

# NRAO's New Archive and VO

Experiences integrating VO into the AAT/PPI



**Stephan Witz**

Science Support & Archive Team Lead

Atacama Large Millimeter/submillimeter Array

Karl G. Jansky Very Large Array

Robert C. Byrd Green Bank Telescope

Very Long Baseline Array



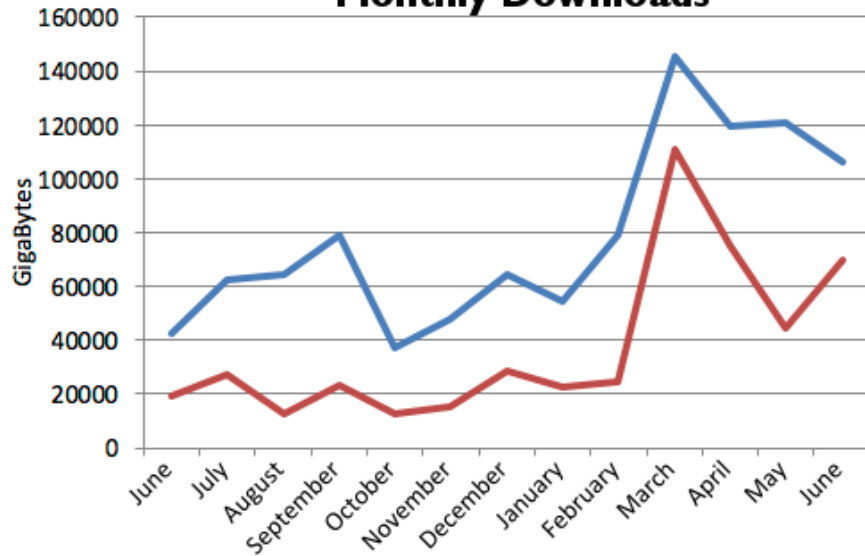
# New Archive: Requirements & Roadmap

- Requirements for next generation NRAO Archive included:
  - Query ALMA & non-ALMA (legacy VLA, Jansky VLA, VLBA & GBT),
  - Authentication & authorization for proprietary data, ALMA's CAS server for ALMA data, NRAO's CAS server for NRAO data
  - Push the button with no other input, get back a list of everything, ~600k file sets spanning 40 years
  - User-specified reprocessing using CASA and NRAO's cluster
- Phase 1: September 2015, limited proof of concept with all the major components hooked up
- Phase 2: September 2016, Phase 1 + more searching options, new responsive interface, authentication + authorization
- Phase 2.5 & 3.0 delayed by VLA Sky Survey, expected January & March 2017, more processing options, production deployment



