

# Current status of JVO ALMA FITS Archive

Yuji Shirasaki (NAOJ/ADC)

# About JVO

- Japanese Virtual Observatory (JVO) Project

- Develop a data search system using the Virtual Observatory (VO) standard.

- JVO portal

- provides seamless access to the VO service of the world,
- reduced data of ALMA, Nobeyama, and Subaru
- Gaia source catalog EDR3
- <http://jvo.nao.ac.jp/portal>



Japanese Virtual Observatory  
(Astronomy Data Center)  
National Astronomical Observatory of Japan

## JVO ALMA / Nobeyama FITS archive

VO Search	Subaru Reduced Data Archive	ALMA FITS Archive	Nobeyama FITS Archive	Gaia Catalog	Other resources
<b>Summary</b>	<b>Summary</b>	<b>Summary</b>	<b>Summary</b>	<b>Summary</b>	<b>Summary</b>
JVO portal VO search page version 2. You can access to the VO Crawler DB at this page or JVO Sky. Older version of JVO portal (version 1) is linked from here.	Download the reduced Subaru data. Suprime-Cam, MOIRCS, HDS.	Search, View, and Download the ALMA data cube in FITS format. SV data, Archive, WebQL demo.	Search, View, and Download the Nobeyama data cube in FITS format. FITS, COMING, StarFormation.	Gaia related source catalogs. Gaia DR1, Gaia DR1 TGAS, Gaia DR2	AKARI FIS Image   Subaru Deep Survey Catalog   IRSF LMC/SMC/BC Survey

NEWS

- 2019-10-04 Data of NBO 45m Legacy Project COMING was updated. New data for 15 galaxies were added, and data for 19 galaxies were updated.
- 2019-09-15 Search JVE for VO Crawler DB is available at VO Crawler DB.
- 2019-04-25 FITS WebQL button was implemented on the VO search result page. You can look at the FITS images found by the VO search interface using FITS WebQL. Try out MultiScope etc.
- 2018-10-17 FITS WebQL v2 (Beta) was released. New feature "FITS Cube slicer" is available.
- 2018-08-07 Gaia DR2 is now available at JVO portal.
- 2018-06-07 The data of Nobeyama 45-m Legacy projects was released.
- 2017-10-19 VO Search update: new VO search interface named JVOIndex and JVOExplorer are open to the public.
- 2017-09-08 JVO ALMA FITS archive update: the BETA version 3 (renamed FITSWebQL) of the interactive ALMA WebQL is now available.
- 2017-06-26 JVO ALMA FITS archive update: Advanced Search GUI was implemented. You can search the data by specifying various criteria.
- 2017-06-26 JVO ALMA FITS archive update: Functionality for filtering the projects list by their science category was implemented.
- 2017-06-12 Subaru Suprime-Cam mosaic images were reprocessed with the most recent reduction pipeline for which mosaic success rate was improved. They are available at: JVO Suprime-Cam mosaic image archive.
- 2017-03-13 JVO portal top page and the VO search GUI was updated.
- 2017-03-08 Subaru WebQL experimental version is available at data download page of JVO Suprime-Cam mosaic image archive. Try it with a sample image
- 2016-10-15 ALMA WebQL v2 was updated.
- 2016-10-15 Gaia source catalog is now available at JVO portal.
- 2016-04-10 Subaru Suprime-Cam Archive was updated. All the data were reprocessed with the most recent data reduction pipeline.

JVO Help Desk  
help\_desk@jvo.nao.ac.jp

# JVO ALMA FITS Archive

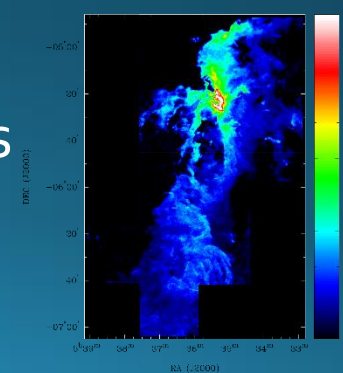
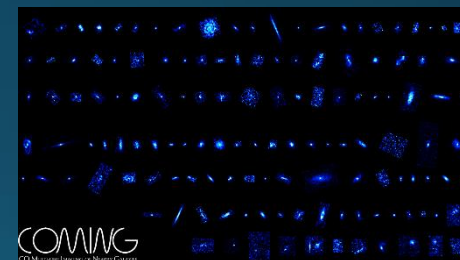
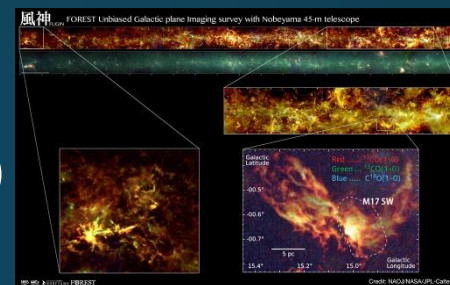


