

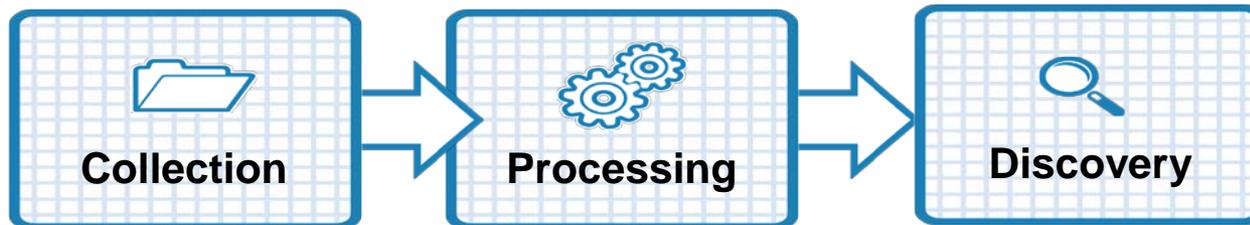
# ARM

CLIMATE RESEARCH FACILITY

# Updates from ARM Data World

Presented To  
ARM/ASR Joint User Facility PI Meeting

Presented By  
Giri Palanisamy  
March 17, 2015



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Presentation Summary

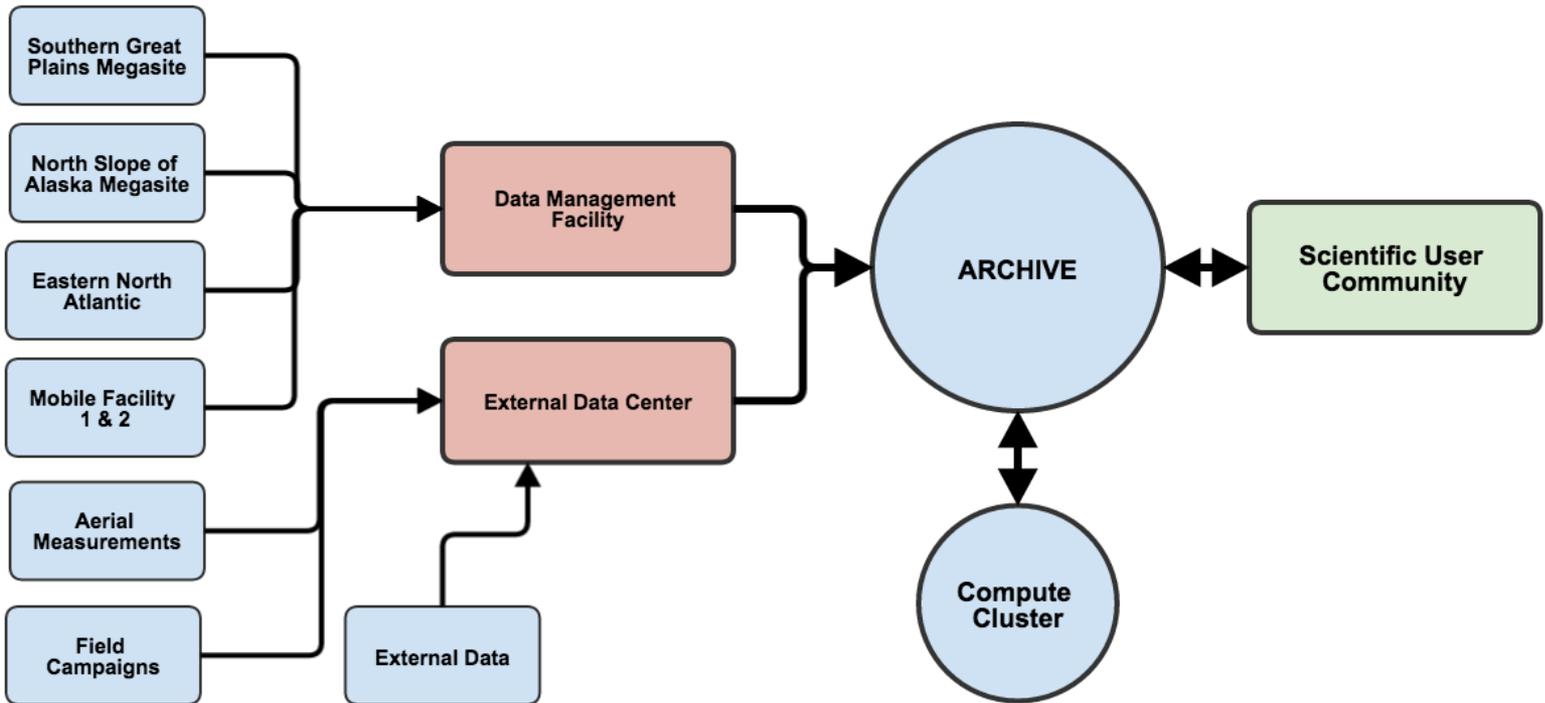
---

- Overview
- Next generation data challenges
- Data Operations Update
- Current Activities
- Planned/Upcoming Activities