# Archive Usage & Ingestion

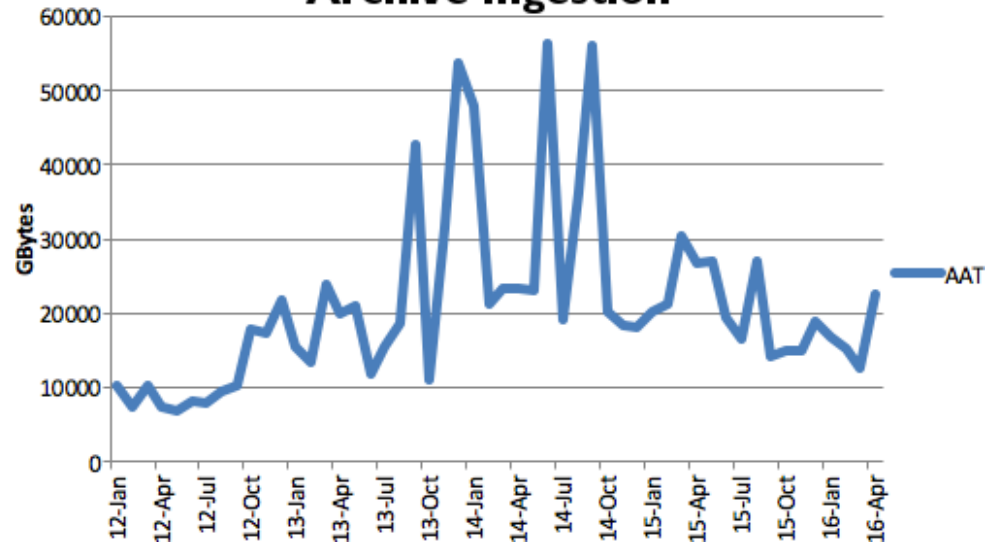
## Monthly Downloads



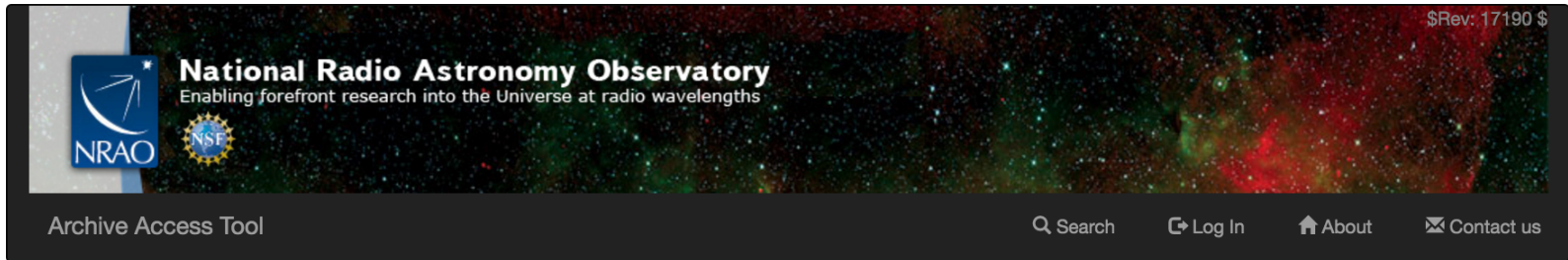
- ~75 users downloading files a day
- ~ 500gb – 2tb a day

- 96/3/1% VLA/VLBA/GBT

## Archive Ingestion




# Basic Query Interface



The header banner features a dark space background with colorful nebulae. On the left is the NRAO logo and the text "National Radio Astronomy Observatory" with the tagline "Enabling forefront research into the Universe at radio wavelengths". On the right, it says "\$Rev: 17190 \$". Below the banner is a dark navigation bar with "Archive Access Tool" on the left and "Search", "Log In", "About", and "Contact us" on the right.




Please consult the [release notes](#) before testing.

quasar Search 

[Advanced Interface](#)

[I'm Feeling Lucky](#)

[Staff](#) | [Policies](#) | [Diversity](#)




The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.




# Advanced Query Interface

\$Rev: 17190 \$



**National Radio Astronomy Observatory**  
Enabling forefront research into the Universe at radio wavelengths



Archive Access Tool

Search Log In About Contact us

Please consult the [release notes](#) before testing.

**Instrument Parameters** >

**Project Parameters** >

**Position Parameters** ▾

<b>Source Position Coordinate System:</b> Equatorial ▾	<b>Search Radius:</b> <input type="text"/> " ▾	<b>Right Ascension:</b> <input type="text"/> ° ▾	<b>Declination:</b> <input type="text"/> ° ▾
<b>Source Name:</b> <input type="text" value="Enter source name here..."/>	<b>FWHM:</b> <input type="text"/> " ▾		




**Data Parameters** >

Search


[Basic Interface](#)  
[Reset Page](#)

[Staff](#) | [Policies](#) | [Diversity](#)

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.



# Results: Table Display



**National Radio Astronomy Observatory**  
Enabling forefront research into the Universe at radio wavelengths

\$Rev: 17190 \$

Archive Access Tool

🔍 Search
🔑 Log In
🏠 About
✉ Contact us

1250 Results Found [List View](#)

10 25 50 100

Telescope	Project	Fileset Name	Observation Time
VLA	16A-108	16A-108.sb32160792.eb32383620.57572.77415369213	2016-07-03 18:34 – 20:29
VLA	12A-296	12A-296.sb7710843.eb8023847.55954.756201296295	2012-01-28 18:10 – 23:09
VLA	12B-361	12B-361.sb12587823.eb14274141.56270.09179542824	2012-12-09 02:12 – 02:57
VLA	12B-361	12B-361.sb12583684.eb14274733.56270.621956979165	2012-12-09 14:55 – 16:25
VLA	12B-361	12B-361.sb12578923.eb13834177.56240.226042685186	2012-11-09 05:25 – 06:10
VLA	12B-361	12B-361.sb12245932.eb13836046.56241.181692708335	2012-11-10 04:21 – 05:06
VLA	12B-361	12B-361.sb12250123.eb14422571.56285.62746512731	2012-12-24 15:11 – 15:56
VLA	12B-361	12B-361.sb12250123.eb14447788.56291.58538069444	2012-12-30 14:02 – 14:47
VLA	12B-361	12B-361.sb12564059.eb14409721.56281.57116119213	2012-12-20 13:42 – 14:57
VLA	12B-361	12B-361.sb12573457.eb14459360.56293.5591928588	2013-01-01 13:25 – 14:39

