Arecibo Observatory

The Arecibo Observatory Archives: A 305m wide data management effort

Julio Alvarado Negron
Big Data Program Manager @ Arecibo Observatory
julio.alvarado@ucf.edu

IVOA Interop 2021

2021.05.25

ARECIBO OBSERVATORY
PUERTO RICO

Agenda

Guide topics

1. Big Data @ AO
2. 3 main sciences
3. A series of snaps
4. The data by the numbers
5. Analytic look at observations
6. The ALFA
7. Notable discoveries
8. Astronomy notable projects
9. What are we doing...
10. How can AO and IVOA collaborate
11. Q&A

ARECIBO OBSERVATORY
PUERTO RICO

Big Data @ AO

Overview

For over 57 years the Arecibo Observatory has been archiving observations data of over 1,800 proposals.

AO had the capability to generate over 80TB with existing instruments previous to the collapse. Majority of that data came from the ALFA.

AO archive holds an approximate **3PB** of observations data.

Making use of the archives

- Data Management and Governance practices implementation
- Facilitate access to community to Arecibo's data
- Enables access to High-Performing Computing
- Implement best practices and lessons learned from partner observatories and research community

ARECIBO OBSERVATORY
PUERTO RICO

A Full Spectrum Pioneer of Sciences Since 1963

Radio Astronomy

Is the study of radio waves produced by a astronomical objects such as Sun, planets, pulsars, stars, etc. Arecibo radio telescope sensitivity allows astronomers to detect faint radio signals from far-off regions of the universe.

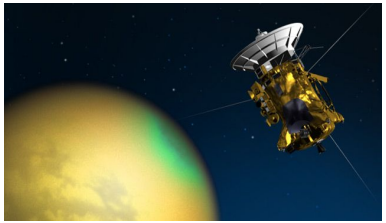
Fast Radio Bursts, Pulsars, Spectral line, Exoplanets, VLBI.

More Info Here

Atmospheric Sciences

Is the investigation of the earth's gaseous envelope. The Arecibo Radio Telescope can measure the growth and decay of disturbances in ionosphere (altitudes above 30 miles). The "big dish" is also used to study plasma physics processes in the electrically charged regions where radio waves are influenced most.

More Info Here



Planetary Radar

The Arecibo Observatory was the world's most powerful planetary radar system. The 305 meter Arecibo telescope equipped with a 1 MW transmitter at S-band (12.6 cm, 2380 MHz) was used for studies of small bodies in the solar system, terrestrial planets, and planetary satellites including the Moon.

More Info Here

ARECIBO OBSERVATORY
PUERTO RICO

A Sequence of Snaps

First Cable Snaps

On August 10th a first cable snaps causing damage to the dish.

Second Cable Snaps

On November 6th a second cable snaps causing major damage to the dish.

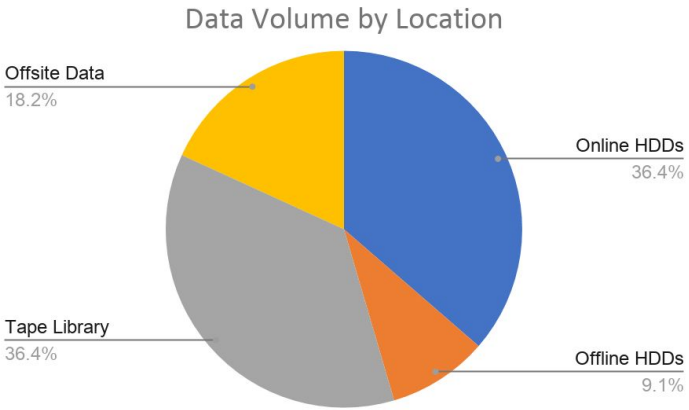
December's Check Mate

A main support cable broke from Tower 4, causing the platform to fall over the dish.

The team got together and realized that the data safety, integrity and accessibility was a priority.

ARECIBO OBSERVATORY
PUERTO RICO

The Data in Numbers and Infra Limitations



The Big Picture

Arecibo Observatory holds over 3PB of data onsite. This amount is spread between active hard drives, offline disks and the tape library.