- Major characteristics
  - data search per file basis
  - VO search functionality
  - FITS WebQL: interactive quick look viewer for image and spectrum
  - Only FITS files are available at the JVO site
  - 700k FITS files (220 TB) are available
- The FITS data at the JVO site
  - are copies of the original FITS data in the ALMA science archive
- Raw data are only available at the ALMA science archive

# JVO Nobeyama FITS archive

Calibrated FITS data cube of Legacy projects:

- **FUGIN** (PI: Umemoto & Minamidani)  
CO Galactic Plane Survey (1<sup>st</sup> & 3<sup>rd</sup> quadrant)  
 $l = 10^\circ \sim 50^\circ$ ,  $l = 198^\circ = 236^\circ$ ,  $|b| \leq 1^\circ$   
(~200 GB)
- **COMING** (PI: Sorai)  
CO Multi-line Imaging of Nearby Galaxies  
147 galaxies (~1.5 GB)
- **Star Formation** (PI: Nakamura)  
Wide-field mapping of nearby molecular clouds  
Aquila Rift, M17, Orion A (~12 GB)



# Status of JVO ALMA FITS archive in 2021

- **ALMA line search service (experimental)**
  - Line detection was carried out on the ALMA public FITS cube.
  - Information on the detected lines were registered into JVO ALMA line DB.
  - Matching with the molecular lines.
- **FITS WebQL v5 (experimental)**
  - Alpha release of v5 with distributed computing is publicly available.
  - Currently four servers are running behind the scenes.
  - Comfortable response even to 300 GB FITS file
- **Accident loss of data in the last month**
  - Half of the ALMA FITS data were lost from the JVO storage system, which is caused by RAID misconfiguration.
  - Most of the data have been recovered.

# ALMA line search service

Target Name Simbad Name Project Code Coords Frequency **Line** Advanced Download Change Log

(un)check all  disable help

# of lines per spectrum  $\geq$   &  $<$    
 f\_LSRK [GHz] =   $\pm$

(\*) Please note that search by the rest-frame frequency properly works for data which contains enough information for determining the source velocity (or redshift). We rely on the FITS keyword "RESTFRQ" which is assumed to be set to the middle of the frequency range in the source rest-frame.

f\_rest [GHz] =   $\pm$

(\*) Please note that "search by the species" below is provided for searching data in which emission or absorption lines are detected in the frequency range expected for given species. This doesn't guaranty that the detected lines are originated from the specified species. The result relies on the accuracy of the FITS keyword "RESTFRQ".

species  include all the checked (and input) species

(popular)  13CN  13CO  13CS  C17O  C18O  C34S  CCH  CN  CO  CS  DCN  DCO  H13CN  H13CO  H2CO  HC3N  HCN  HCO  HDO  HN13C  HNC  N2H  SiO  SO

(diatomic)  AlCl  AlF  AlO  CF  CH  CN  CO  CP  CS  FeO  HD  HCl  HF  OH  KCl  NH  O2  PN  PO  SH  SO  SiC  SiN  SiO  SiS  TiO

(Triatomic)  AlNC  CCH  C2S  C2P  H3  CH2  H2Cl  H2O  H2S  HCN  HCO  HCP  HCS  KCN  MgCN  NH2  N2O  NaCN  OCS  SO2  SiC2  SiCN  TiO2

