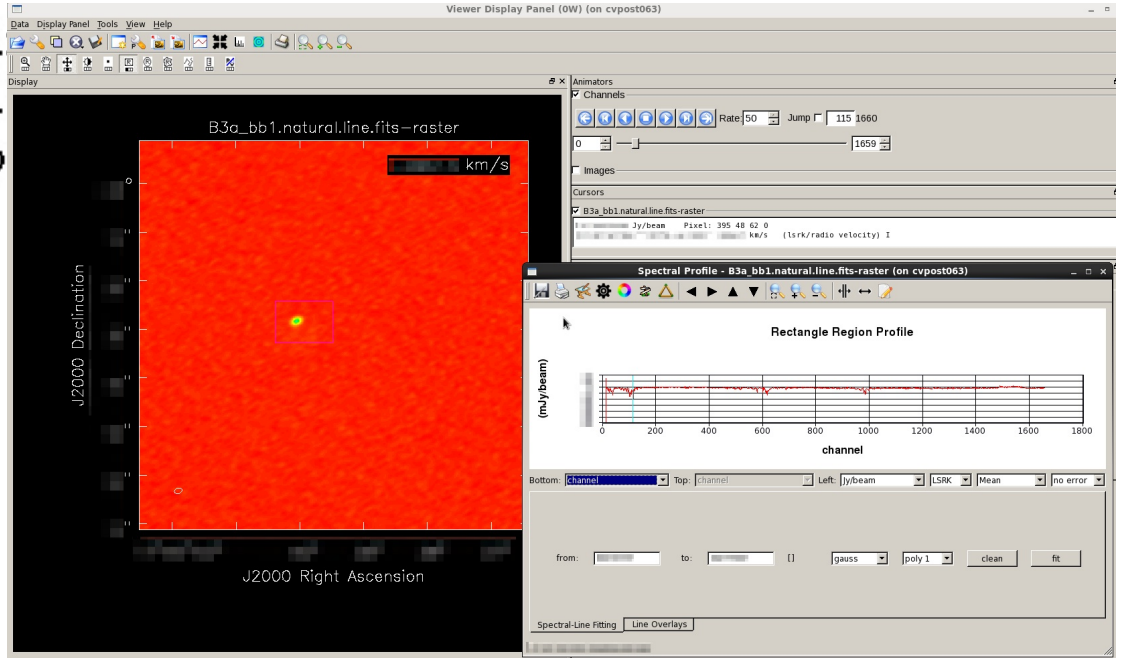
README — Edited

```
Atacama Large Millimeter/submillimeter Array (ALMA)

#####

Cycle: 2 (Early science)
Project code: 2013.1.00857.S
SB name: Circumnu_a_03_TE
PI name: Elisabeth Mills
Project title: The Density (and Destiny) of the Circumnuclear Disk
Configuration:  " x  " (natural)
Proposed rms:  mJy per  MHz channel ( km/s for  GHz)
CASA version used for reduction: 4.2.1
Comments from Reducer:

These data appear very good, requiring little flagging. The PI should
note there do seem to be a few bad channels in the reconstructed image
cube ( and ) that they may wish to flag out.
The rms achieved in channels which appear empty of emission or
absorption is  mJy in natural weighting. The beam of  x  "
is slightly bigger than the requested  ", but since the proposers note
in the TJ that they mean to separate sources about  " in size, this
resolution should meet their requirements. Slightly better resolution is
achieved in uniform weighting ( x ").
The image cubes are continuum subtracted. The continuum peak is
Jy, a S/N of .
Several line detections can be seen in the spectra.
```

QA2 Data Products Package: product directory

- Contains “best efforts” images of science target
  - used to confirm PI requirements for sensitivity and resolution.
  - PI should re-image target prior to publication.

```
product
├── B3a_bb1.natural.cont.fits
├── B3a_bb1.natural.flux.fits
├── B3a_bb1.natural.i
├── B3a_bb1.natural.l
└── B3a_bb1.natural.p
```