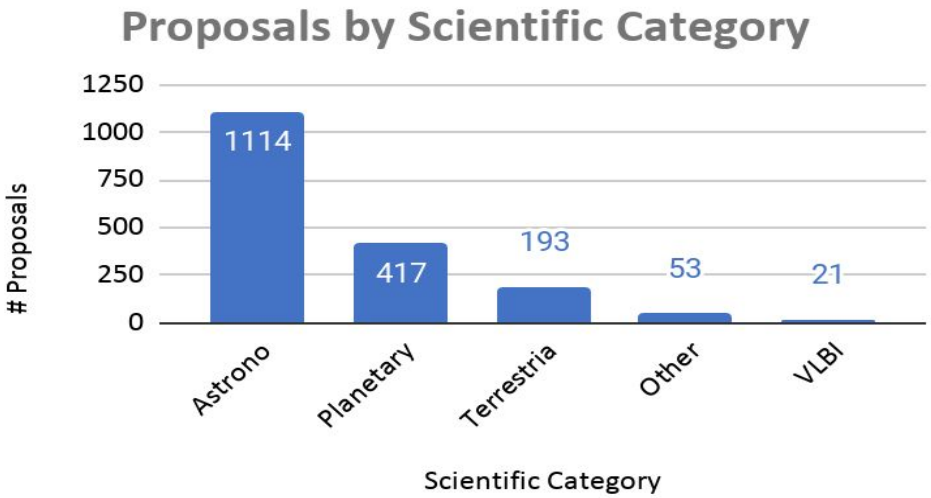
Arecibo also has copies of data stored on various institutions across the globe, to which we refer to as offsite data.

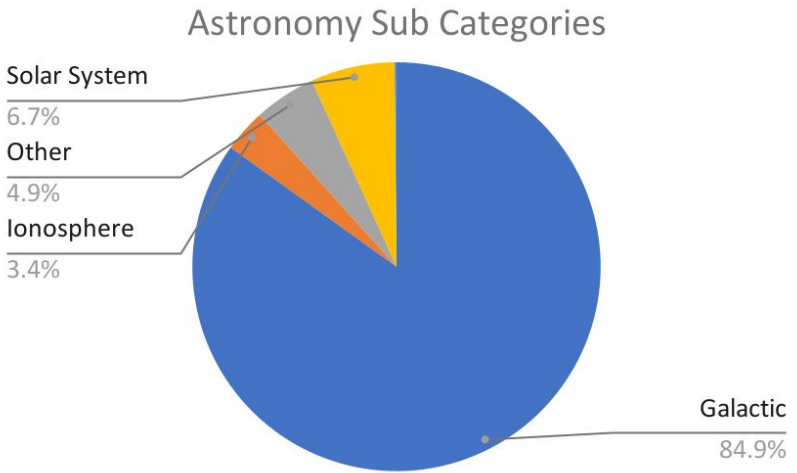
What to do with the data

After the collapse, Arecibo and working teams partnered to safely store the data at TACC. Having the data centralized allows us to build a long term data management strategy in benefit of the scientific community.

