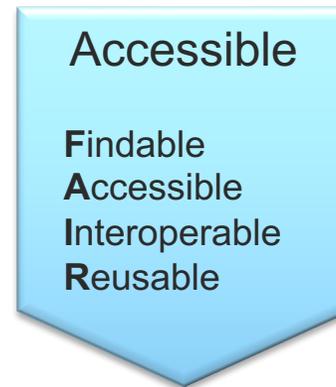


Open-Science for the Aerosol Community

ADAM THEISEN, GIRI PRAKASH, MAXWELL GROVER

What is Open Science?

- ▶ From NASA's Open-Source Science Initiative (<https://science.nasa.gov/open-science-overview>)
 - **Transparent** – Scientific process and results should be visible, accessible, and understandable
 - **Inclusive** – Process and participants should welcome participation by and collaboration with diverse people and organizations
 - **Accessible** – Data, tools, software, documentation, educational resources, and publications should be accessible to all (Findability, Accessibility, Interoperability, and Reuse -FAIR)
 - **Reproducible** – Scientific process and results should be open such that they are reproducible by members of the community

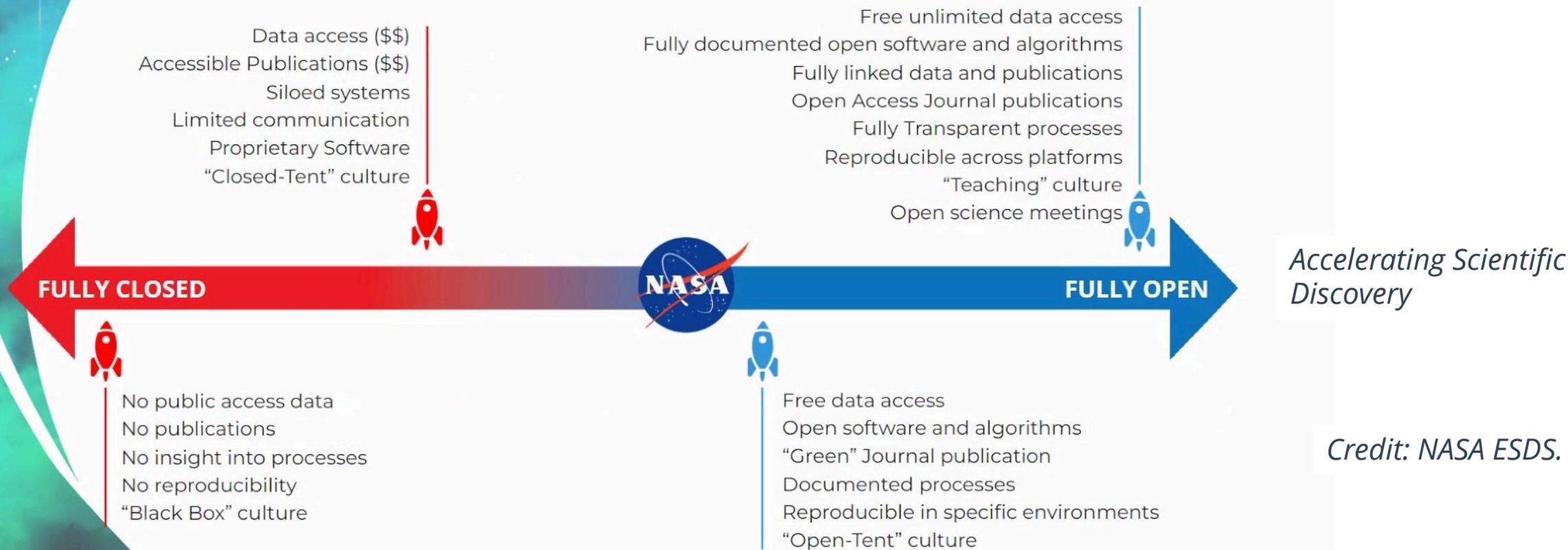


Why Open Science?