QA2 Data Products Package: qa directory

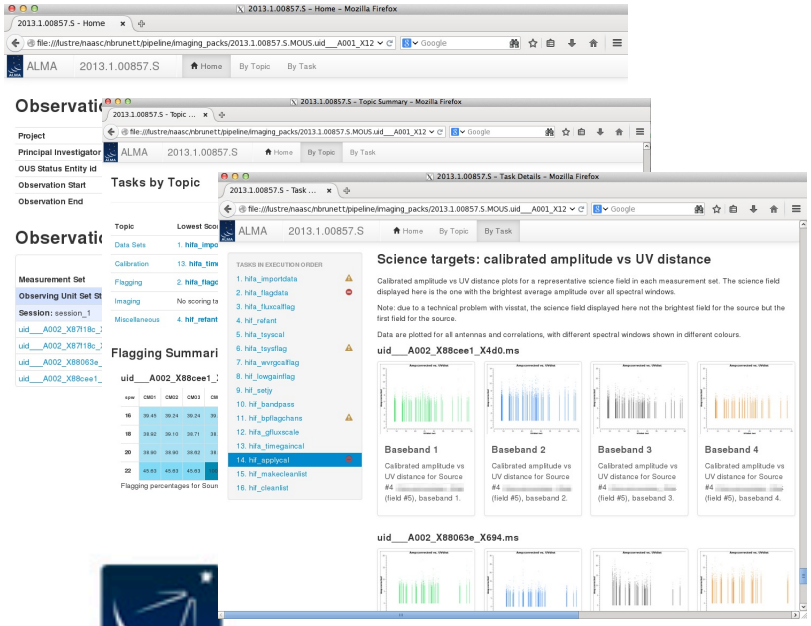
Contains diagnostic information, tables, and plots. Pipeline produces an HTML weblog for convenient viewing of information.

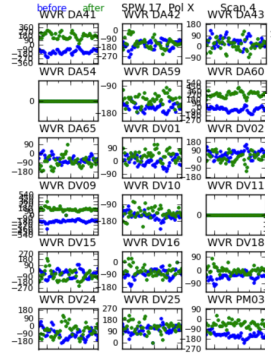
Pipeline calibration

```
qa
├── pipeline-20141111T191030/
│   └── html/
└── uid___A001_X121_X2f5.weblog.tar.gz
```

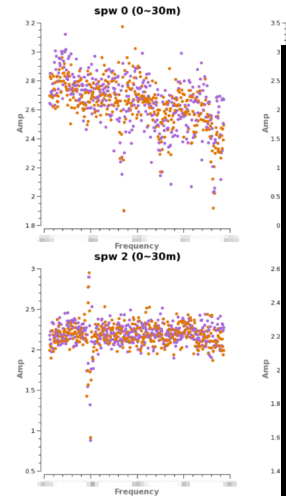
Manual calibration

```
qa
├── uid___A002_X87544a_X22d1__qa2_part1.png
├── uid___A002_X87544a_X22d1__qa2_part2.png
├── uid___A002_X87544a_X22d1__qa2_part3.png
└── uid___A002_X87544a_X22d1__textfile.txt
```

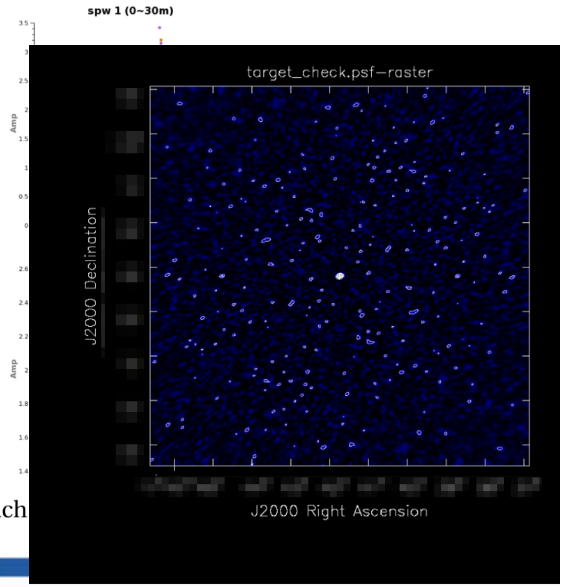




Phase: before/afte



Target Spectrum for each



Target psf

QA2 Data Products Package: Log Directory – *Pipeline Calib.*

Contains CASA log files from QA2 processing

```
log
├── casapy-20141129-145932.log
├── casapy-20141129-150412.log
├── casapy-20141129-150456.log
├── casapy-20141223-170951.log
├── casapy-20141229-164757.log
└── uid___A001_X121_X2e5.casa_commands.log
```