(Four atoms)  l-C3H  C3N  H3O  H2CN  H2CO  H2CS  HCCN  HOCO  HCNO  HOCN  HNCO  HNCS  NH3  HNCS  c-SiC3  HMgNC

(Five atoms)  C3H2  H2C3  CH2CN  H2C2O  H2COH  C4H  HC3N  HCOOH  NH2CN  CNCHO  SiC4

(Six atoms)  H2C3O  HNCHCN  CH3CN  CH3OH  CH3SH  HC3NH  NH2CHO  C5H  C5N  HC2CHO  HC4N  H2CCNH  C5S

(Seven atoms)  H2COCH2  CH3CCH  CH3NH2  CH2CHCN  H2CHCOH  C6H  HC5O  CH3CHO  CH3NCO

(Eight atoms)  HCCCH2CN  CH2OHCHO  CH3OCHO  CH2CHCHO  H2CCCHCN  C7H  H2NCH2CN

(Nine atoms)  CH3C4H  CH3OCH3  CH3CH2CN  CH3CH2OH  C8H  HC7N  c-C6H5CN

(free text) =

match score  $\geq$    
 match fraction  $\geq$    
 offset [km/s]  $<$    
 offset sigma [km/s]  $<$

## Search criteria

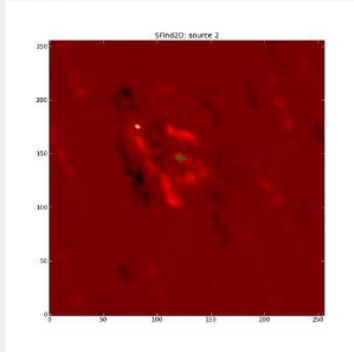
- # of lines per spectrum
- Frequency of detected line peak (LSRK, Rest)
- Molecular species
- ...

