



Using ALMA Archive

ALMA Science Archive substructure and operation

Erik Muller,
Ken'ichi Tatematsu,
Daisuke Iono.

EA-ARC ALMA user support
(NAOJ Chile Observatory)



ALMA Science Archive - ASA:

Synopsis



1. ASA: Motivated by Scientists
2. Structure of the Archive facility
3. Obtaining data from the archive
4. Structure of Archived data.
5. PI publication requirements
6. Future development.

What do scientists want?

1. Reduced, science-ready, high-level data (PIs want proprietary protection)
2. Processible, accessible data
3. Data they can reprocess/analyse with their “pet” data analysis software

Solution:

1. Provide processed, calibrated data, with clear description of the processing
2. Provide general format data
3. Provide raw data for the PI to *completely* re-process, if desired.

ALMA science Archive caters for TWO “clients”:

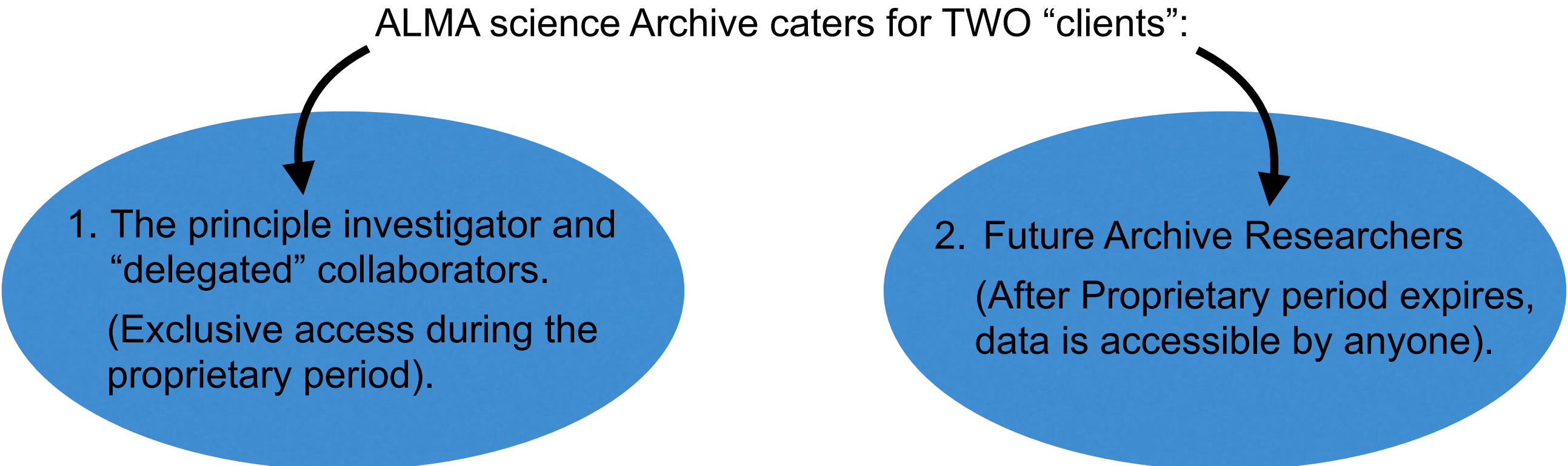
What do scientists want?

1. Reduced, science-ready, high-level data (PIs want proprietary protection)
2. Processible, accessible data
3. Data they can reprocess/analyse with their “pet” data analysis software

Solution:

1. Provide processed, calibrated data, with clear description of the processing
2. Provide general format data
3. Provide raw data for the PI to *completely* re-process, if desired.

ALMA science Archive caters for TWO “clients”:



1. The principle investigator and “delegated” collaborators.
(Exclusive access during the proprietary period).

2. Future Archive Researchers
(After Proprietary period expires, data is accessible by anyone).



Processing staff



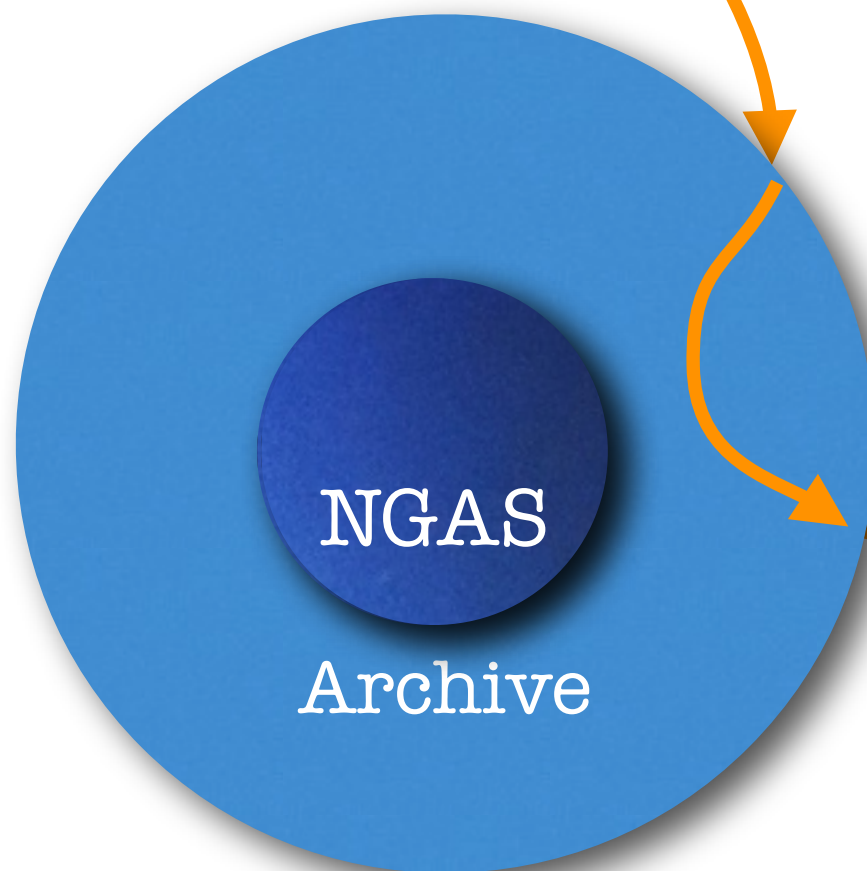


Proposal submission





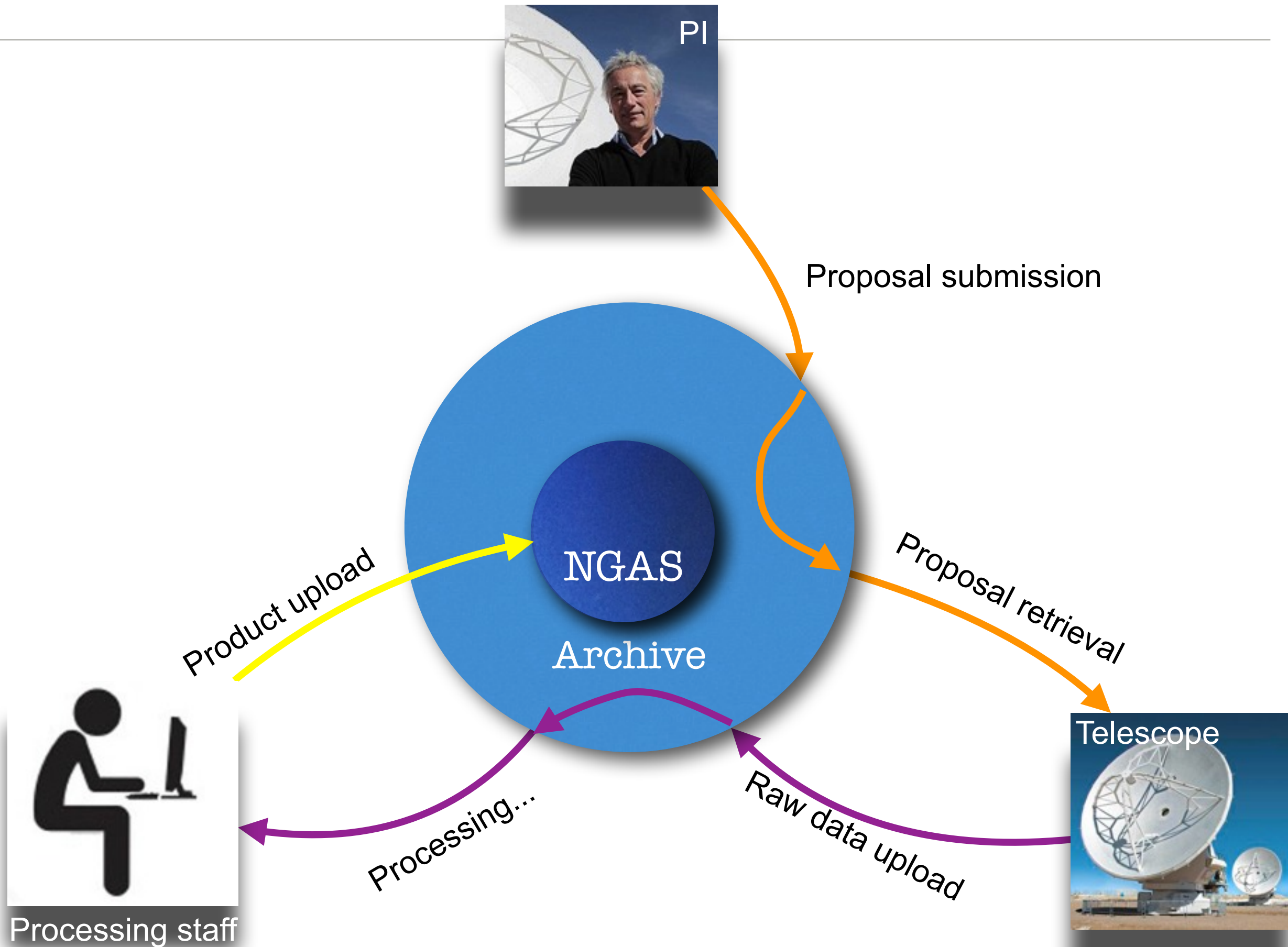
Proposal submission

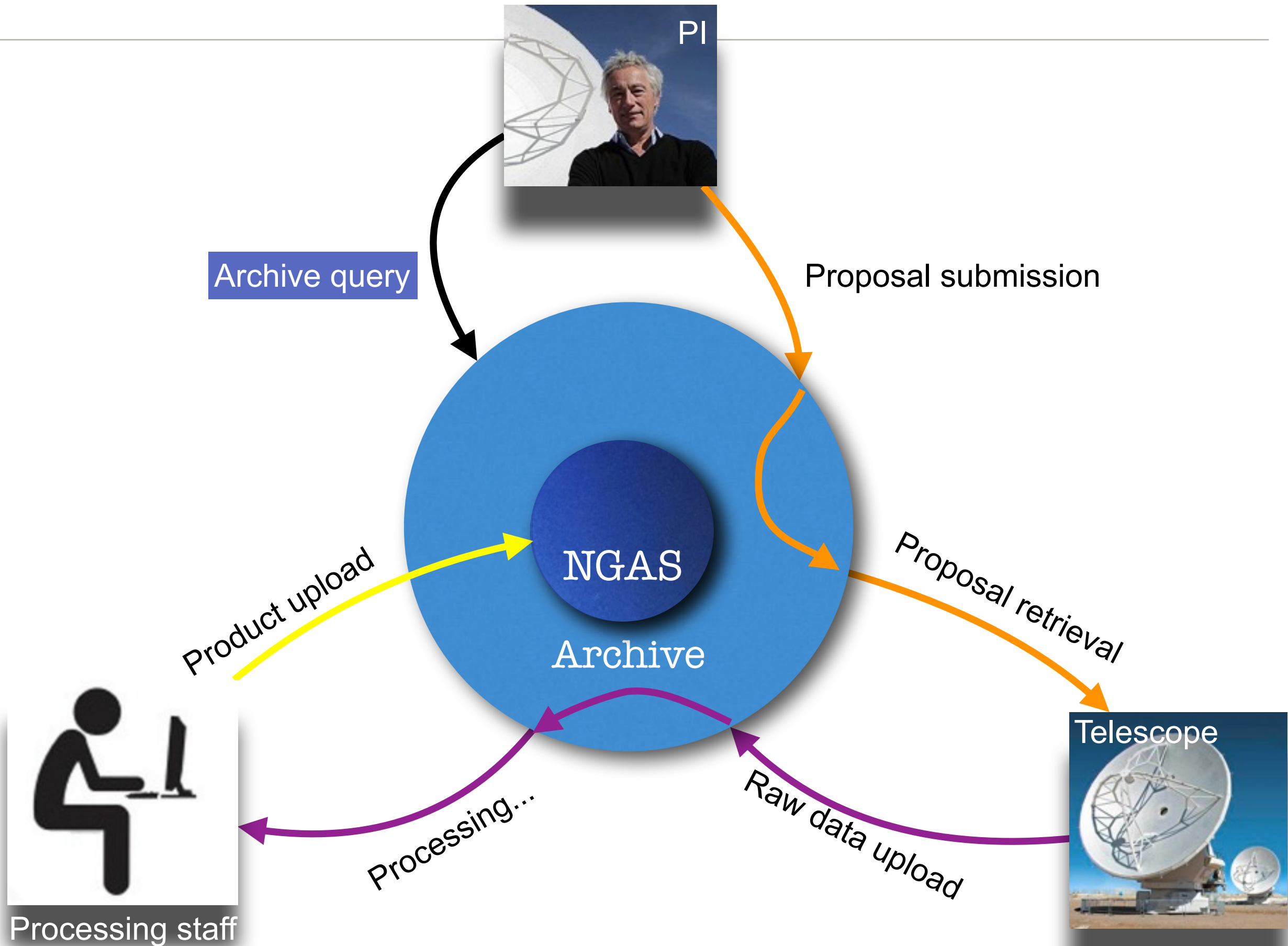


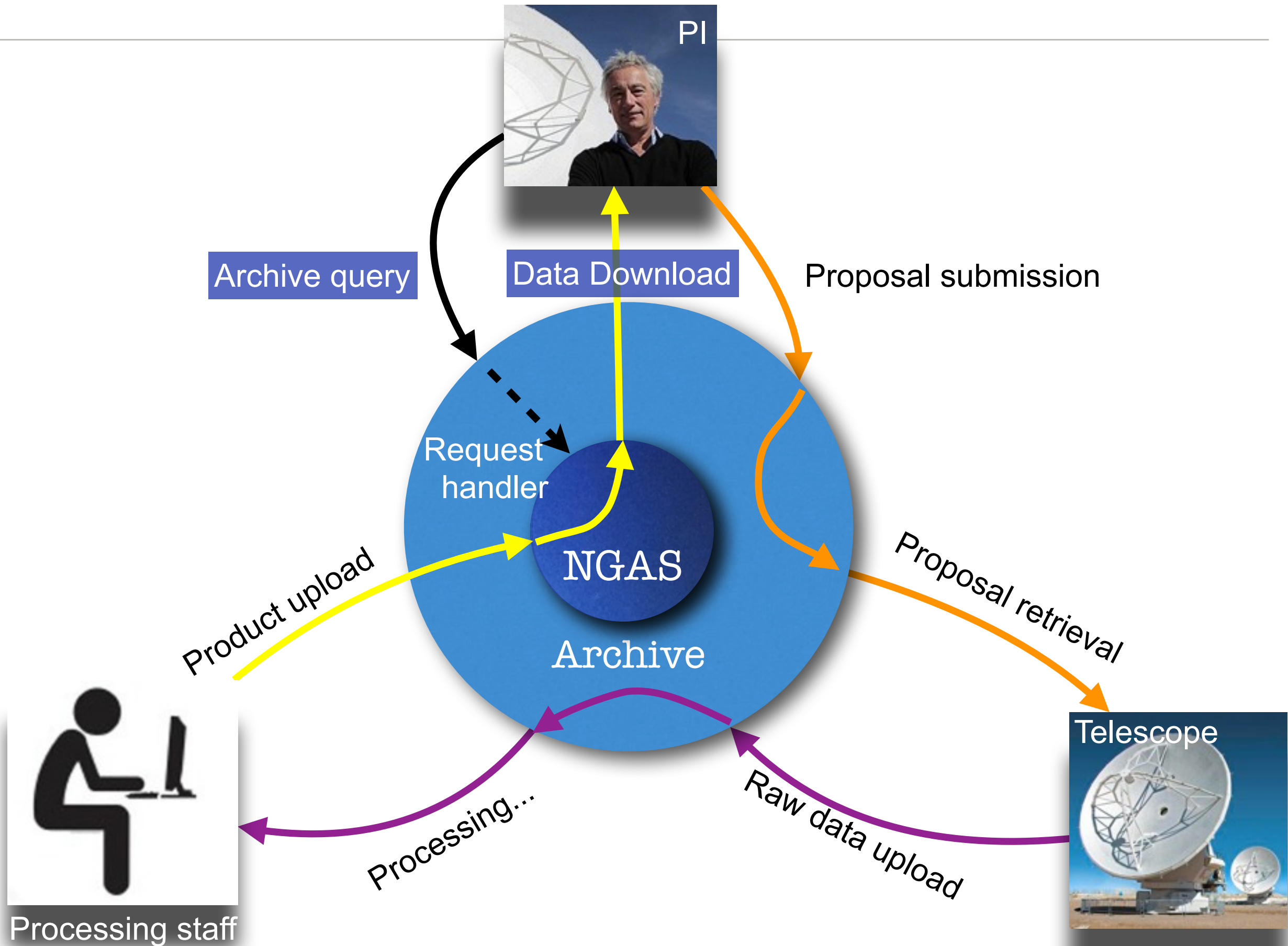
Proposal retrieval



Processing staff







Current status

- Presently, EA is delivering data primarily by FTP transfer.
(Resulting in a shorter delivery time, after staff-processing stage is complete)
- We intend to migrate to delivery by Archive internet data distribution, in the near future (~3-5 weeks).
- Some deliveries are already being made using Archive internet data-distribution.
- Delivery is currently effected through one or more 'tarballs'.
- Raw data ('ASDMs') not 'delivered' by default in cycle 1+ data, but are available directly by internet request.
- Future researchers can optimise pipeline processing for their science

ALMA Science Archive - PI delivery:

PI notice

- PI is currently notified by email, using ALMA helpdesk.
- Email will contain instructions on how to obtain your data, one of two methods:
 - **1. by FTP (File transfer Protocol), from EA-ARC firewall-protected repository**