NASA's Open-Source science is the *activation* of an open science community



A continuum of open-source science



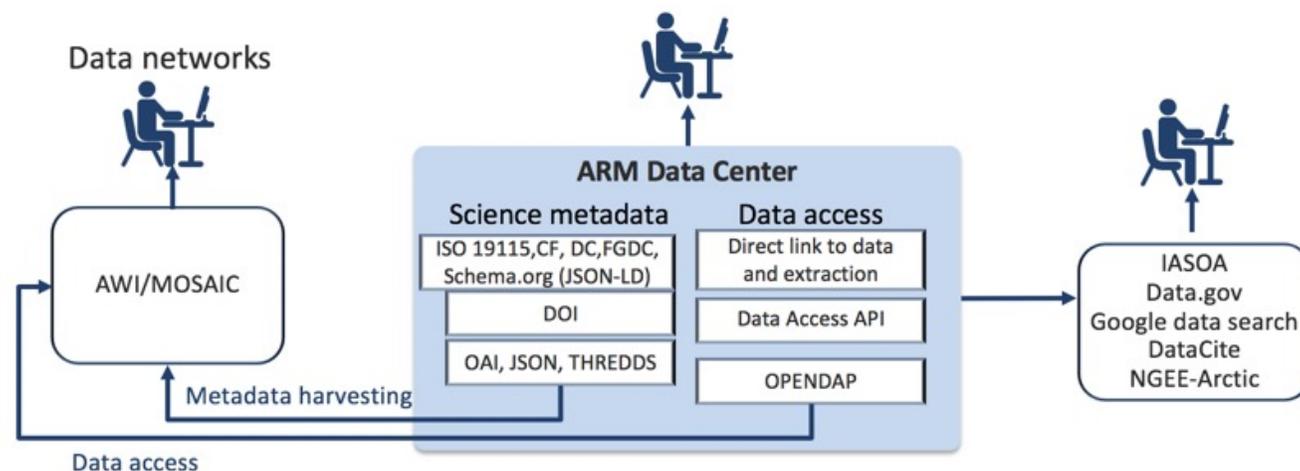
ARM Data Sharing Options

GIRI PRAKASH

ARM Data Center
ASR/ARM PI Meeting, 2023
<https://arm.gov/>

ARM Data sharing to Other Repositories

- ▶ ARM provides metadata in internationally adopted standards for data sharing
 - ISO 19115, FGDC, Schema.org (JSON-LD), DublinCore, EML, etc.
- ▶ Data Identifiers/DOIs:
 - ARM uses a dedicated DOI prefix (10.5439) for all of its datastreams. These DOIs will be provided as part of the metadata records.
- ▶ Data Acknowledgement: ARM will provide guidance for citing and acknowledging its data.
- ▶ ARM POC: Giri Prakash



ARM Data sharing to Other Repositories

- ▶ **Data Access:**
 - Ensure the latest version of data available for users
 - Data endpoints are provided in the metadata
 - Direct link via API-based live data service
- ▶ Provide access to data quality, plots, and other ancillary details
- ▶ Options for users to get notified of any data quality changes or new data versions

30EBBR

Selected data level
 bt Start: 1993-07-20 End: 2022-05-08
 bt: QC checks applied to measurements

Description: Energy Balance Bowen Ratio (EBBR) station: surf. heat flux and related data, 30-min
Site: Southern Great Plains (SGP)
Location: Lamont, OK (Extended and Co-located with C1)
Facility Code: E13
Category: Radiometric
Data Type: Routine Data
Source Instrument/Data: Energy Balance Bowen Ratio Station
Sampling Interval: 30 seconds
Start Date: 1993-07-20
End Date: 2022-05-08
DOI: 10.5439/1023895

Data Timeline & Quality
 Filter DQRs based on zoom Resolution: 14 Days Viewing: latent_heat_flux
 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
 Legend: ROUTINE, INCORRECT, SUSPECT, MISSING, NOTE, LIMITED ACCESS

Data Plots
 Primary Measurements | File Header Information

VARIABLES:
Soil heat flux: Variable: Surface soil heat flux, average of fluxes 1-5 (surface_soil_heat_flux_avg) Recommended
Latent heat flux: Variable: Latent heat flux (latent_heat_flux) Recommended
Sensible heat flux: Variable: Sensible heat flux (sensible_heat_flux) Recommended
Net broadband total irradiance: Variable: Net radiation (net_radiation) Recommended