# ARM Data Flow – The Big Picture

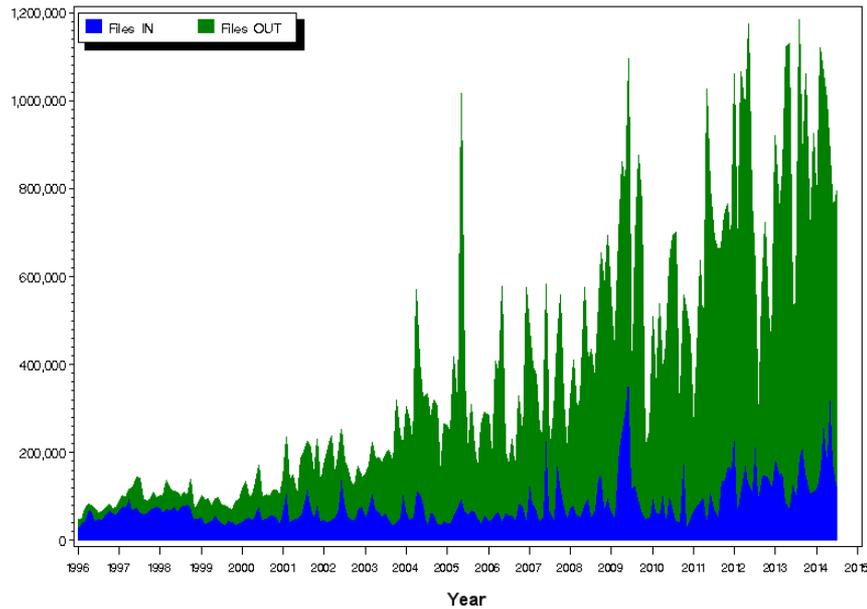
## ARM Data Operations Group

### Research Sites

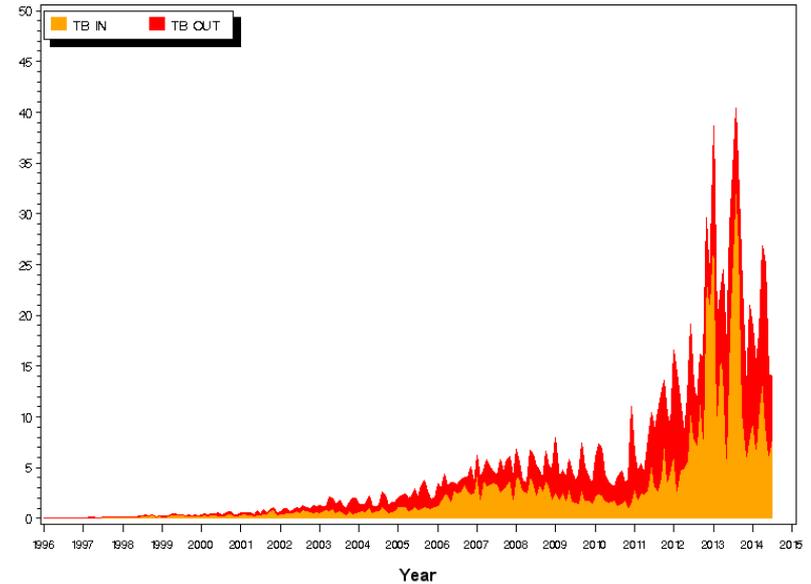


# ARM Data by the Numbers

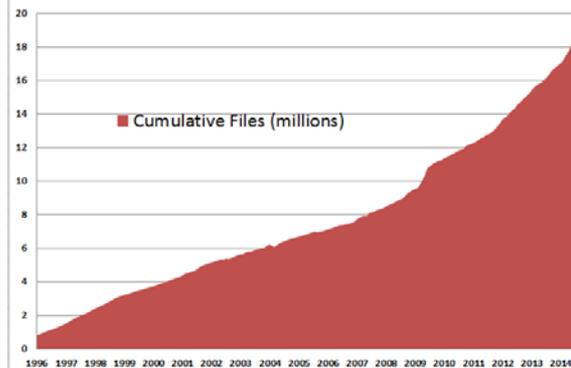
Number of Filenames In/Out per month



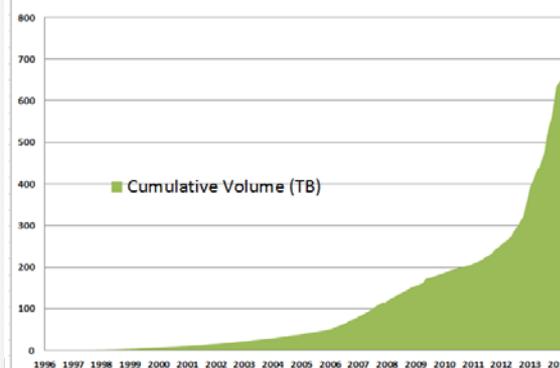
TB In/Out per month



History of Cumulative Number of Files Stored in ARM Archive

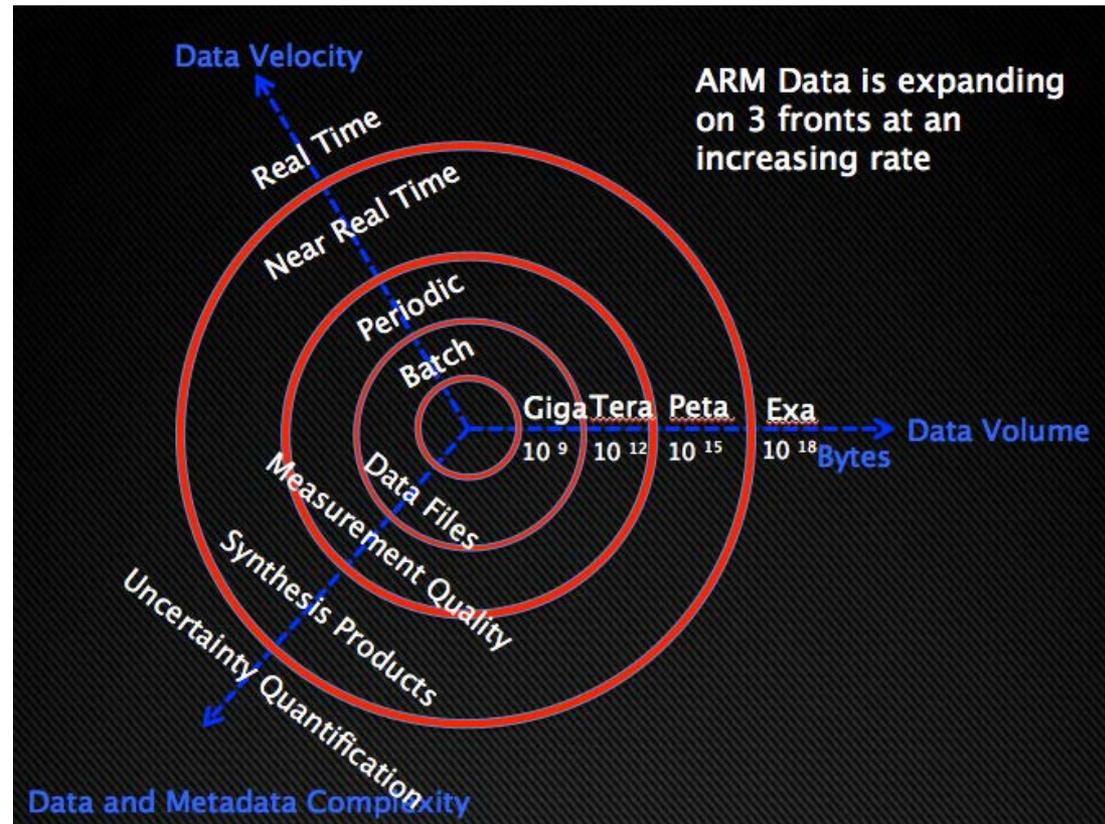


History of Cumulative Terabytes (TB) Stored in ARM Archive



# Next Generation Data Challenges

- ARM data – both volume and complexity
- Improve the ARM user experience
- Increased interest by larger user community to use ARM data
- Provide interoperable data that could be used in different applications at different scales



Prepared by: Jimmy Voyles

# Improving ARM Data Lifecycle Architecture and Tools

## ARM Data Strategy Team



# ARM Data Center – User Resources



- Shared storage resource (~ 250 TB)
- Large online data collection (data/archive) for immediate access (~ 100 TB)