Download Reprocess

◀ 1 ... 4 **5** 6 ... 125 ▶

**12B-361**

**Title:**  
The Radio Proprieties of High Redshift Broad Absorption Line Quasars

**Abstract:**  
[Click to display](#)

**Proposal:**  
[Click to search](#)

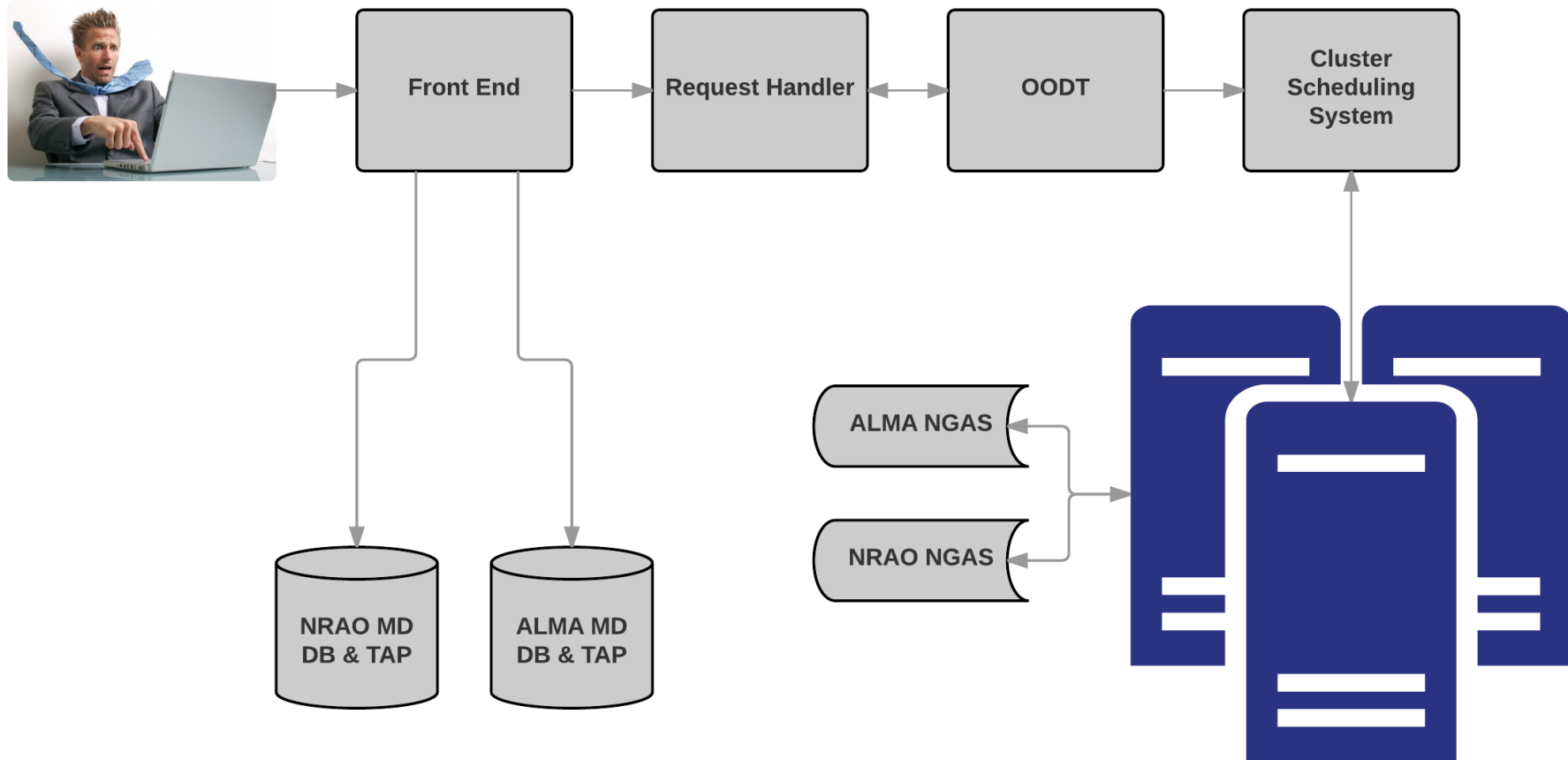
**Data Type:**  
visibility

**File Size:**  
55.701 GB

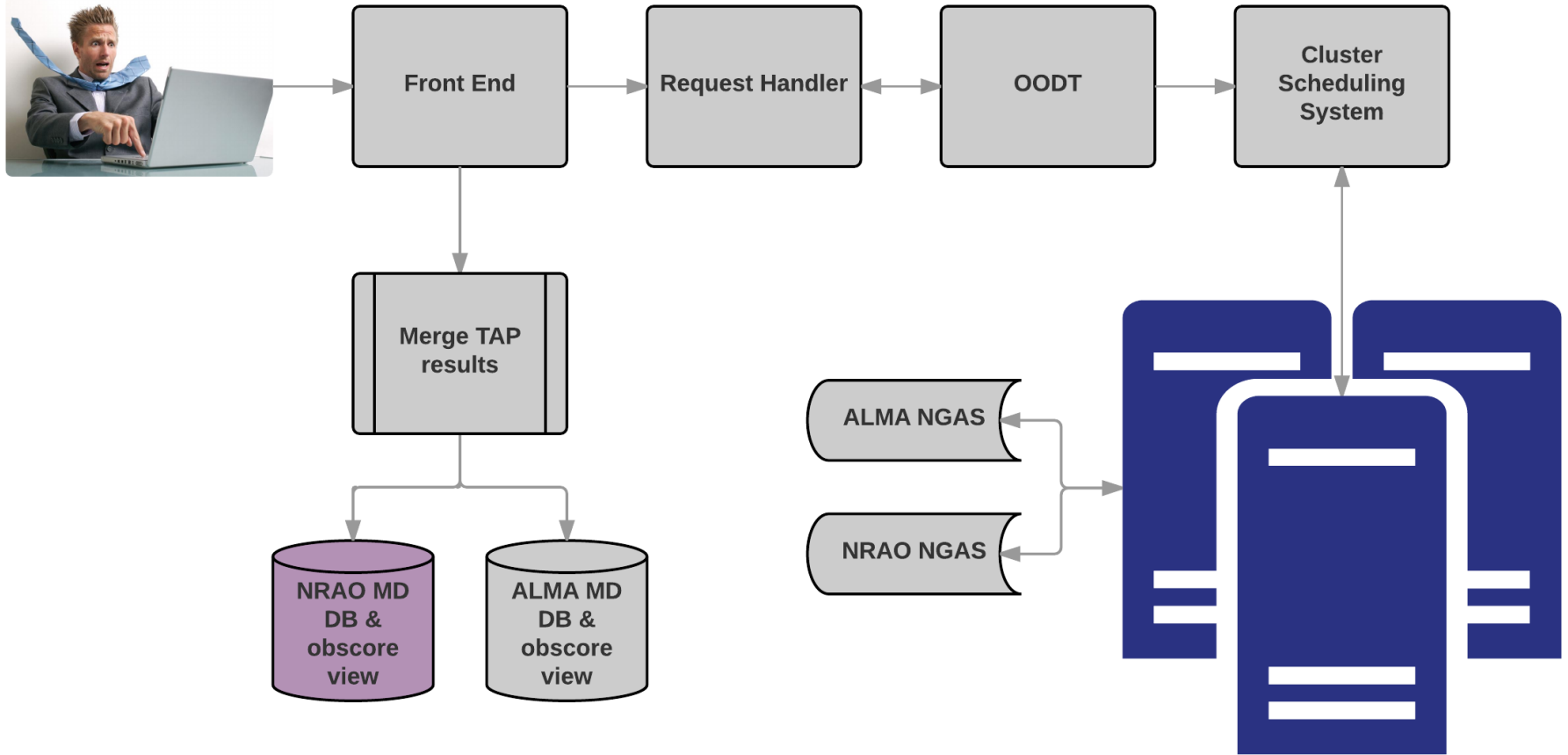
**Sources:**  
J1436+6336  
SDSSJ115+2257  
SDSSJ110+1902