ARECIBO OBSERVATORY
PUERTO RICO

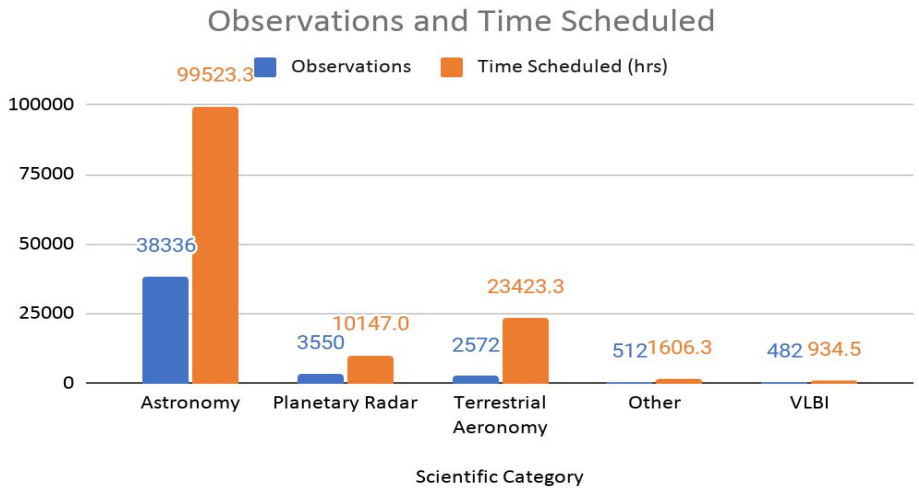
50+ Years of Contributions

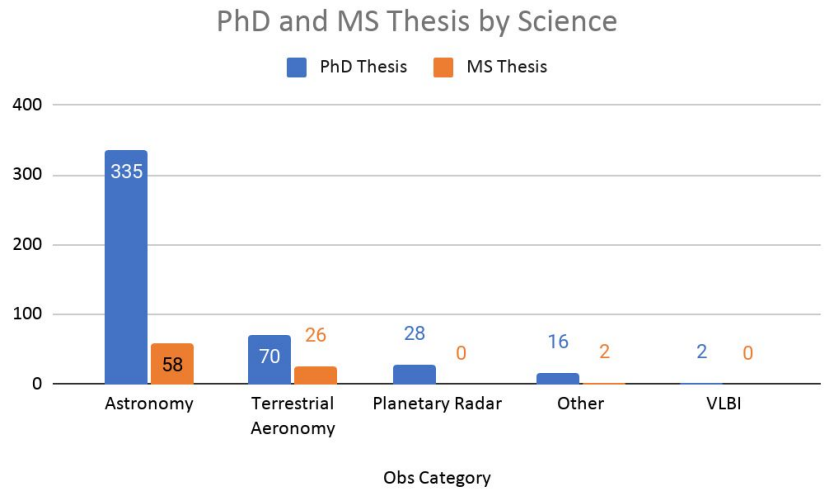




ARECIBO OBSERVATORY
PUERTO RICO

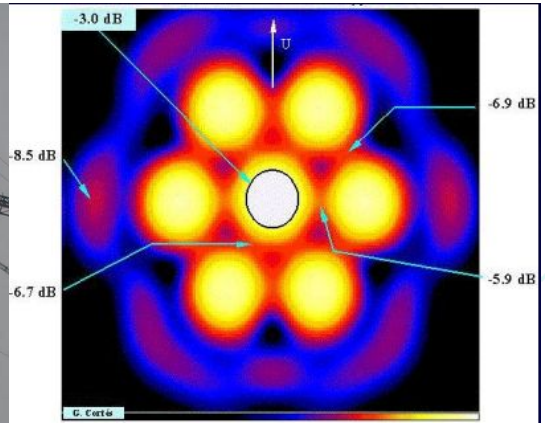
50+ Years of Contributions





ARECIBO OBSERVATORY
PUERTO RICO

Knowing the Sources and Discoveries



ALFA

The Arecibo L-band Feed Array (ALFA) is a seven feed system that allows large-scale surveys of the sky to be conducted with unprecedented sensitivity using the 305-m Arecibo telescope in Puerto Rico. ALFA, operating near 1.4 GHz, consists of a cluster of seven cooled dual-polarization feeds, a fiber-optical transmission system, and digital back-end signal processors.

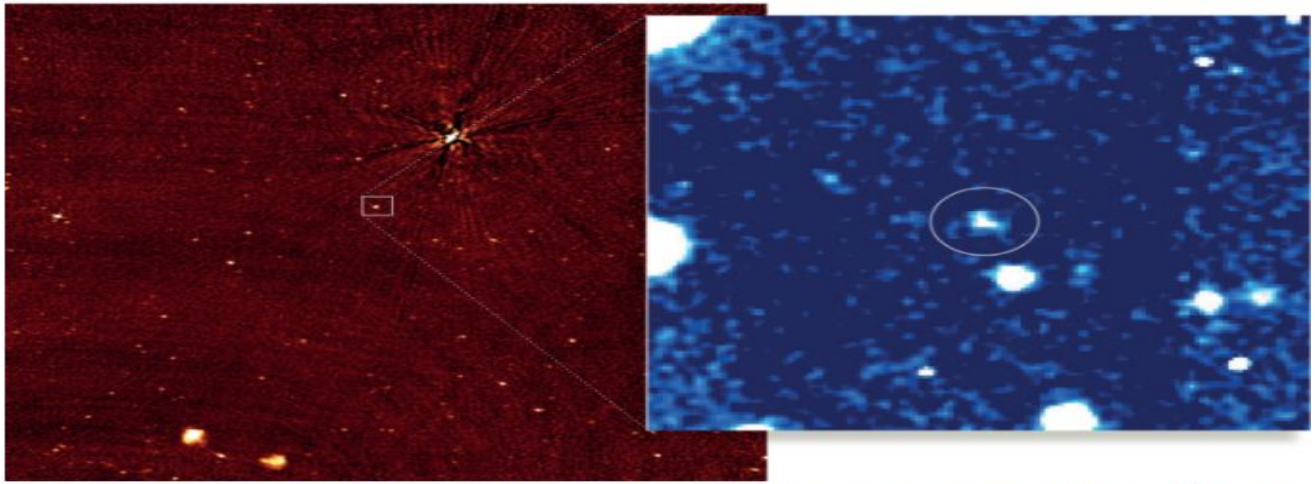
Most of this projects are considered “surveys” due to their nature. The radar is left static in a position while the Earth rotates, allowing to “drift scan” the sky above Arecibo.

It could generate an aggregate of **875MB/s**, **76TB** per day.

Using ALFA for ALFALFA

ARECIBO OBSERVATORY
PUERTO RICO

Knowing the Sources and Discoveries



A repeating source of radio waves discovered by Arecibo (radio image, left) was the first fast radio burst traced back to its home galaxy. The burst originated in a dwarf galaxy about 2.5 billion light-years away (visible light image, right).