- `casapy` log files contain pipeline logs for data import, applying calibration, flagging, and imaging
- `casa_commands` log file:
  - Record of all CASA commands run
  - Does not include heuristic and flagging calculations
  - Cannot be executed
  - Shows which CASA tasks are run by each pipeline task

QA2 Data Products Package: Log Directory – *Manual Calib.*

Contains CASA log files from QA2 processing

```
log
├── casapy-20140918-205633.log
└── uid___A002_X87544a_X22d1.log.tgz
```

Contains some or all CASA logs for data reduction process. Number of files varies, depending on the approach taken by the manual reducer.

QA2 Data Products Package: Calibration directory

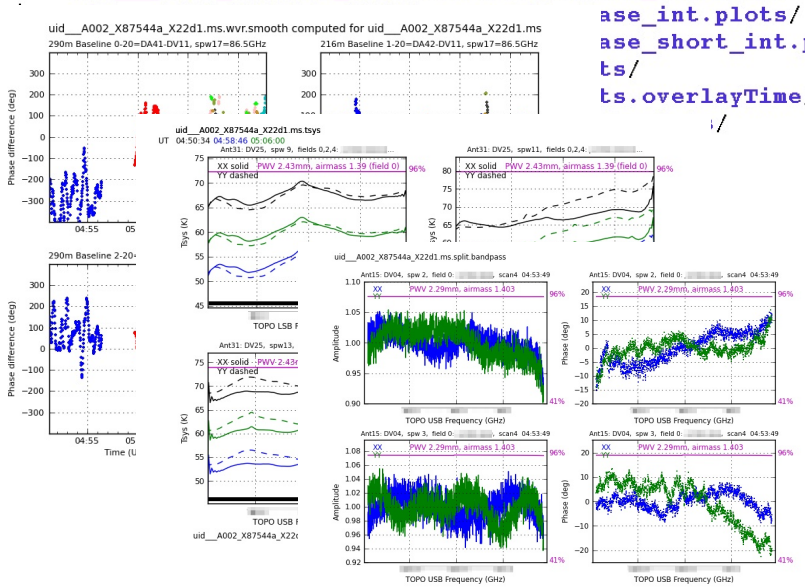
Contains CASA calibration tables, flagging inputs, diagnostic plots (manual calibration only)

Pipeline calibration

```
calibration
├── flux.csv
├── uid___A001_X121_X2f5.session_1.caltables.tar.gz
├── uid___A002_X87f18c_X116b_flagtemplate.txt
├── uid___A002_X87f18c_X116b_flagtemplate.txt~
├── uid___A002_X87f18c_X116b.ms.calapply.txt
├── uid___A002_X87f18c_X116b.ms.flagversions.tar.gz
├── uid___A002_X87f18c_Xed9_flagtemplate.txt
├── uid___A002_X87f18c_Xed9_flagtemplate.txt~
├── uid___A002_X87f18c_Xed9.ms.calapply.txt
├── uid___A002_X87f18c_Xed9.ms.flagversions.tar.gz
├── uid___A002_X88063e_X694_flagtemplate.txt
├── uid___A002_X88063e_X694_flagtemplate.txt~
├── uid___A002_X88063e_X694.ms.calapply.txt
├── uid___A002_X88063e_X694.ms.flagversions.tar.gz
├── uid___A002_X88cee1_X4d0_flagtemplate.txt
├── uid___A002_X88cee1_X4d0_flagtemplate.txt~
├── uid___A002_X88cee1_X4d0.ms.calapply.txt
└── uid___A002_X88cee1_X4d0.ms.flagversions.tar.gz
```