# ALMA line search service

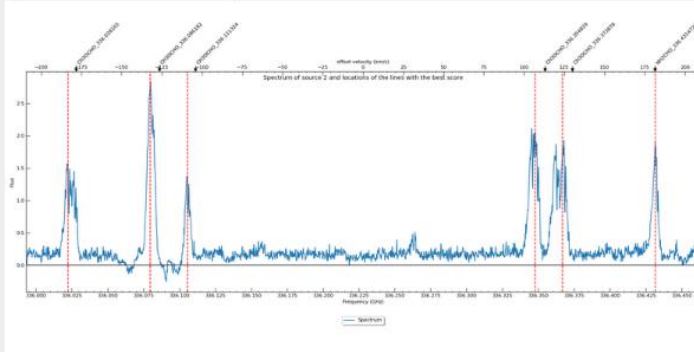
## Source 2 spectrum

src ID	ra	dec	X	Y	SN
2	05:35:14.181	-05:22:34.20	121.24	146.28	37.9

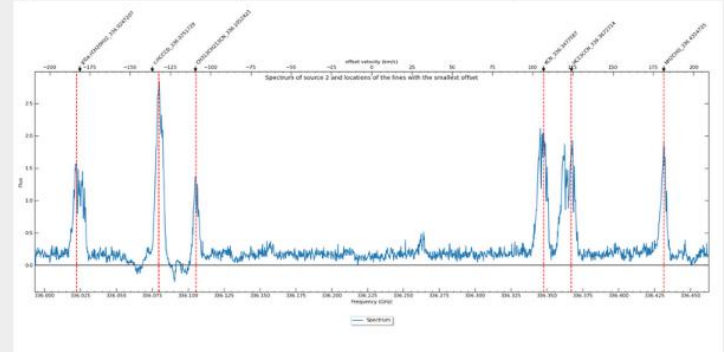
Integrated image and source 2



Spectrum and transition frequencies selected based on match-score



Spectrum and transition frequencies selected based on velocity offset

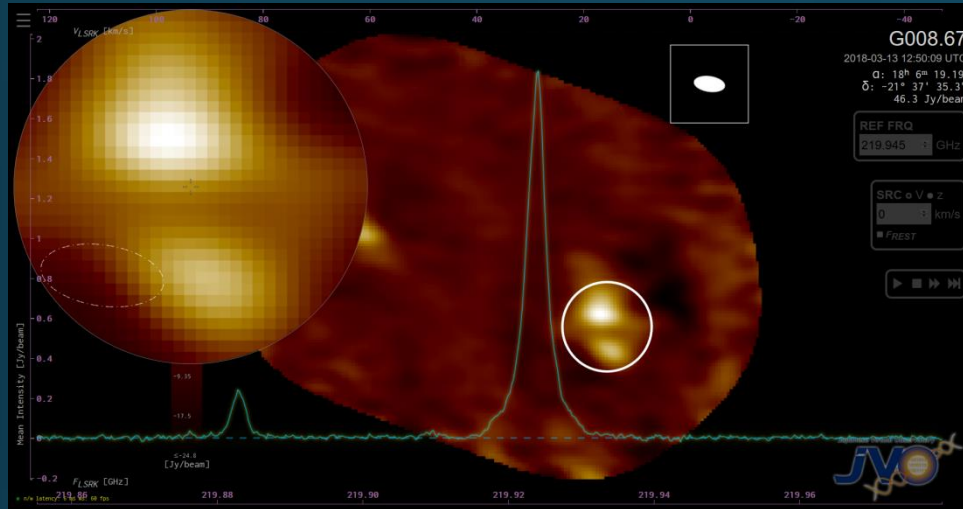


Frequency of detected lines and candidate species

src ID	line ID	peak frequency (LSRK) [GHz]	peak frequency (Rest) [GHz]	species1 name ?	species1 frequency [GHz]	species1 score ?	species1 fraction ?	species1 offset velocity ? [km/s]	species2 name ?	species2 frequency [GHz]	species2 offset velocity ? [km/s]
2	0	336.01800	336.02248	CH3OCHO	336.02817	8	1.0	5.1	g'Ga-(CH2OH)2	336.02472	2.0
2	3	336.07600	336.07961	CH3OCHO	336.08618	8	1.0	5.9	c-HCCCD	336.07517	-4.0
2	6	336.10100	336.10549	CH3OCHO	336.11132	8	1.0	5.2	CH313CH213CN	336.10524	-0.2
2	9	336.34400	336.34769	CH3OCHO	336.35483	8	1.0	6.4	KCN	336.34776	0.1
2	14	336.36300	336.36697	CH3OCHO	336.37388	8	1.0	6.2	HC13CCN	336.36727	0.3
2	15	336.42700	336.43143	NH2CHO	336.43147	3	0.7	0.0	NH2CHO	336.43147	0.0

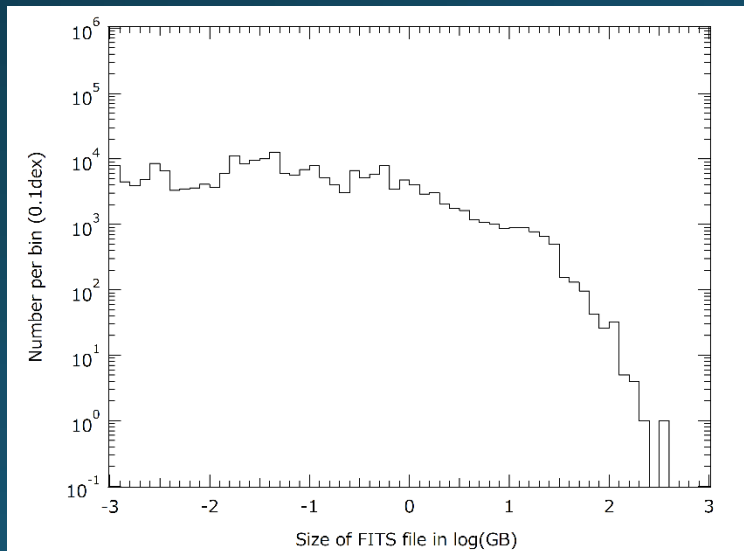
- Spectrum at the location of peak intensity in the image
- Detected emission / absorption lines are indicated by red lines
- Candidates of molecular species
- **Caution!**: conversion to REST frame relies on the RESTFREQ in the FITS header. → check if velocity (redshift) is correctly derived!

# Status of FITS Web QL



- FITS Cube viewer running on the Web browser. Display image and spectrum for a selected region.
- The size of the FITS file is getting larger and larger. Current max is 330 GB.
- file size will grows to ~1 TB.

Distribution of FITS file size (public).

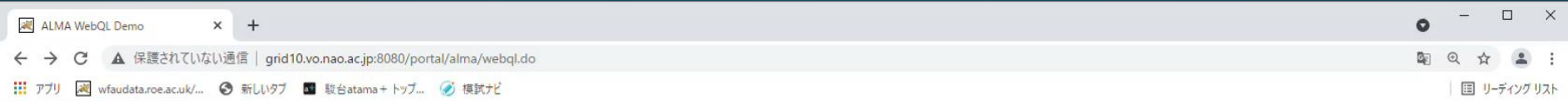


## Solutions

1. Upgrade to a server with larger memory (768GB). → high-cost, still have a limit in memory size.
2. Adapt distributed computing with consumer PCs. → low-cost, scalable by adding PCs