- Custom staging of large datastreams with specific retention times (data/project)

- Visualization and software Development Cluster
- ARM Data Processing cluster
- GPU Computing Cluster

Breakout session: Next Generation ARM Data Analysis Tools - Mather, Sivaraman, Palanisamy [Ash Grove – Wednesday, 1.30 pm]

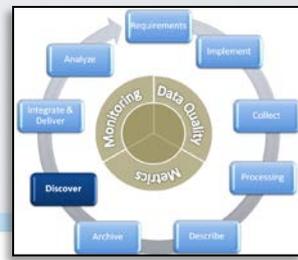
# Data Discovery Tool - Update

- Powerful data search capability to find and access ARM regular, PI and Field Campaign Data products
- Provides data availability in a timeline graphics
- Seamless access to data quality and data plots
- Provides options for data extraction and filtering based on data quality

The screenshot displays the ARM Data Discovery web application. The search term 'aerosol optical depth' is entered in the search bar. The interface is divided into several sections:

- Left Sidebar:** Contains navigation menus for 'CATEGORIES' (Radiometric, Aerosols, Cloud Properties, Atmospheric State, Surface Properties), 'MEASUREMENTS' (Aerosol optical depth, Shortwave narrowband diffuse downwelling irradiance, etc.), 'SITES', 'FIELD CAMPAIGNS', and 'SOURCES TYPES'.
- Search Results:** Shows a list of measurement entries with details such as 'AEROSOLBESTURN s1 @ SGP C1 // AEROSOL BEST ESTIMATE, FROM 1ST TURNER ALGORITHM'. Each entry includes a timeline visualization with green bars indicating data availability. A red arrow points to the 'ROUTINE DATA' button in the search results area.
- Timeline:** A horizontal timeline at the bottom of the search results shows data availability from 1997 to 2014. A red arrow points to the '2003-02-27' date marker.
- Search Results Panel (Right):** Displays 'Search Results' for 'aerosol optical depth'. It includes filters for 'DATA UNRELIABLE', 'DATA QUESTIONABLE', 'DATA MISSING', and 'DATA NOTE'. A red arrow points to the 'Aerosol optical depth' filter.
- Data Selection Summary (Bottom Right):** A pop-up window showing details for selected data, including '30EBBR b1 @ SGP E32' and '30ECOR b1 @ SGP E33'. It lists measurement variables and provides options for data extraction.