H. FALCKE/*NATURE 2017*

Fast Radio Bursts

Fast radio bursts, or FRBs, are brief, brilliant blasts of radio waves with unknown origins. The first FRB known to give off multiple bursts was FRB 121102, which Arecibo first spotted in 2012 and again in 2015.

Arecibo's discovery backed up the theory from the Charles Parkes telescope in Australia that FRB's are events that come farther than the Milky Way.

Radio bursts are observed during 90 days followed by a silent period of 67 days. The same behaviour then repeats every 157 days.

- Arecibo Discoveries

ARECIBO OBSERVATORY
PUERTO RICO

Notable Projects

Notable projects in Radio Astronomy

- **The Arecibo Galactic Chemistry Survey (Yale University - 2009)**
  a. Used Mock spectrometers to make a 1-10 GHz molecular-line census of the galactic high-mass star-forming region W51
  b. Survey expected to facilitate discoveries of molecules considered pre-biotic.
- **Origin of Carbon Radio Recombination Line formation in the Galactic plane (Anish Roshi - AO - 12/2021)**
  a. Plan to observe Carbon Radio Recombination Lines (CRRL) from the Galactic plane. The origin of these line formation is not well understood.
- **Confirmational Study of OH 18 cm emission from Intermediate-Velocity Molecular Clouds (Univ of North Georgia - 2019)**
  a. This was based on Rohser et al 2016b, produced a list of 239 candidate IVMCs based on HI/Far-IR correlation
  b. Survey the X positions observable from Arecibo for the presence of OH 18cm emission to confirm the candidate IVMCs.

Notable Projects

Notable projects in Radio Astronomy

- **The G-ALFA Continuum Transit Survey (GALFACTS - Univ of Calgary - 2005)**
    a. Major observational advance in imaging of the polarized radiation from the Milky Way and other galaxies
    b. Understanding of the magnetic field of the Milky Way, the properties of the magneto-ionic medium, and the role of magnetic fields in galactic processes
    c. Instrument(s): P-ALFA/GALFA Continuum spectrometer

The GALFACTS Survey

ARECIBO OBSERVATORY
PUERTO RICO

What are we doing...

...to give back life to Arecibo archives?

1. **Data Inventory**
   - ✓ Understand the variety of sciences, uses and attributes of our projects
2. **Data Transfer**
   - ✓ Get the data archives to a safe and accessible place
3. **Data Catalog**
   - ✓ Structure and organize our archives to facilitate data management
4. **Voice of the Community**
   - ✓ Showcase the archives in forums like this one to get community feedback and interest
5. **Enable discoveries**
   - ✓ Develop a set of processes and tools to facilitate community access to AO archives
   - ✓ Create a friendly portal for users to browse and access our archives

ARECIBO OBSERVATORY
PUERTO RICO

How can AO and IVOA collaborate

Bringing AO to community standards

Arecibo is collaborating with community partners such as CICoE and EPOC to develop a robust Data Management framework.

Some possible IVOA standards applications:

- **Dataset characterization/modeling**
  - ✓ With near 3M FITS files, and counting, AO could leverage IVOA standards on data modeling and semantics to present attributes in a standardized manner.
  - ✓ Currently running Python scraping scripts to the FITS files to get the header attributes
- **Data Access**
  - ✓ Leverage IVOA's DAL to provide a robust data access layer to AO archives
  - ✓ Provide the community with tools that allows effective data sharing
  - ✓ Research tools/processes to share newly captured data with IVOA

CI CoE PILOT

IVOA

EPOC
Engagement and Performance Operations Center

ARECIBO OBSERVATORY
PUERTO RICO
UCF · YEI · UMET

ARECIBO OBSERVATORY
PUERTO RICO

How can AO and IVOA collaborate

Bringing AO to community standards

- **User interaction**
  - ✓ Learning from what has been done, best practices, standards
  - ✓ Develop user friendly applications to users
  - ✓ Visualization, catalog, browsing tools
- **Arecibo in-house tools**
  - ✓ Data catalog enabling search-engine like access to the proposals, observations log, RAW file details, etc.
  - ✓ File Catalog - registry of files uploaded to TACC: file name, path, size, project.
  - ✓ FITS catalog - registry of FITS files attributes, could be used for filtering, etc.
  - ✓ **Bring all the pieces together with the help of community accepted standards.**

ARECIBO OBSERVATORY
PUERTO RICO
UCF • YEI • UMET

ARECIBO OBSERVATORY
PUERTO RICO

Nowadays the value of our archives is not what's discovered, but what can we discover.

Q&A time!

ARECIBO OBSERVATORY
PUERTO RICO