 - 2. from the Archive data distribution system, using your ALMA login.

PI - Notification for Delivery by FTP transfer

Dear Albert Einstein,

We are pleased to announce that data of the Scheduling Blocks "alpha_cen_12m" containing execution block uid__A00_X123_X456 from your project 2012.1.00111.S with the title

"Alpha centauri and all that is beautiful"

are now available to you for download. A tarball of these data can be obtained from

(1) Using web browser:

FTP URL: ftp://

(2) Using FTP-specific client software:

Host: alma-dl.mtk.nao.ac.jp

User: [user]

Path: [path]

Only passive-mode FTP is available.

The password will be sent to you via separate email. Please reply to this ticket if you do not receive a password within 1 hour.

The account will expire at 19th april,16:00 JST

.

ALMA Science Archive - PI delivery:

PI notice

- PI is currently notified by email, using ALMA helpdesk.
- Email will contain instructions on how to obtain your data, one of two methods:
 - 1. by FTP (File transfer Protocol), from EA-ARC firewall-protected repository
 - **2. from the Archive data distribution system, using your ALMA login.**

PI - Notification for Delivery by Archive data distribution

Dear Albert Einstein,

We are pleased to announce that data of the Scheduling Blocks "alpha_cen_12m" containing execution block uid___A00_X123_X456 from your project 2012.1.00111.S with the title