Instrument Contacts:
 Ryan Sullivan (Lead Mentor), Evan Keeler (Associate Mentor), Brian Ermold (Developer)

Additional Resources:
 Instrument/VAP Info, Related Publications, Instrument Handbook

Actions: Visualize Data, Tag this Data, Add to Cart

Sample ARM metadata

- ▶ Sample ARM metadata records can be fetched from:
 - <https://adc.arm.gov/metadata/guc/sitemap.xml>
- ▶ Example of HTML metadata page: <https://www.archive.arm.gov/metadata/guc/html/30ecor.html>
 - The JSON metadata is embedded within the source
- ▶ ISO metadata records are available as xml files:
 - <https://www.archive.arm.gov/metadata/guc/xml/30ecor.xml>
- ▶ Repositories and data portals can harvest ARM metadata records and make them available in their data portals using the above metadata records. By doing this, users will discover ARM data within the external portal but will be directed to ARM discovery or API endpoint for downloading the corresponding ARM data.

ARM Data in External Portals

► Here are some external portals that consumes ARM metadata in their portal:

- MOSAIC portal: <https://marine-data.de/?site=data&qf=keywords.project.name%2FMOSAiC&qf=organisations.provider.abbreviation%2FARM>
- IASOA: <https://psl.noaa.gov/iasoa/dataataglace>
- Arctic Data Federation: <https://search.dataone.org/portals/polderdemo/Data>
- DataOne: <https://search.dataone.org/data>

DAM DEUTSCHE ALLIANZ MEERESFORSCHUNG | **HELMHOLTZ** RESEARCH FOR GRAND CHALLENGES

HOME EXPEDITIONS PLATFORMS DATA VIEWER

Search for author, expedition, project, ...

MAP

DATASETS PUBLICATIONS REPORTS MAPS SAMPLES SORT BY: RELEVANCE DATE ↓

SELECTED FILTERS: MOSAIC × ARM ×

AOS: Fine Condensation Particle Counter - aospcf (2020)
 Andrews, Elisabeth; Kuang, Chongai; Salwen, Cynthia; Singh, Ashish
<http://dx.doi.org/10.5439/1046184>

Marine W-Band ARM Cloud Radar (MWACR) CF-Radial Spectral Data, Co-Polarization Mode - mwacrfspccopol (2020)
 (Andrei) Lindenmaier, Iosif; Castro, Vagner; Wendler, Tim; Matthews, Alyssa; Johnson, Karen; Giangrande, Scott
<http://dx.doi.org/10.5439/1608613>

Ka-Band ARM Zenith RADAR (KAZR) CF-Radial Spectral Data, general mode, co-polarized mode - kazrcfrspgcepol (2020)
 (Andrei) Lindenmaier, Iosif; Nelson, Danny; Matthews, Alyssa; Castro, Vagner; Wendler, Tim;

Temporal Coverage

Author

(Andrei) Lindenmaier, Iosif (7)
 Andreas, Afshin (1)
 Andrews, Elisabeth (8)
 Cadeddu, Maria (2)

Arctic Data Federation
 An Arctic Data portal for data from dozens of repositories across the DataONE Federation.

About Data Contacts

Search CURRENT SEARCH CLEAR ALL

Search these datasets Repository: ARM ×

Repository Person Year of publication Year of data coverage Scientific Name

ARM Person 1800 to 2022 1800 to 2022 Scientific Name

Datasets 1 to 25 of 1,591

1 2 3 ... 64 Next Sort by Most recent

ARM Damao Zhang. 2015. **Micropulse Lidar cloud mask using machine learning model from Cromwell et al 2019 - mplcmaskml**. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1614821, version: 43318cdeecud48cae94da3099477322.

ARM Laura Riihimaki. 2019. **Portable Radiation Package on ice, includes automated and manual qc - iceradrhihimaki**. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1814821, version: 43318cdeecud48cae94da3099477322.