Manual calibration

```
calibration/
├── uid___A002_X87544a_X22d1.calibration.plots/
│   ├── uid___A002_X87544a_X22d1.ms.split.ampli_short_inf.plots/
│   ├── uid___A002_X87544a_X22d1.ms.split.ap_pre_bandpass.plots/
│   ├── uid___A002_X87544a_X22d1.ms.split.bandpass.plots/
│   ├── uid___A002_X87544a_X22d1.ms.split.bandpass_smooth20ch.plots/
│   ├── uid___A002_X87544a_X22d1.ms.split.flux_inf.plots/
│   ├── uid___A002_X87544a_X22d1.ms.split.phase_inf.plots/
│                                                  ase_int.plots/
│                                                  ase_short_int.plots/
│                                                  ts/
│                                                  ts.overlayTime/
│                                                  /
├
└
```



QA2 Data Products Package: Script Directory - *Pipeline Calib.*

Contains scripts for calibration, imaging and (optionally) flux equalization

```
script
4 ├── casa_piperestorescript.py
3 ├── casa_pipescript.py
5 ├── PPR_uid___A001_X121_X2e6.xml
2 ├── scriptForImaging.py
1 └── scriptForPI.py
```

1. top-level wrapper restores calibrated MS(s)
2. regenerates imaging products
3. calibrates data from scratch using pipeline tasks
4. Restores calibrated MS – called by scriptForPI.py
5. pipeline input parameters

QA2 Data Products Package: Script Directory - *Manual Calib.*

Contains scripts for calibration, imaging and (optionally) flux equalization

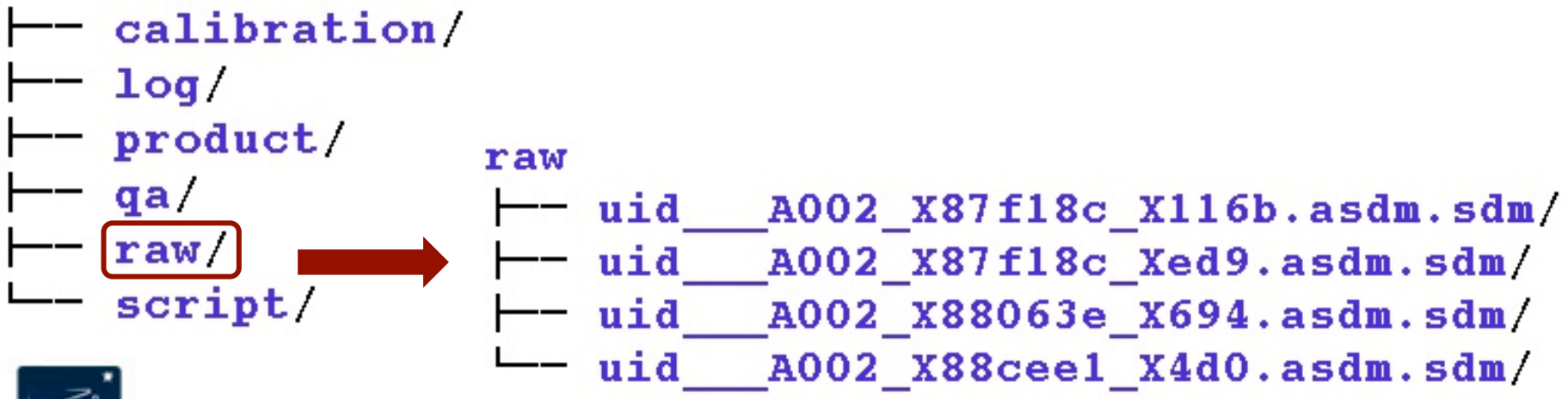
```
script
2 ├── scriptForImaging_22d1_natural.py
2 ├── scriptForImaging_22d1_uniform.py
1 ├── scriptForPI.py
3 └── uid___A002_X87544a_X22d1.ms.scriptForCalibration.py
```

1. top-level wrapper restores calibrated MS(s)
2. regenerates imaging products
3. calibrates single Execution Block

`scriptForFluxCalibration.py` – combines calibrated MSs and sometimes sets all data to same absolute flux scale

Raw Data

- Downloaded through the Request Handler.
- Format: ALMA Science Data Model (ASDM)
  - Directory tree containing binary and XML files
  - Converted to Measurement Set by CASA
- Must be placed in “raw” directory when re-running calibration scripts.