"Alpha centauri and all that is beautiful"

are now available to you for download. A tarball of these data can be obtained from the ALMA Science Archive.

Please use your ALMA userid "username" to login to the ALMA Science Portal,
<http://almascience.nao.ac.jp>

and follow the "Data" -> " link to the archive to search for and download your data, or access to

http://almascience.nao.ac.jp/aq/asdm_uid= uid___A00_X123_X456

A download request of both processed and raw data can be submitted by selecting all data (topmost checkbox) and hitting "submit download request".

.
.
.

Delivery content for Cycle 1 deliveries + onwards.

A blue folder icon with the text "FITS" in white.

FITS

- Data products (FITS cube)

A blue folder icon with the text "LOGS" in white.

LOGS

- Observing Logs

A blue folder icon with the text "Scripts" in white.

Scripts

- Calibration, processing + Imaging Scripts

A blue folder icon with the text "Cal. Tables" in white.

Cal. Tables

- Calibration tables

A blue folder icon with the text "QA rept." in white.

QA rept.

- Quality Assessment (QA0, QA2) Reports

A yellow document icon with the text "README" in black.

README

- README file (text file describing contents of delivery)

A large blue folder icon with the text "ALMA" in white and a red "TOP SECRET" stamp.

ALMA

TOP SECRET



Delivery content for Cycle 1 deliveries + onwards.