# Data Plots and Data Quality Reports – Helping Data Discovery



**ARM**  
CLIMATE RESEARCH FACILITY

Home | People | Site Index

U.S. DEPARTMENT OF ENERGY  
Office of Science

About Science Campaigns Sites Instruments Measurements Data News Publications Education

06/01/2009 06/30/2009 Back to Previous Page

### Data Plots for Millimeter Wavelength Cloud Radar

**Site & Facility:** North Slope Alaska(nsa) - Central Facility, Barrow AK (C1)  
**Instrument:** Millimeter Wavelength Cloud Radar  
**Datastream:** nsammcrmomC1.b1

Select a Measurement to View:

- MMCR Mean Doppler Velocity, Time-Height field, Stratus mode
- MMCR Mean Doppler Velocity, Time-Height field, Cirrus mode
- MMCR Mean Doppler Velocity, Time-Height field, General mode
- MMCR Mean Doppler Velocity, Time-Height field, Precipitation mode
- MMCR Reflectivity, Time-Height field, Stratus mode

Click a Thumbnail for Larger View and DQ Information

Previous Month Next Month

20090601 20090602 20090603 20090604 20090605  
 20090606 20090607 20090608 20090609 20090610  
 20090611 20090612 20090613 20090614 20090615  
 20090616 20090617 20090618 20090619 20090620

Data Quality Information and Quick Look for nsammcrmomC1.b1 (06/15/2009)

Close this window Back to previous page

> Data Quality Color: (show/hide)

Measurement	Auto QC Color
MMCR Mean Doppler Velocity, Time-Height field, Stratus mode	<input type="checkbox"/>
MMCR Mean Doppler Velocity, Time-Height field, Cirrus mode	<input type="checkbox"/>
MMCR Mean Doppler Velocity, Time-Height field, General mode	<input type="checkbox"/>
MMCR Mean Doppler Velocity, Time-Height field, Precipitation mode	<input type="checkbox"/>
MMCR Reflectivity, Time-Height field, Stratus mode	<input type="checkbox"/>
MMCR Reflectivity, Time-Height field, Cirrus mode	<input type="checkbox"/>
MMCR Reflectivity, Time-Height field, General mode	<input type="checkbox"/>
MMCR Reflectivity, Time-Height field, Precipitation mode	<input type="checkbox"/>

■ Missing ■ Undetermined ■ Review Pending ■ Good ■ Suspect ■ Incorrect

**Data Quality Report:**

- D080430.1 : NSA/MMCR/C1 - Reflectivities too high (Details)

**Data Quick Look:**

nsammcrmomC1.b1.20090615.000011.cdf

Light (km) 1 Reflectivity Mode 1 Velocity Mode 1 Spectral Width

**Data Quality Report**

**DQRID : D080430.1**

Start Date	Start Time	End Date	End Time
02/13/2008	0215	Open Ended	-1

**Subject:** NSA/MMCR/C1 - Reflectivities too high

**DataStreams:** nsammcrmomC1.b1, nsammcrspecmomC1.a0

**Description:** Comparison with the Canadian NRC radar's aboard their Convair seem to indicate that the Barrow MMCR is reporting reflectivity approximately 10 dBZ higher than they should be.

We are currently analyzing the radar's calibration and trying to figure out where this offset may be coming from.

If using reflectivity values, 10 dBZ should be subtracted. I estimate that this should be accurate to within +/- 2 dBZ. Once we find the cause and can come up with an accurate offset, this DQR will be revised and a request to reprocess the data will be submitted.

**Suggestions:**

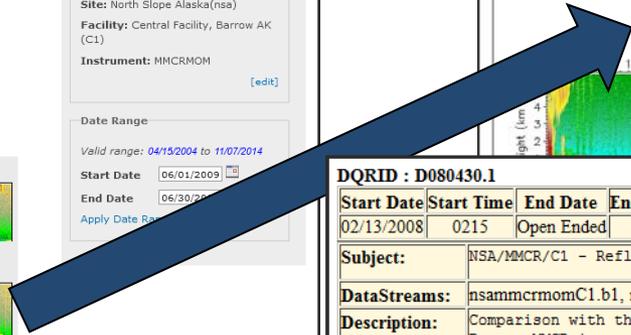
Subtract 10 dBZ from reflectivity values

**Measurements:** nsammcrmomC1.b1:

- MMCR Reflectivity(Reflectivity)

nsammcrspecmomC1.a0:

- MMCR Reflectivity(Reflectivity)



# ARM Data Integrator (ADI)

- ADI allows user to merge and consolidate ARM data
- All ingests and VAPs at the ADC are being migrated to the ADI Framework
- Archive is currently working with ADI team to integrate this as part of the data ordering process

Data Selection Summary

Signed in as gp8.