Calibration Process

- Remove affects of atmospheric water vapor
- Correct frequency-dependent
- Correct time-varying phases and amplitudes
- Set absolute flux scale
- Remove problematic data (flagging)

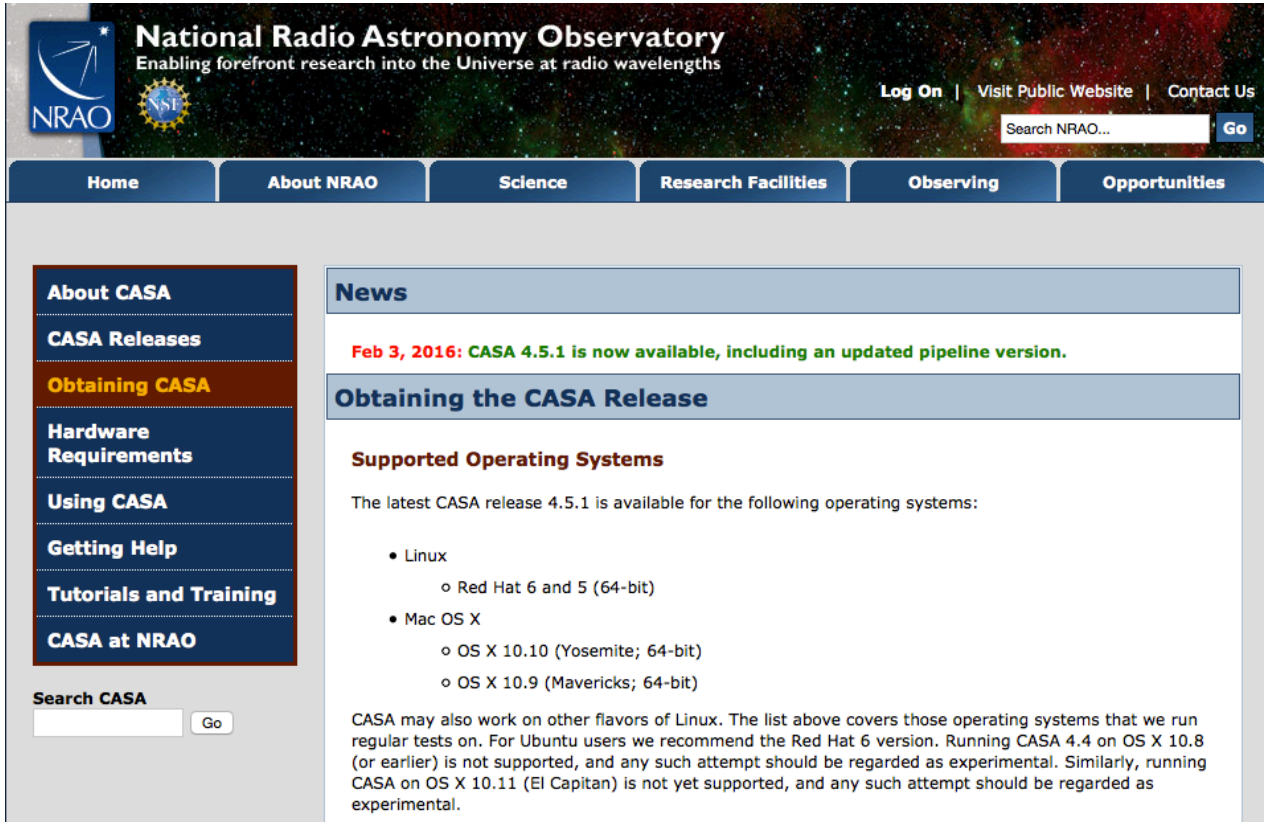
Example pipeline calibration

```
from recipes.almahelpers import fixsyscaltimes
__rethrow_casa_exceptions = True
h_init()
try:
    hifa_importdata(vis=['uid___A002_X88e746_X735'], sess
    fixsyscaltimes(vis = 'uid___A002_X88e746_X735.ms')
    fixplanets(vis = 'uid___A002_X88e746_X735.ms', field
    hifa_flagdata(pipelinemode="automatic")
    hifa_fluxcalflag(pipelinemode="automatic")
    hif_refant(pipelinemode="automatic")
    hifa_tsyscal(pipelinemode="automatic")
    hifa_tsysflag(pipelinemode="automatic")
    hifa_wvrgcalflag(pipelinemode="automatic")
    hif_lowgainflag(pipelinemode="automatic")
    hif_setjy(pipelinemode="automatic")
    hif_bandpass(pipelinemode="automatic")
    hif_bpflagchans(pipelinemode="automatic")
    hifa_gfluxscale(pipelinemode="automatic")
    hifa_timegaincal(pipelinemode="automatic")
    hif_applycal(pipelinemode="automatic")
    hif_makecleanlist(intent='PHASE,BANDPASS,CHECK')
    hif_cleanlist(pipelinemode="automatic")
finally:
    h_save()
```