- Data products (FITS cube)
- Observing Logs
- Calibration, processing + Imaging Scripts
- Calibration tables
- Quality Assessment (QA0, QA2) Reports
- README file (text file describing contents of delivery)

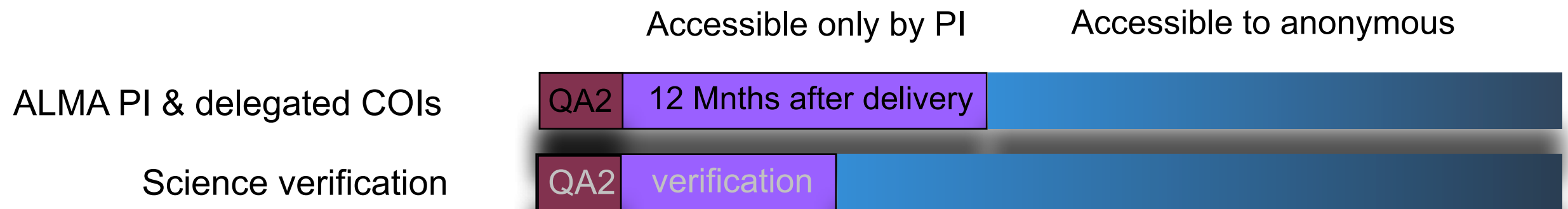


Delivery to PI via network



Proprietary rules:

- PI Data is available for public use after 12 calendar months (6 months for director discretionary time)
- Science verification data available after verification



Proprietary-expired data are all available through the Science portal.



Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Access Alma science portal:
<http://almascience.nao.ac.jp>

Navigate to
DATA -> Archive

Choose **Archive query**

Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Access Alma science portal:
<http://almascience.nao.ac.jp>

Navigate to
DATA -> Archive

Choose **Archive query**

1. Enter search parameters
2. Click Search.
3. Refine parameters

Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Access Alma science portal:
<http://almascience.nao.ac.jp>

Navigate to
DATA -> Archive

Choose **Archive query**

1. Enter search parameters
2. Click Search.
3. Refine parameters

Specify file selection in
Request Handler

Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Access Alma science portal:
<http://almascience.nao.ac.jp>

Navigate to
DATA -> Archive

Choose **Archive query**

1. Enter search parameters
2. Click Search.
3. Refine parameters

Specify file selection in
Request Handler

Download with
provided wget scripts

Download from
internet

Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Access Alma science portal:
<http://almascience.nao.ac.jp>

Navigate to
DATA -> Archive

Choose **Archive query**

Choose project download
(**'delivery list'**)

1. Enter search parameters
2. Click Search.
3. Refine parameters

Specify file selection in
Request Handler

Download with
provided wget scripts

Download from
internet

Getting proprietary-expired data:

1. Submit query on 'Archive query'
OR 'delivery list'
2. Refine query and request data
on Request handler
3. Download.

Access Alma science portal:
<http://almascience.nao.ac.jp>

Navigate to
DATA -> Archive

Choose **Archive query**

Choose project download
(**'delivery list'**)

1. Enter search parameters
2. Click Search.
3. Refine parameters

1. Find target project,
2. Select with checkbox.
3. Click 'Request marked deliveries'

Specify file selection in
Request Handler

Download with
provided wget scripts

Download from
internet



ALMA Science Archive - Archive research



Archive Query:

When using JAVA 7.51+: security settings should be modified.



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

Position

Source name (Sesame)

Source name (ALMA)

RA Dec

Search radius

0:10:00

Energy

Frequency

Bandwidth

Spectral resolution

Band

Time

Observation date

Integration time

Polarisation

Polarisation type

Observation

Water vapour

Scan intent

Observe target

Project

Project code

Project Title

PI Name

Options

Results view

☒ raw data ☐ project

- Enables 'standard' search patterns
 - target, position, band. Observing date & conditions. Also: project code, PI name
- Enables advanced filtering capabilities
 - Limit search by linking logical filter parameters (band, sensitivity, etc. etc.)
- Download links are returned to user.



ALMA Science Archive - Archive research Request Handler



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

Request Handler

Login

Archive Requests Req #85,290,726

Request #85290726 by Anonymous User ✓

ALMA

Select All

Unselect All

Download Selected

Requested Projects / OUSets / Executionblocks

UNIX/Linux Shell script if you prefer command line download of complete request: [downloadRequest85290726script.sh](#)

Data entities 1-1 of 1

Select	Project / OUSet / Executionblock	File	Size	Access
<input type="checkbox"/>	<input type="checkbox"/> uid__A002_X305ca0_X792			
<input type="checkbox"/>		2011.0.00099.S 2012-02-24 001 of 004.tar	19.1GB	✓
<input type="checkbox"/>		2011.0.00099.S 2012-02-24 002 of 004.tar	16.8GB	✓
<input type="checkbox"/>		2011.0.00099.S 2012-02-24 003 of 004.tar	19.5GB	✓
<input type="checkbox"/>		2011.0.00099.S 2012-02-24 004 of 004.tar	13.2GB	✓
Data entities 1-1 of 1			68.7GB	

- Enables selection of specific tarballs (note, duplicity exists in some cases, for cycle 0 data)
- Provides a text linux script invoking **wget** commands, which may be run from the terminal.



Cycle 0 - vs - cycle 1

- Cycle 0 data includes the raw data.
(in an ASDM format file e.g. uis____X123_X123_X12.asdm.sdm.tar)
- Cycle 0 data also contains the 'measurement set' data, which is calibrated
- Cycle 0 data also contains the products in FITS format
- Cycle 1 data tarballs to NOT include the raw data, nor the calibrated measurement set.
- Cycle 1 data tarballs include the products, in FITS format.