299 file(s) // 15 MB

2014-01-01

2014-10-19

30EBBR b1 @ SGP E13

Order Complete Datastream  Extract Specific Measurements

Measurement: Sensible heat flux

Variable: Heat flux, sensible, at 1.5-m height, 30-min intervals // sensible\_heat\_flux

2014-01-01

2014-10-19

306 file(s) // 593 MB

MFRSR b1 @ SGP E13

Order Complete Datastream  Extract Specific Measurements

Measurement: Shortwave narrowband total downwelling irradiance

Variable: Radiation, shortwave, hemispheric irradiance, 500 nm wavelength // hemisp\_narrowband\_filter2

2014-01-01

2014-10-19

Consolidate data?  No  Yes

Primary source

File format

oints from DQR(s) known to be

30EBBR b1 @ SGP E13

MFRSR b1 @ SGP E13

QCRAD1LONG c1 @ SGP E13

CANCEL SUBMIT DATA REQUEST

**ADI Tutorial - Sivaraman [Potomac – Today, 12.30 pm]**

# ARM Data Product Registration and Submission Form (OME)

ARM CLIMATE RESEARCH FACILITY

Data Product Registration and Submission Tool

U.S. DEPARTMENT OF ENERGY Office of Science

\*This record has been approved

The tool only supports ASCII characters and UTF-8 formatting. If you are copying text from other documents, please, avoid any special characters and other formats. But, HTML tags like hyperlinks are supported

Back Preview Metadata \* = required [ logout ]

**Data Product**

First, within the ARM program, is this data product/set only associated with ARM field campaign?  
Answer 'yes' to this question if the dataset was collected for an ARM field campaign. Answer 'no' if this is an independent dataset that did not use ARM infrastructure, or a product developed for regular ARM data, or for use with more than one field campaign.

Yes  No

Please select the ARM campaign:

New Particle Formation Study 2013 (NPF52013)

**Contact information about the author of this metadata record.**

Full Name (First Name followed by Last Name):\*  
Begin typing a letter (slowly) to see the list of existing fields. This list is supplied from arm database.

Smith, James ID

Organization Name: \*  
NCAR

Position Name:

Telephone Number: \*  
3034971468

Email: \*  
jimsmith@ucar.edu

Address:  
Street: PO Box 3000

- Data Type
- Description and keywords
- Contact information
- Data Quality
- When and Where
- Related Citations
- Analytical Tools
- Save, revisit and Submit

# ARM Data Citation Service

## Benefits:

- Allow users to cite exact ARM data used in their research/publication
- Allow ARM to provide proper data citation credits to the PIs and collaborators
- Allow future data users and the project to easily track the data used in various articles

Need DOIs for your Data?  
- Contact ARM Data Center

## Example:

*Atmospheric Radiation Measurement (ARM) Climate Research Facility. 1994, updated daily. SONDEWNPN. Oct. 2010–March 2011, 36° 36' 18.0" N, 97° 29' 6.0" W: Southern Great Plains Central Facility (C1). Compiled by R Coulter, J Prell, M Ritsche, and D Holdridge. ARM Data Archive: Oak Ridge, Tennessee, USA. Data set accessed 2011-04-13 at <http://dx.doi.org/10.5439/1021460>.*

The screenshot shows the ARM Data Center website interface. The top navigation bar includes 'Home', 'People', 'Site Index', and a search bar. The main content area displays the 'Datastream : SONDEWNPN' page, which includes a description of the Balloon-borne sounding system (BBSS), active dates (1994.04.12 - 2014.11.07), and measurement categories. A red circle highlights the DOI: 10.5439/1021460 and the 'GENERATE CITATION' button. The bottom of the page shows a 'Measurements' section with a table of variables and units.