ARM Lynn Ma and Richard Wagener. 1997. **Cimel Sunphotometer (CSPHOT): almucontars sky radiance data, version 3 - csphotalmv3**. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1485686, version: 70b3c54fbb0b7db330d25df4e9a2d8e.

ARM 2012. **KAZR-ARSLC product after reflectivities have been statistically aligned, on a monthly basis, with those of the CloudSat Cloud Profiling Radar - arslkazrcloudsat**. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1728666, version: 8184b4637b0dc78e90ee87888cc0081.

Hide Map

Open Science Tools and Activities

Software to Bridge Communities

- ▶ Data sharing can be hard!
 - Share metadata to link to data between organizations
 - Unlikely that all these organizations are going to change their processes and data formats to a common standard
- ▶ Sharing software is easy!
 - Come to a standard in the software space
 - Transparent processing
 - Consistent processing between organizations
 - Contributions that impact a broader community

Py-ART Usage (10 Years)

Fork 251

★ Starred 431

downloads 4.5k/month

ACT Usage (4.5 Years)

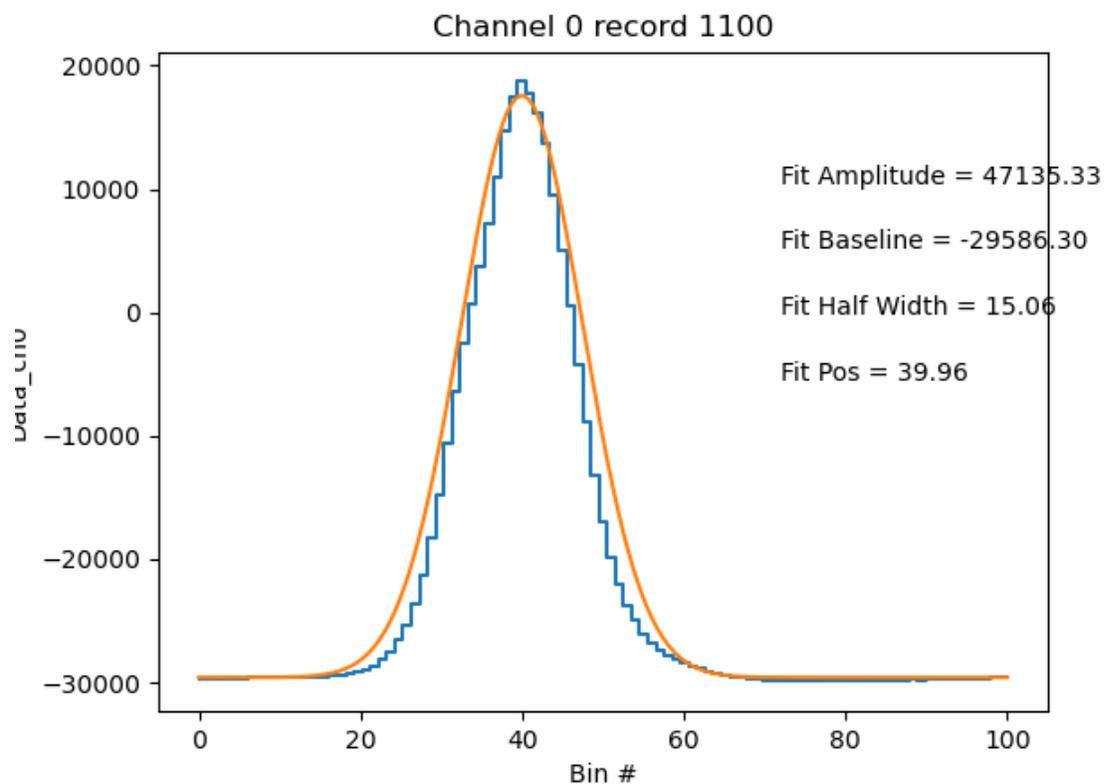
Fork 32

★ Starred 114

downloads 67k total

Early Open Aerosol Success!