# FITS WebQL v5 (distributed computing version) Demo



## ALMA WebQL Demo

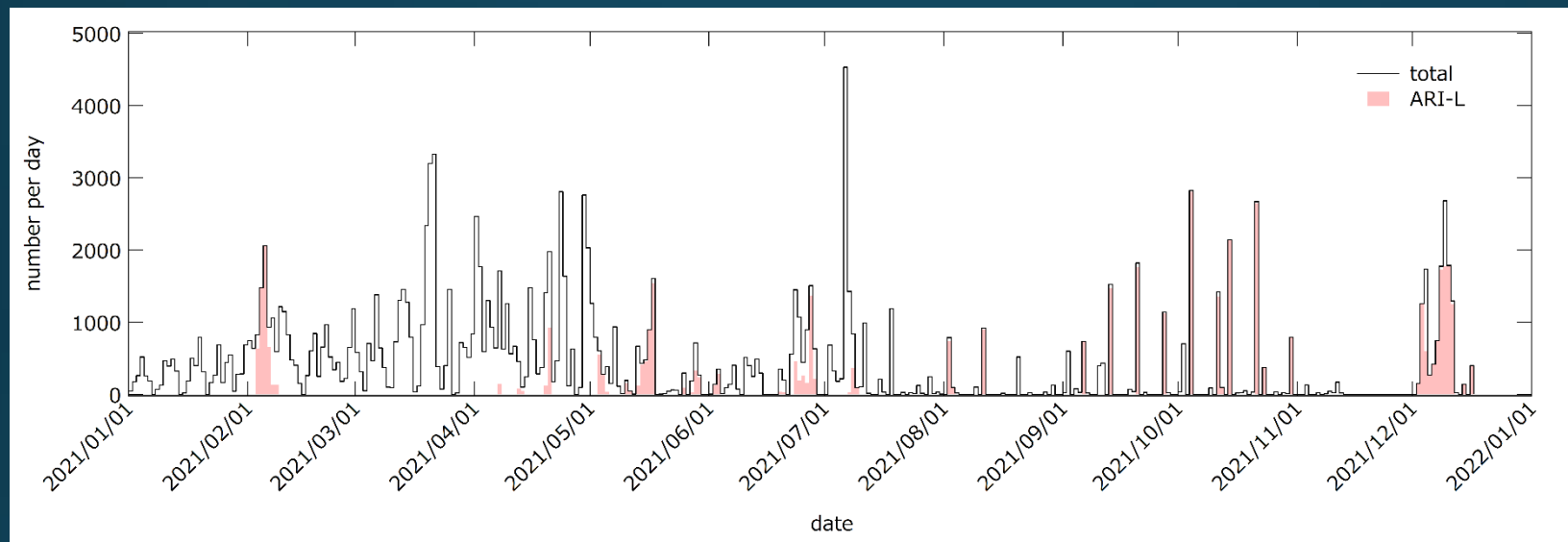
### News

- 2021-12-09: FITS WebQL v5 (experimental) is available.

#	Target Name	Project Code	Dataset ID	Cube Pix	File Size	WebQL
1	G10.47+0.03	<a href="#">2016.1.00929.S</a>	<a href="#">ALMA01116054</a>	250x250x1918	460 MB	<a href="#">FITSWebQLv4</a> <a href="#">FITSWebQLv5</a>
2	IRAS16293-2422	<a href="#">2013.1.00278.S</a>	<a href="#">ALMA01020358</a>	140x140x384	30 MB	<a href="#">FITSWebQLv4</a> <a href="#">FITSWebQLv5</a>
3	Cotton Candy Nebula	<a href="#">2016.1.01530.S</a>	<a href="#">ALMA01157161</a>	1890x1890x1918	26 GB	<a href="#">FITSWebQLv4</a> <a href="#">FITSWebQLv5</a>
4	Centaurus A	<a href="#">2013.1.00803.S</a>	<a href="#">ALMA01006956</a>	2000x2000x650	10 GB	<a href="#">FITSWebQLv4</a> <a href="#">FITSWebQLv5</a>
5	M83	<a href="#">2012.1.00762.S</a>	<a href="#">ALMA01003454</a>	1120x1600x952	6.4 GB	<a href="#">FITSWebQLv4</a> <a href="#">FITSWebQLv5</a>
6	HD 149597 (Procyon's Companion B)	<a href="#">2011.0.00165.S</a>	<a href="#">ALMA01000710</a>	216x216x71	12 MB	<a href="#">FITSWebQLv4</a>