The screenshot shows the 'Generate Citation' dialog box on the ARM Data Center website. The dialog box is titled 'Generate Citation' and contains the following information:

- Author:** Atmospheric Radiation Measurement (ARM) Climate Research Facility
- Original Publication Date:** 1994
- Update Period:** hourly
- Location Accessed:** Site: Summit Station, Greenland (SMT); Facility:  External Data (satellites and others) (X1)
- Dates Used:** Start: [ ] End: [ ]
- Editor(s) or Compiler(s):** D. Holdridge, J. Kyrouac and R. Coulter
- Date Accessed:** 2014-11-10
- Citation(s):** Atmospheric Radiation Measurement (ARM) Climate Research Facility. 1994, updated hourly. Balloon-Borne Sounding System (SONDEWNPN). << start date used >> to << end date used >>, Summit Station, Greenland (SMT) << facility >>. Compiled by D. Holdridge, J. Kyrouac and R. Coulter. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed 2014-11-10 at <http://dx.doi.org/10.5439/1021460>

The dialog box also includes a 'Remove Modify' button, a 'Send me a copy' checkbox, and a 'DONE' button.

# Machine Readable DQR Web Service

## Simple Steps to Use Web Service:

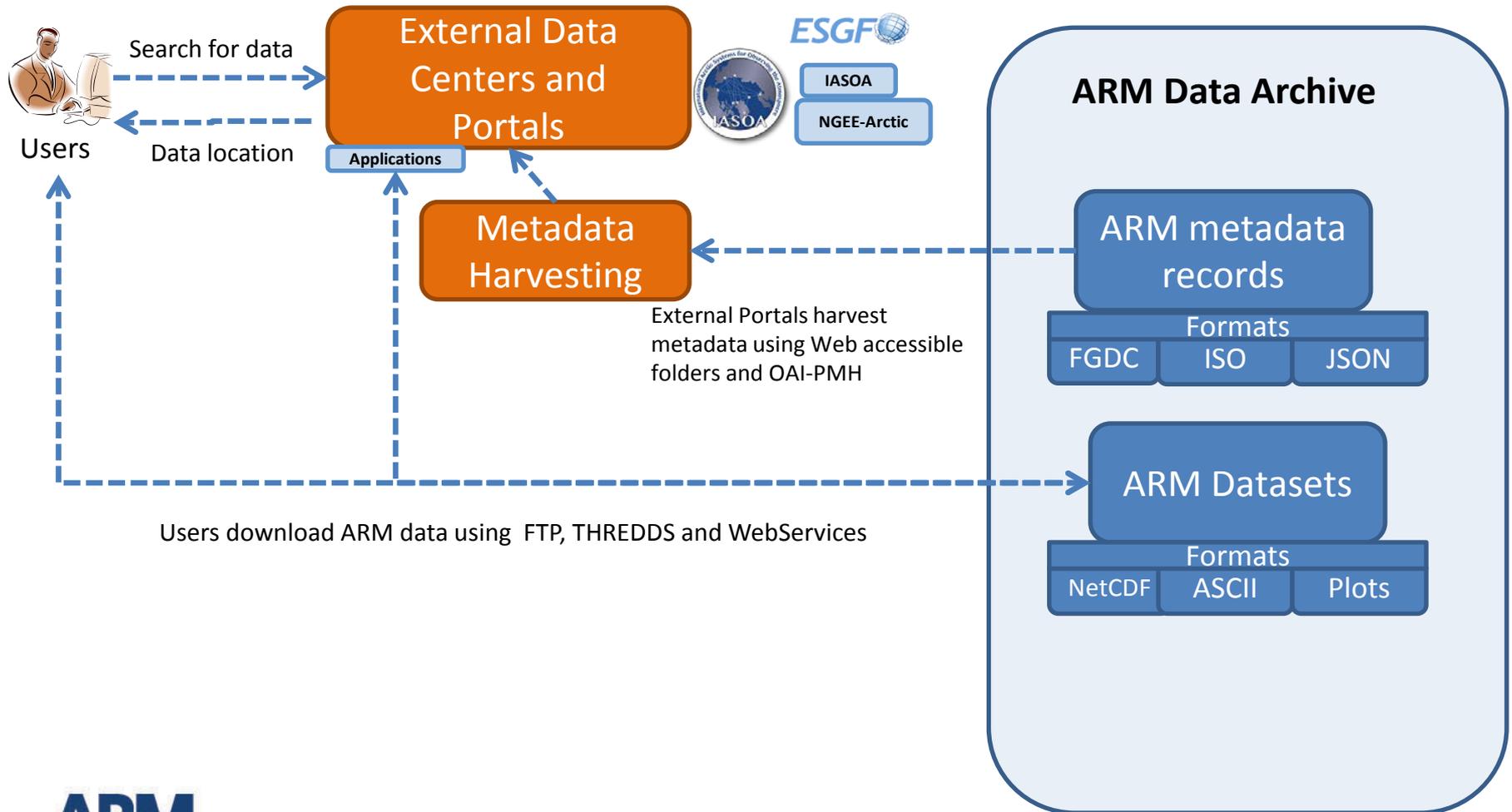
- Build query with at least these two parameters:
  - Datastream
  - varname
- Send the query and check return status code for success
- If status 'okay', parse timeblocks from the response
- Discard or flag data that web service indicates has failed your criteria