- ▶ PySP2 - <https://github.com/ARM-DOE/PySP2>
 - ▶ Python package for processing and visualizing SP2 data
 - ▶ Open and transparent processing that ARM is now using for processing SP2 data
 - ▶ Received contributions from non-ARM funded researchers that then benefit ARM and the broader community



Art Sedlacek and Bobby Jackson

Atmospheric data Community Toolkit (ACT)

► Python library for working with time-series data from atmospheric research instrumentation

Discovery

IO

ACT

Plotting

Utilities

Corrections

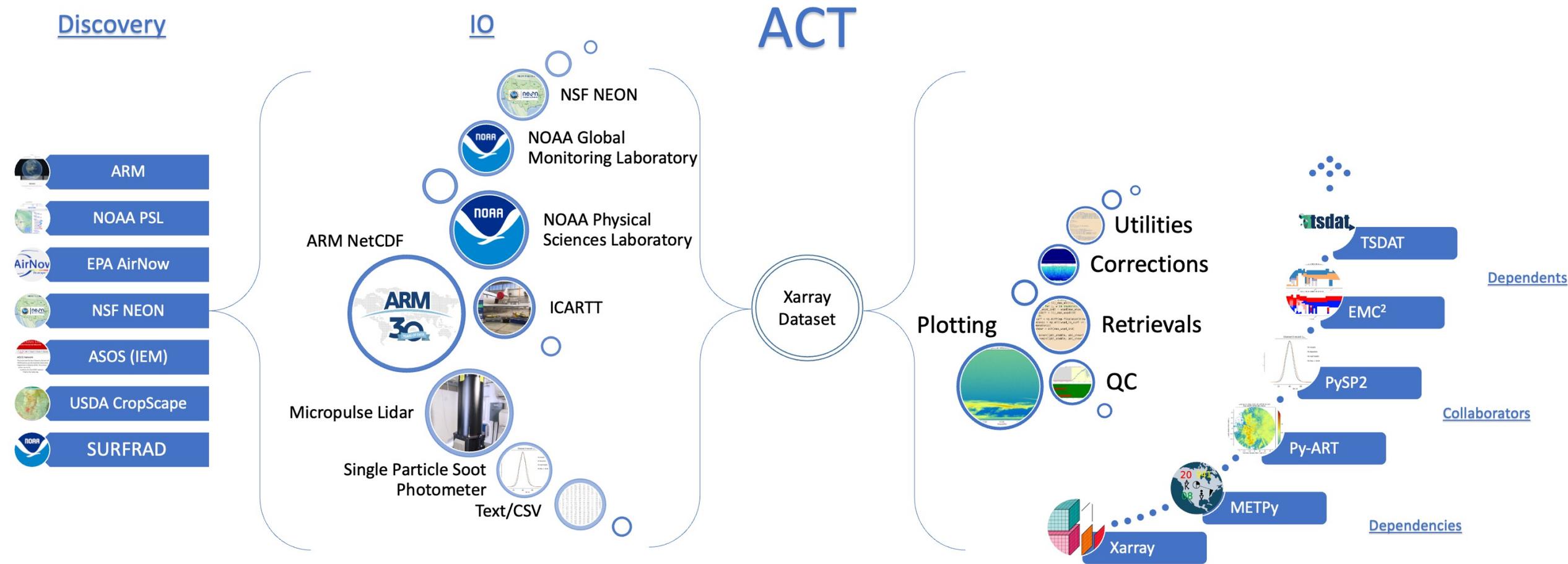