# ALMA data release history in 2021

Number of FITS files ingested to the JVO ALMA FITS Archive per day



ALMA Cycle	Total projects (FITS provided)	Projects available for ARI-L	fraction	year
Cycle 2	381	177	46%	2013
Cycle 3	541	437	81%	2015
Cycle 4	602	206	34%	2016
Cycle 5	551	0	0	2017
Cycle 6	525	0	0	2018
Cycle 7	295	0	0	2019

Stats. in 2021  
(as of 12/16)

# of files (total)  
**150k** (695k)

File size  
**50 TB** (220 TB)

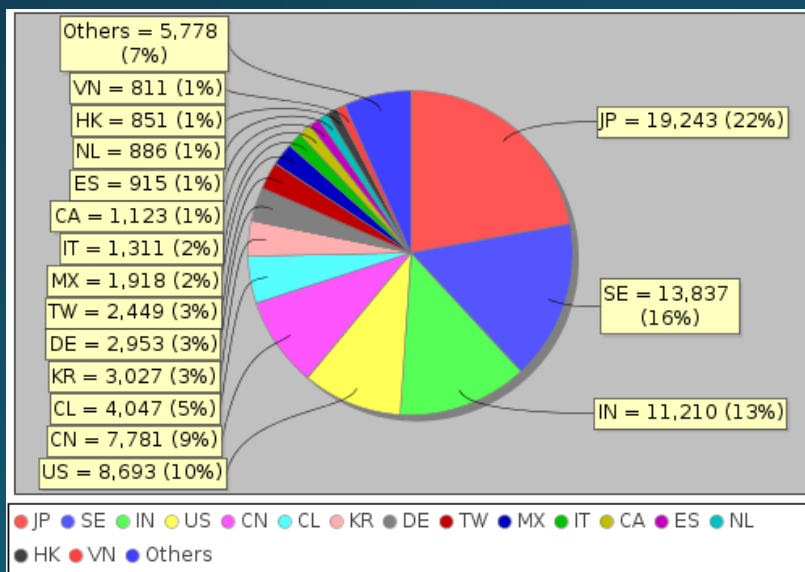
# Usage statistics

- Usage stat. in 2021 (2020)

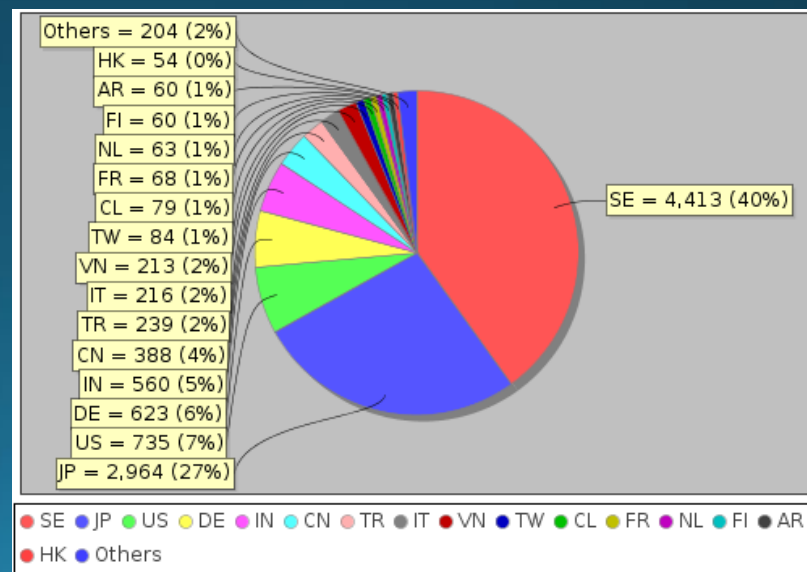
	ALMA	Nobeyama
Page access	115k (82k)	11k (20k)
DL count	32k (40k)	16k (19k)
DL size	2.7TB (5.2TB)	2.0TB (1.7TB)