[http://www.archive.arm.gov/dqrws/ARMDQR?datastream=nsa\\_mfrsrC2.b1&varname=direct\\_horizontal\\_broadband&dqrfields=dqid,starttime,endtime,metric,subject&timeformat=yyyymmdd.hhmmss&searchmetric=missing,suspect](http://www.archive.arm.gov/dqrws/ARMDQR?datastream=nsa_mfrsrC2.b1&varname=direct_horizontal_broadband&dqrfields=dqid,starttime,endtime,metric,subject&timeformat=yyyymmdd.hhmmss&searchmetric=missing,suspect)

D070813.1|20070601.000000|20070831.235900|suspect|NSA/MFRSR/C2 - Instrument shading problem  
D050811.4|20050808.050000|20050810.210000|suspect|NSA/MFRSR/C2 - Intermittent data  
D041028.4|19990912.235900|20000417.235900|missing|NSA/MFRSR/C2 - No data collected during winter  
D041028.4|20001124.232200|20010504.225900|missing|NSA/MFRSR/C2 - No data collected during winter  
D041028.4|20071124.230200|20080202.010200|missing|NSA/MFRSR/C2 - No data collected during winter  
D041028.4|20031111.011800|20040331.215300|missing|NSA/MFRSR/C2 - No data collected during winter

<http://www.archive.arm.gov/dqrws/>

# ARM Metadata and Data Sharing With Other Portals





# Other Key Updates

## **New Hires:**

- ARM Metadata Team @ BNL hired David Troyan as a metadata manager
- ARM Data Center @ ORNL hired a Post Doc (Dr. Bhargavi Sriram) to help with data operations

## **ADC Infrastructure:**

- Developing whitepaper for Mega Site data lifecycle plan including the interface with High resolution modeling pilot program (with Data Strategy Team)
- Site Data System
  - Successfully deployed a new data storage system for ACAPEX
  - Exploring new storage solution for all sites
- DMF successfully processed ACAPEX data
- Archive is migrating to new open source Database (PostGresql)
- Metadata team:
  - Completed ARM Database Analysis
  - Developing ARM Surface Characterization Database

# Thanks!

Please come and talk to us ☺

## Lunch Tutorials:

- Welcome to ARM for New Investigators:
  - Part I - Mather [Great Falls] today at 12.30 pm
  - Part II, Data Services - Palanisamy [Great Falls] Wednesday 12.30 pm

## Posters:

- Next Generation ARM Needs Next Generation Website (Rolanda Jundt)
- Registering and Submitting Your PI Data with Ease Using the ARM Data Product Registration Tool (OME) (Biva Shrestha)
- Machine-Readable Data Quality Reports (Sean Moore)

We all need people who will give us feedback. That's how we improve.

Bill Gates