Request #85314508 by Anonymous User ✓

ALMA

[Requested Projects / OUSets / Executionblocks](#)

UNIX/Linux Shell script if you prefer command line download of complete request: [downloadRequest85314508script.sh](#)

Data entities 1-1 of 1

Select	Project / OUSet / Executionblock	File	Size	Access
<input type="checkbox"/>	uid__A001_Xa0_X17e5			
		2012.1.00941.S_uid__A002_X5d7935_X1cb_001_of_001.tar	85.9MB	⊘
		2012.1.00941.S_uid__A002_X5d7935_X1db_001_of_001.tar	243.8MB	⊘
		uid__A002_X67f61e_X3a4.asdm.sdm.tar	1.2GB	⊘
		uid__A002_X67f61e_X4c2.asdm.sdm.tar	1.2GB	⊘
		uid__A002_X67f61e_X57a.asdm.sdm.tar	1.2GB	⊘
		uid__A002_X6a1245_X976.asdm.sdm.tar	1.2GB	⊘
		uid__A002_X7330a2_X11b5.asdm.sdm.tar	1.9GB	⊘
Data entities 1-1 of 1			6.9GB	

ALMA Science Archive - Archive research

What next? for PIs and Archive researchers.

- You will receive one, or many tarballs - untar these (macs can't do this properly from a on-screen click. Untar the folders from the command line)
- Examine the README file, it contains information about the contents of the tarballs
- Examine the PRODUCTS directory, it contains FITS-format images of the target frequencies (or continuum), made by the PI

Atacama Large Millimeter/submillimeter Array (ALMA)

#####

Cycle: 0 (Early science)

Project code: 2011.0.00199.SProject title: Bursting Water Maser Feature in Orion

KL

Configuration: Extended

Proposed rms: 170mJy

CASA version used for reduction: 3.4

Comments from Reducer:

DA42, DV01, DV07, and DV23 were removed because of large pointing error.

DV17 was removed due to negative Trec.

DV19 XX pol. was removed due to out of spec Trec temperature and high Tsys.

DA51 was removed due to bad bandpass gain solutions.

DV05 was removed due to poor S/N in bandpass.

The rms was achieved to be 105 mJy (the required rms is 140 mJy) by the integration in 7.5 GHz in frequency width but the image may contain non-real detection.

WVR correction seems to make phase worse in some antennas. It is probably better not to use them in those antennas.

emuller@alma-work:~/2011.0.00199.S\$ more `find -name README`

Atacama Large Millimeter/submillimeter Array (ALMA)

(C0, C1+) README contains

1. Basic information about the data.
2. Notes from the data reducer: flagging, high tsys, etc.
3. A directory tree

ALMA Science Archive - Archive research

What next? for PIs and Archive researchers.

- You will receive one, or many tarballs - untar these (macs can't do this properly from a on-screen click. Untar the folders from the command line)
- Examine the README file, it contains information about the contents of the tarballs
- Examine the PRODUCTS directory, it contains FITS-format images of the target frequencies (or continuum), made by the PI

```
-- 2011.0.00199.S/  
| |-- sg_ouss_id/  
| | |-- group_ouss_id/  
| | | |-- member_ouss_2013-03-01_id/  
| | | | |-- README  
| | | | |-- product/  
| | | | | |-- uid___A002_X590461_X620.ms.split.image.continuum.source4.image.fits  
| | | | | |-- uid___A002_X590461_X620.ms.split.image.continuum.source4.mask/  
| | | | |-- calibration/  
| | | | | |-- uid___A002_X590461_X620.calibration.plots/  
| | | | | |-- uid___A002_X590461_X620.calibration/  
| | | | |-- log/  
| | | | | |-- uid___A002_X590461_X620.calibration.log  
| | | | |-- raw/  
| | | | | |-- uid___A002_X590461_X620.ms.split/  
| | | | |-- calibrated/  
| | | | | |-- uid___A002_X590461_X620.ms.split.cal/  
| | | | |-- qa/  
| | | | | |-- uid___A002_X590461_X620__qa2_part1.png  
| | | | | |-- uid___A002_X590461_X620__qa2_part2.png  
| | | | | |-- uid___A002_X590461_X620__textfile.txt  
| | | | | |-- uid___A002_X590461_X620__qa2_part3.png  
| | | | |-- script/  
| | | | | |-- uid___A002_X590461_X620.ms.scriptForCalibration.py  
| | | | | |-- uid___A002_X590461_X620__scriptForImaging.py
```