Retrievals

QC

Dependents

Collaborators

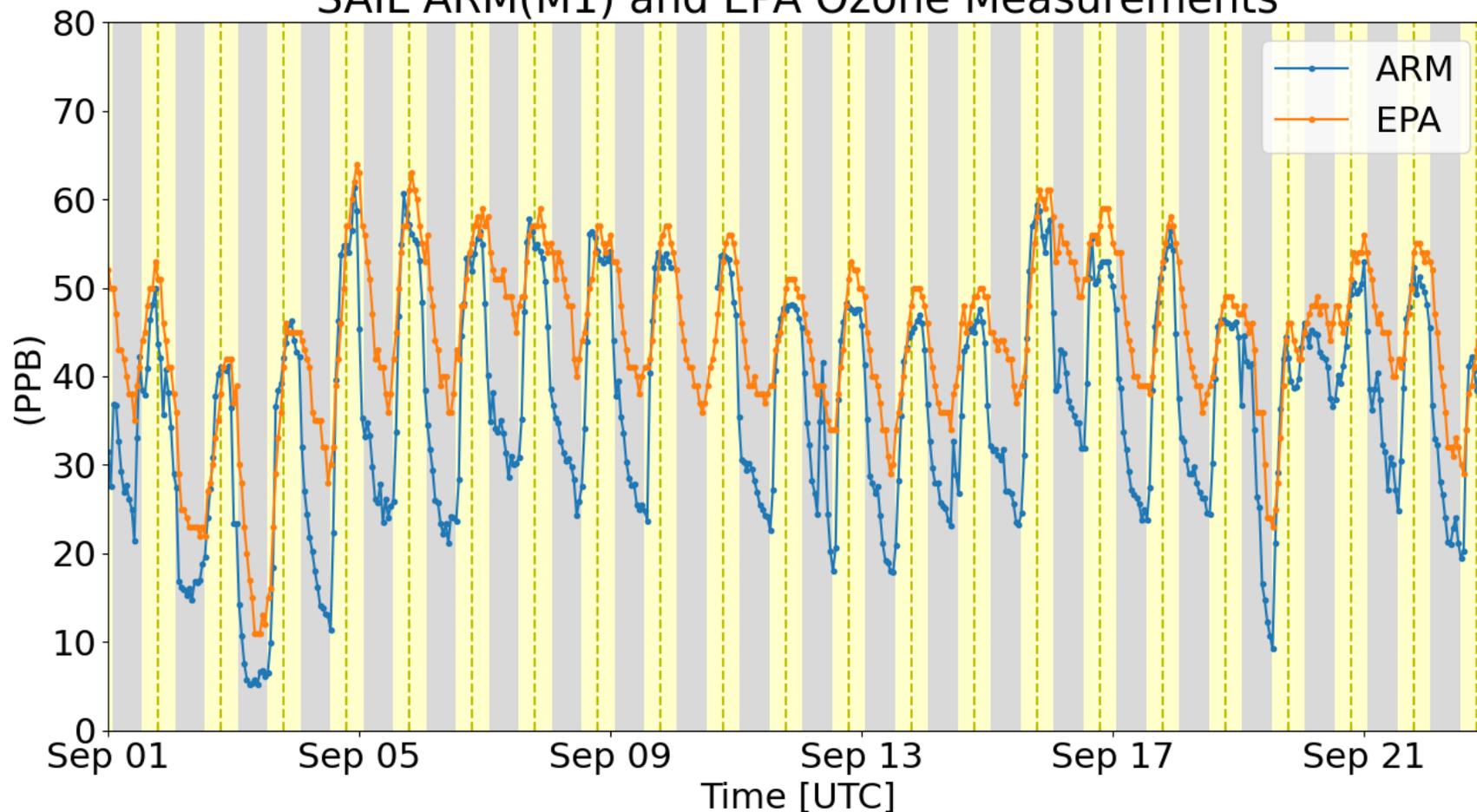
Dependencies



ARM and EPA Ozone Data

Achieve a better understanding of measurement performance

SAIL ARM(M1) and EPA Ozone Measurements



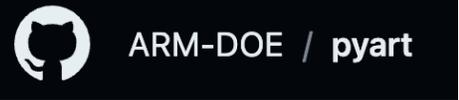
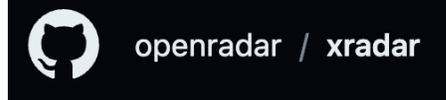
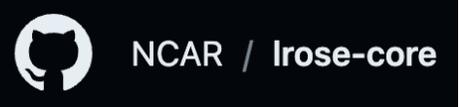
```
lat_lon = '-106.994245,38.9504,-106.959845,38.97245'
```

```
results =
act.discovery.get_airnow_bounded_obs(
    token, '2021-09-01T00',
    '2021-09-23T23', lat_lon,
    'OZONE,PM25', data_type='B'
)
```