- Access stat. by country

ALMA

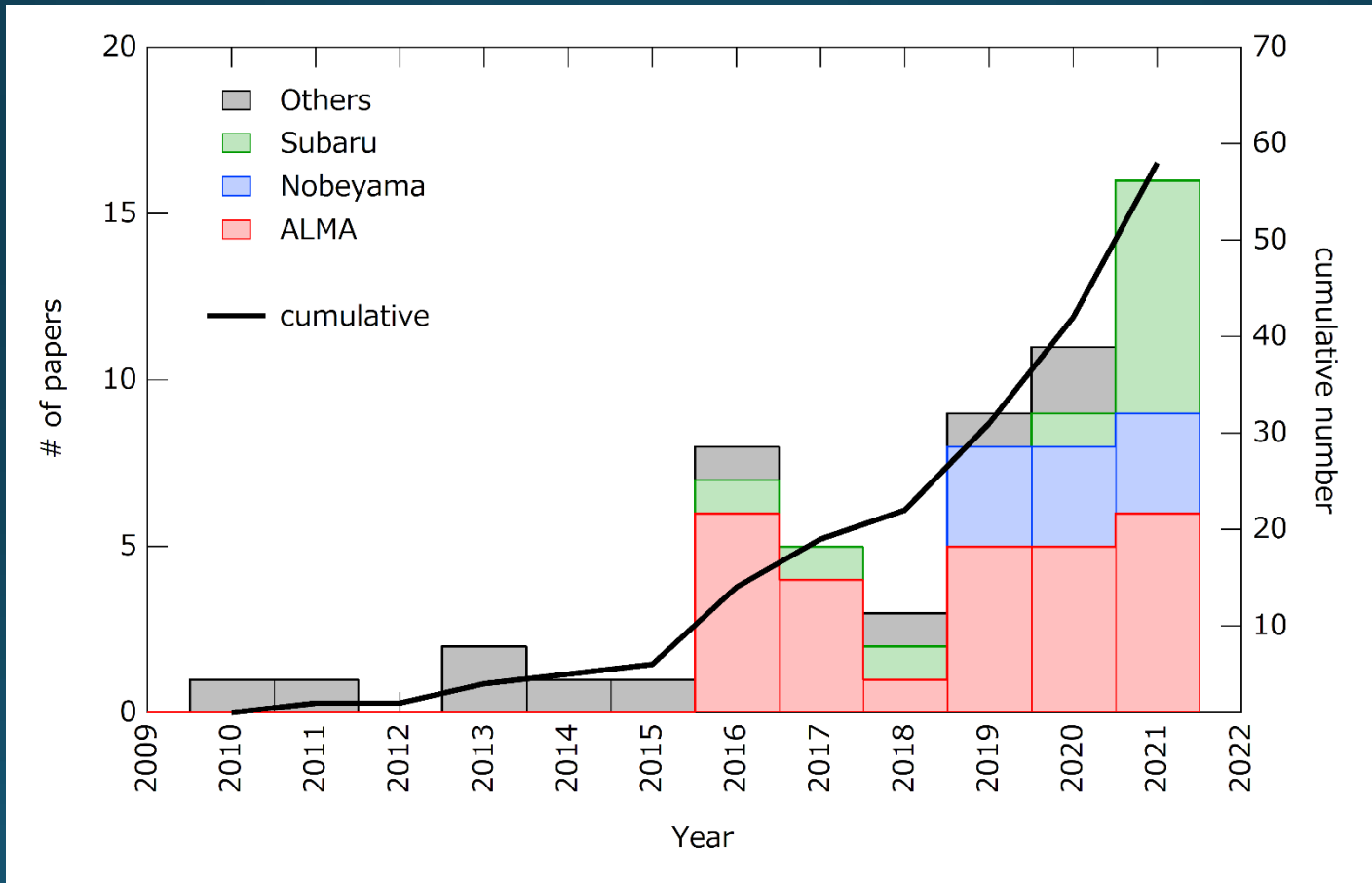


Nobeyama



# Numbers of publications by year

Numbers in 2021 : 6 (ALMA) 3 (Nobeyama) 7 (Subaru)



Publications was searched by ADS full text search.  
Refereed papers only.

# (Near) Future Plan

- **ALMA line search service**
  - Validate source velocity (redshift)
- **FITS WebQL v5**
  - Needs more work. (memory management etc.)
  - Development of new features
    - Download spectrum data in text.
    - Overplot of source catalog.
    - Moment map.
    - ...

If you have any suggestions please contact to  
help\_desk@jvo.nao.ac.jp  
or directory to me at yuji.shirasaki@nao.ac.jp