(C0, C1+) README contains

1. Basic information about the data.
2. Notes from the data reducer: flagging, high tsys, etc.
3. A directory tree



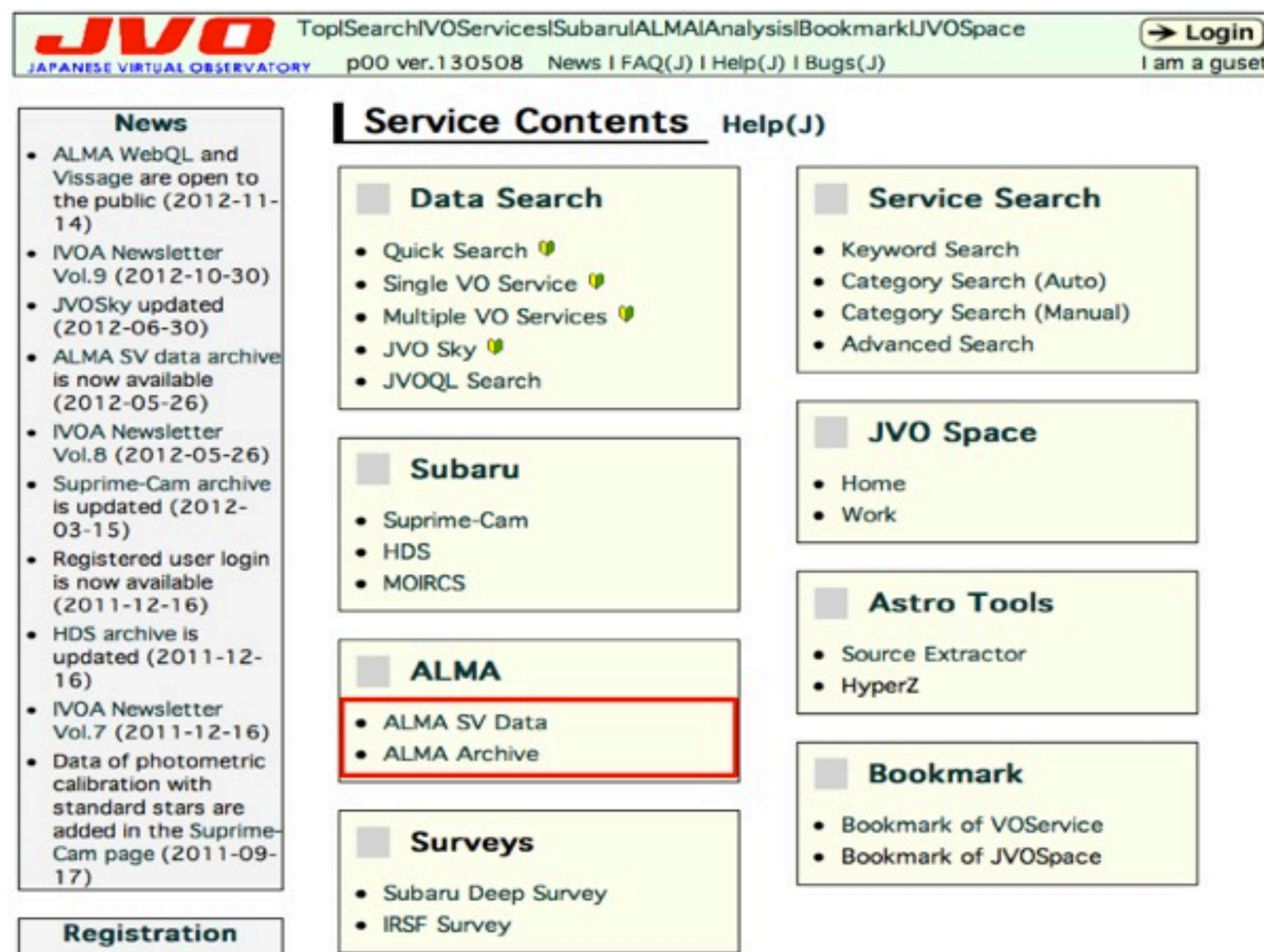
How to make cycle 1 measurement sets (refer to README information):

- Cycle 1 data includes the CASA processing script, which runs on the raw data (asdm format)
- PIs and Archive users can obtain the raw data from the archive, put it in the 'calibrated' directory, and run the script (or adjust it)

Why do we do this?

- Not all PIs/ARs need the measurement set, it is very bulky (~100's GB), and slows transfer
- Not all PIs/ARs need the raw data, but it must be available.

locally-stored data query interface



ALMA Archive : Target Info

Target Name: R Scl

#	dataset id	ra/dec (J2000)	size (arcmin ²)	band	freq. range (GHz)	data type	Cube size (XxYxF) ?	image resol (arcsec)	freq. resol (MHz)	obs date	original fits name
1	ALMA01000002	01h26m58.1-32d32m35	2.16x2.16	Band7	330.232 -- 333.962	intensity cube	432x432x1	0.30	3,729.650	2011-10-03	RScI_continuum.fits
2	ALMA01000003	01h26m58.1-32d32m35	2.16x2.16	Band7	345.773 -- 345.854	intensity cube	432x432x141	0.30	.577	2011-10-03	RScI_line1.fits
3	ALMA01000004	01h26m58.1-32d32m35	2.16x2.16	Band7	345.328 -- 345.398	intensity cube	432x432x121	0.30	.576	2011-10-03	RScI_line2.fits
4	ALMA01000005	01h26m58.1-32d32m35	2.16x2.16	Band7	342.871 -- 342.940	intensity cube	432x432x121	0.30	.572	2011-10-03	RScI_line3.fits

ALMA Science Archive & JVO:

http://alma-intweb.mtk.nao.ac.jp/~kt/ALMA_archive_handbook.pdf


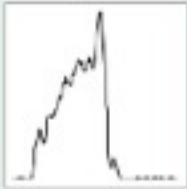
Purpose-built data visualisation with “Vissage”