Obtaining CASA

http://casa.nrao.edu/casa_obtaining.shtml

- Supported OS's
  - Linux – RedHat 5 and 6
  - Mac OS (10.8)
- Binary distribution for each OS, *with and without pipeline*
- Source distribution



Calibration Process

- Initial Calibration
  - Water vapor radiometer (WVR)
  - System temperature (T_{sys})
  - Antenna positions

Calibration Process

- WVR
  - 183 GHz radiometers
  - Mounted on each antenna
  - Estimates corrections for path fluctuations
  - See CASA Cookbook section 4.3.7 for more info
- T_{sys}
  - ALMA records raw correlation coefficients
  - Multiply by T_{sys} to convert to Kelvins
  - Later multiply by antenna gain (Jy/K) to convert to correlated flux density
  - CASA Cookbook 4.3.1 for more info
- Antenna Positions
  - Small corrections to antenna positions included in raw data

Calibration Process

- Next Correct Time and Frequency Dependent Fluctuations
  - Bandpass
  - Temporal phase and amplitude (gain)
- Types of Calibrator Sources
  - Point sources offer simple phase and amplitude behavior (constant w/ UV distance)
  - Bandpass calibrator should be bright for adequate S/N
  - Phase calibrator should be close to science target to minimize differences in atmospheric effects
  - Flux calibrator needs well understood flux densities

Calibration Process

- Bandpass
  - First solve for (additive) temporal phase corrections
  - Usually fine solution interval (integration)
  - Then derive phase and (multiplicative) amplitude corrections as function of frequency
  - Corrects antenna-based frequency responses
  - All derived from bandpass calibrator target
- Gain
  - Solve for short interval temporal phase corrections on all calibrators
  - Then long interval (scan) amplitude corrections
  - Derive flux scaling relation
  - Long interval phase corrections (for transfer to science)

Calibration Process

- A Bit More on Flux Scaling
  - After applying short interval gain corrections data is now in units of flux density
  - Flux calibrator has known flux density (model)
  - Compare observed flux density to model
  - Derive solutions that scale observed flux density to match the model
  - Assume antenna gain does not change between targets
  - Apply same scaling corrections to other target flux densities

Calibration Process

- Once All Calibration Applied, Inspect Bandpass, Flux and Phase Calibrators
  - Look at visibility amplitude and phase vs. time and frequency
  - Amplitudes should be centered around flux scaled values
  - Phase should be centered around zero (calibrator at center of field of view)
- Flag Any Misbehaving Antennas, Scans, Integrations etc.
- Re-Derive and Re-Apply Calibration Solutions
- Inspect Data
- Rinse, Repeat Until Satisfied