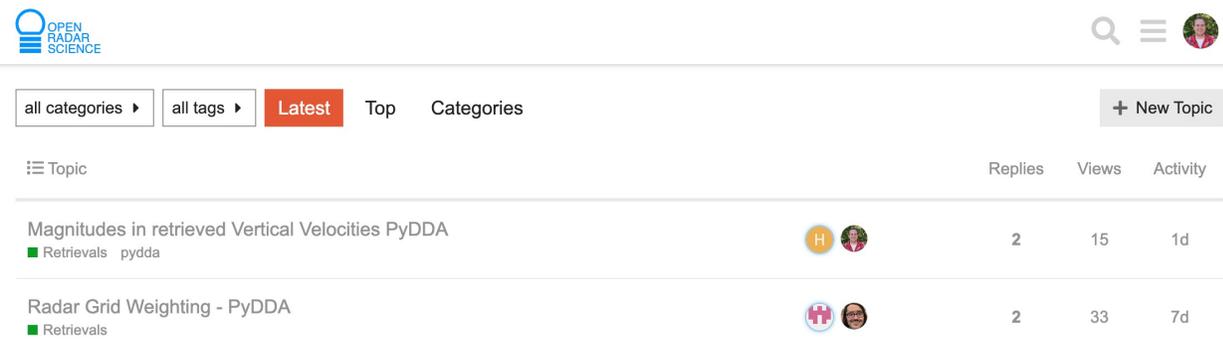
What is the Open Radar Stack? How do we Collaborate?



openradar
Unwritten understanding all code addition to be done by PR and everything to be collaborative. No unexpected actions.

 Python	 ARM-DOE / pyart	 wradlib / wradlib	 openradar / xradar
 C++	 NCAR / Irose-core	The key concept – interoperability , not necessarily one language or tool	
 Java	 baltrad / BaltradDex		

Discussions
+ Support on
the Open
Radar Forum



OPEN RADAR SCIENCE

all categories ▾ all tags ▾ Latest Top Categories + New Topic

Topic	Replies	Views	Activity
Magnitudes in retrieved Vertical Velocities PyDDA Retrievals pydda	2	15	1d
Radar Grid Weighting - PyDDA Retrievals	2	33	7d

How do we continue grow our Open Radar community?

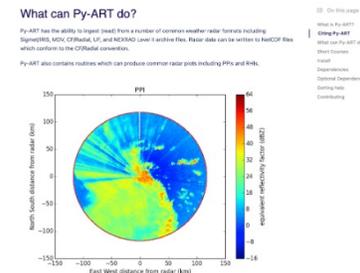
- ▶ Support of existing users
 - Continue to improve outreach + examples
 - Encourage contributions
- ▶ New features
 - Support of xarray/cfradial2 data model– aligning with Pangeo ecosystem
 - New cookbooks focused on working with ARM + other data
- ▶ Outreach opportunities
 - Conferences (ERAD, AMS Radar, AMS Annual)
 - Regional workshops
 - SciPy

Active Py-ART Development - Submitting a Pull Request (PR)

26 April 2022 Joe O'Brien

The motivation for this blog comes from wanting to change the colorscheme within the default Py-ART documentation images to a more (color vision deficiency friendly color scheme).

Here are the images we are working with:



PANGEO



Open Aerosols?

NASA's Open-Source science is the *activation* of an open science community



Accelerating Scientific Discovery

A continuum of open-source science

Data access (\$\$)
Accessible Publications (\$\$)
Siloed systems
Limited communication
Proprietary Software
"Closed-Tent" culture



Free unlimited data access
Fully documented open software and algorithms
Fully linked data and publications
Open Access Journal publications
Fully Transparent processes
Reproducible across platforms
"Teaching" culture
Open science meetings



FULLY CLOSED



FULLY OPEN



No public access data
No publications
No insight into processes
No reproducibility
"Black Box" culture



Free data access
Open software and algorithms
"Green" Journal publication
Documented processes
Reproducible in specific environments
"Open-Tent" culture

Credit: NASA ESDS.

Questions