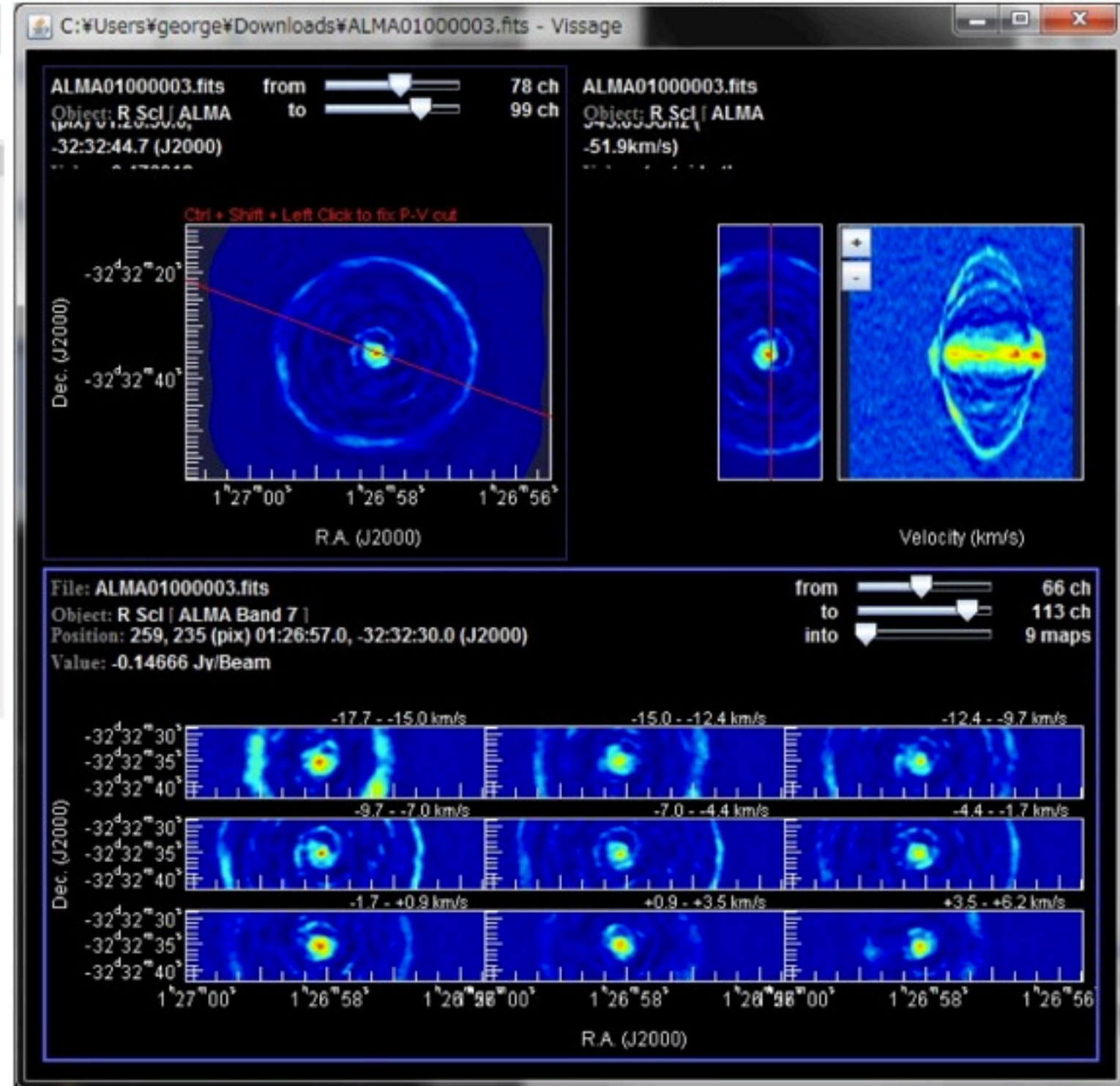
JVO TopSearch/IVOServices/Subaru/ALMA/Analysis/Bookmark/JVOSpace
JAPANESE VIRTUAL OBSERVATORY p00 ver.130508 News / FAQ(J) / Help(J) / Bugs(J)
=> Location: Top Page > ALMA > Archive > Target Info > Dataset Info

ALMA Archive : Dataset Info

Summary Binning Data Desktop Viewer Using the data

Target R Scl	Dataset ID ALMA01000003
Coord. (RA/DEC J2000) 01h26m58.1-32d32m35	Date of Observations 2011-10-03
Image Size (arcmin2) 2.16x2.16	Image Resol. (arcsec) 0.30
Band Name Band7	Data Type Intensity cube
Freq. Range. (MHz) 345,772.633 -- 345,853.951	Spectrum Resol. (MHz) .577
Cube Pix ? 432x432x141	Original Filename RScI_line1.fits

data id	image	spect	file size (byte)	Download	Web QL
ALMA01000003			105,301,440	Download	Web QL



JVO

Vissage



ALMA Science Archive - ASA:

Required acknowledgement information.



Please notify EA-ARC and EA-ARC outreach for any publications making use of ALMA data.

Publications making use of ALMA data (PI and Archive researchers) are required to include the following acknowledgements in publications:

“This paper makes use of the following ALMA data: ADS/JAO.ALMA#2011.0.01234.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.”

Immediate Future improvements

1. Basic layout to be modified slightly
2. 'Frequency support' information to be improved
3. File Download destination support improved
4. Filtering improved
5. online help improved.
6. Visual quick-look planned.

Not happy with the Archive functionality?

Or are you having problems?

YOU are the users, but we need your input to make the system better for everyone.
We can only improve the archive if you tell us about problems, or any suggestions.

If you have any comments or suggestions, please submit them to EA-ARC using the helpdesk system (<https://help.almascience.org>).

Please also ensure you examine the 'help' information on both the Archive query and Request handler pages.

Please contact EA-ARC for face-to-face support information