# Phase I: Design



# Phase 2: Preliminary Design





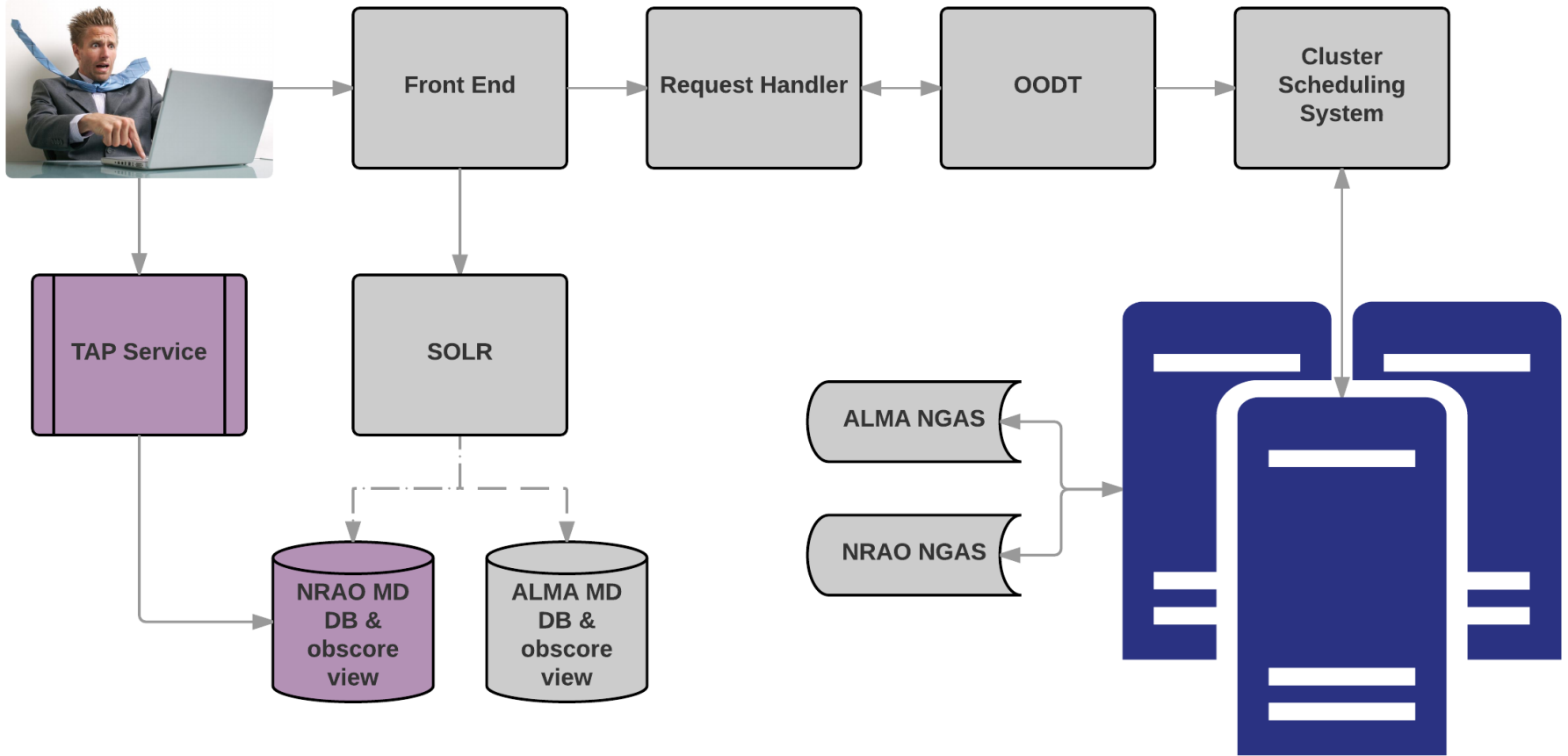
# Phase 2: Problems

*“Everybody has a plan until they get punched in the mouth.”* Mike Tyson

- Loss of two key personnel due to retirement, our local contact for VO technologies & chief developer for current archive and VO
- ALMA’s DB is Oracle, NRAO’s DB is PostgreSQL
- Implemented front end as SPA with dynamic pagination, needed LIMIT and OFFSET, no ADQL OFFSET capability, so different query for each DB (rownum, Oracle, really?)
- Merging VO tables problematic, sorting by relevance with two results sets?
- Struggled with display (part of this likely was team transitioning from Java server code to Javascript SPA & Angular.JS)
- As mentioned, Javascript SPA really wants JSON instead of XML
- We tried SOLR as a way to bridge the gap between the two DBs and the performance sold us, plus gave us natural language text searches



# Phase 2: Final Design





The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

[www.nrao.edu](http://www.nrao.edu) • [science.nrao.edu](http://science.nrao.edu)

