Persistent Identifiers - ARM Use Case

HAROLD SHANAFIELD, CHIRAG SHAH, GIRI PRAKASH

ARM Data Center, Oak Ridge National Laboratory
shanafieldha@ornl.gov

FAIR Facilities and Instrument Workshop, September 13, 2023

Ground-Based Atmospheric Observing Facility

ARM

Atmospheric Radiation Measurement Facility

Since 1992, the world-leading facility for measurements of cloud & aerosol properties, & their impacts on Earth's energy balance

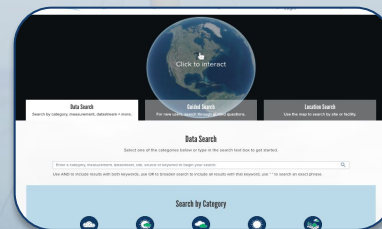Comprehensive measurements across diverse climate regimes

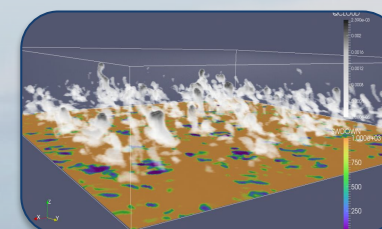Network of 3 fixed-location & 3 mobile observatories

Piloted & unmanned aerial measurement platforms

Extensive data management infrastructure

Freely available data products to support atmospheric research & model development

Large-eddy simulation (LES) model simulations and analysis tools

Serves the international climate research community and has close collaboration with Atmospheric System Research (ASR)

Source: Jim Mather

Comprehensive Sets of Instruments Deployed in Diverse Climate Regimes

- Background atmospheric state
- Surface energy balance
- Aerosol and hydrometeor profiles
- Near-surface aerosol properties
- Aerial measurements

Legend:
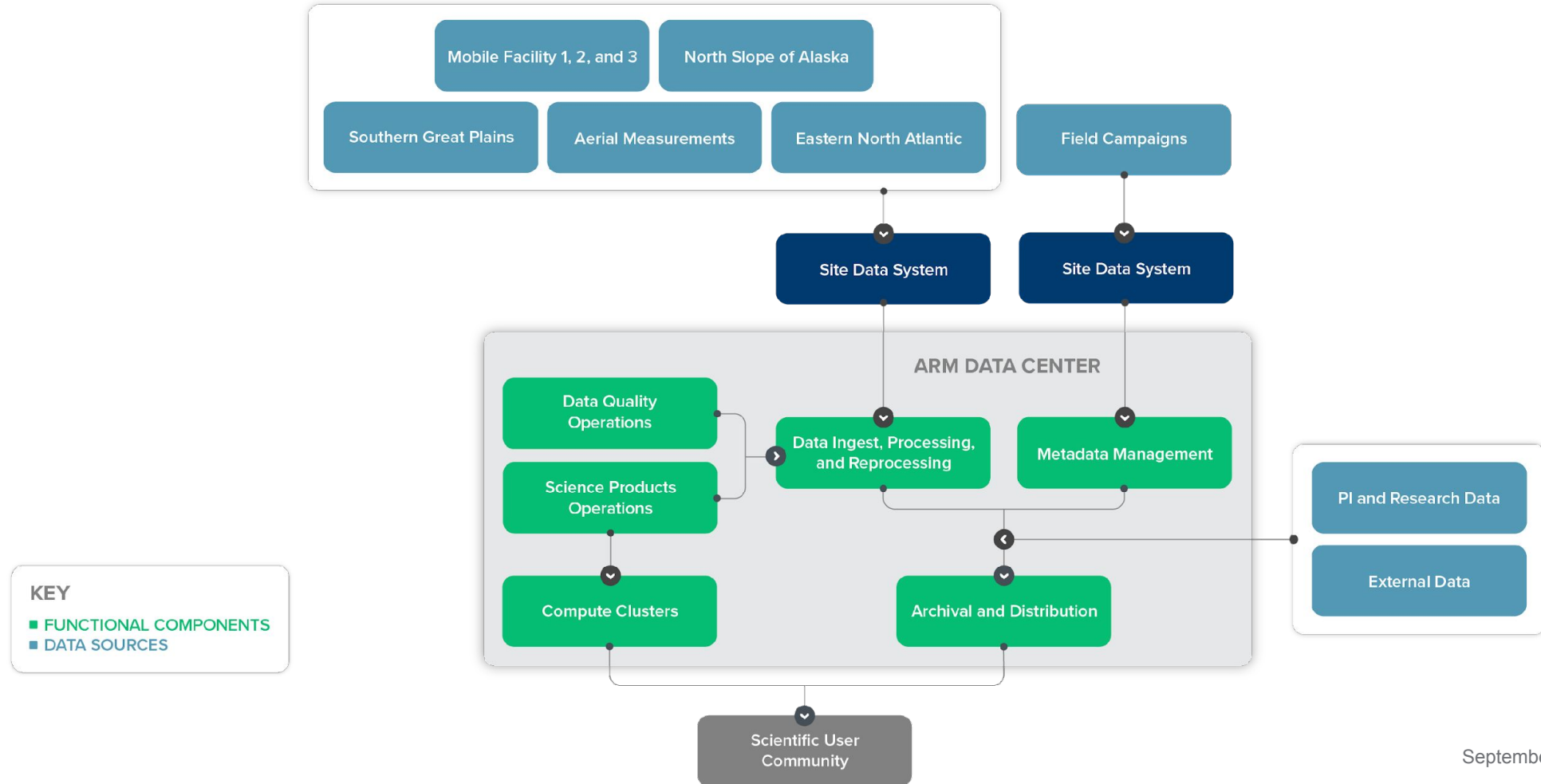- Fixed Site
- Past Fixed Site
- Deployments
- Aerial Deployment
- Ship Deployment

# ARM Data Flow: From Collection to Distribution

- Offers powerful and adaptable infrastructure capabilities to support a wide range of data pipeline requirements, enabling efficient and streamlined processing of data from various sources.

About The ARM Data Center

Provides a robust integrated data and computing ecosystem to advance understanding of atmospheric measurements

- Data flow operations and monitoring
- Advanced data collection systems
- High-performance computing (HPC)
- Comprehensive Data Processing
- Data Interoperability:
  - Advanced strategies for utilizing metadata
  - Data Discovery
  - Data workbench
  - FAIR, Standards, and Protocols
- User Management and Citations
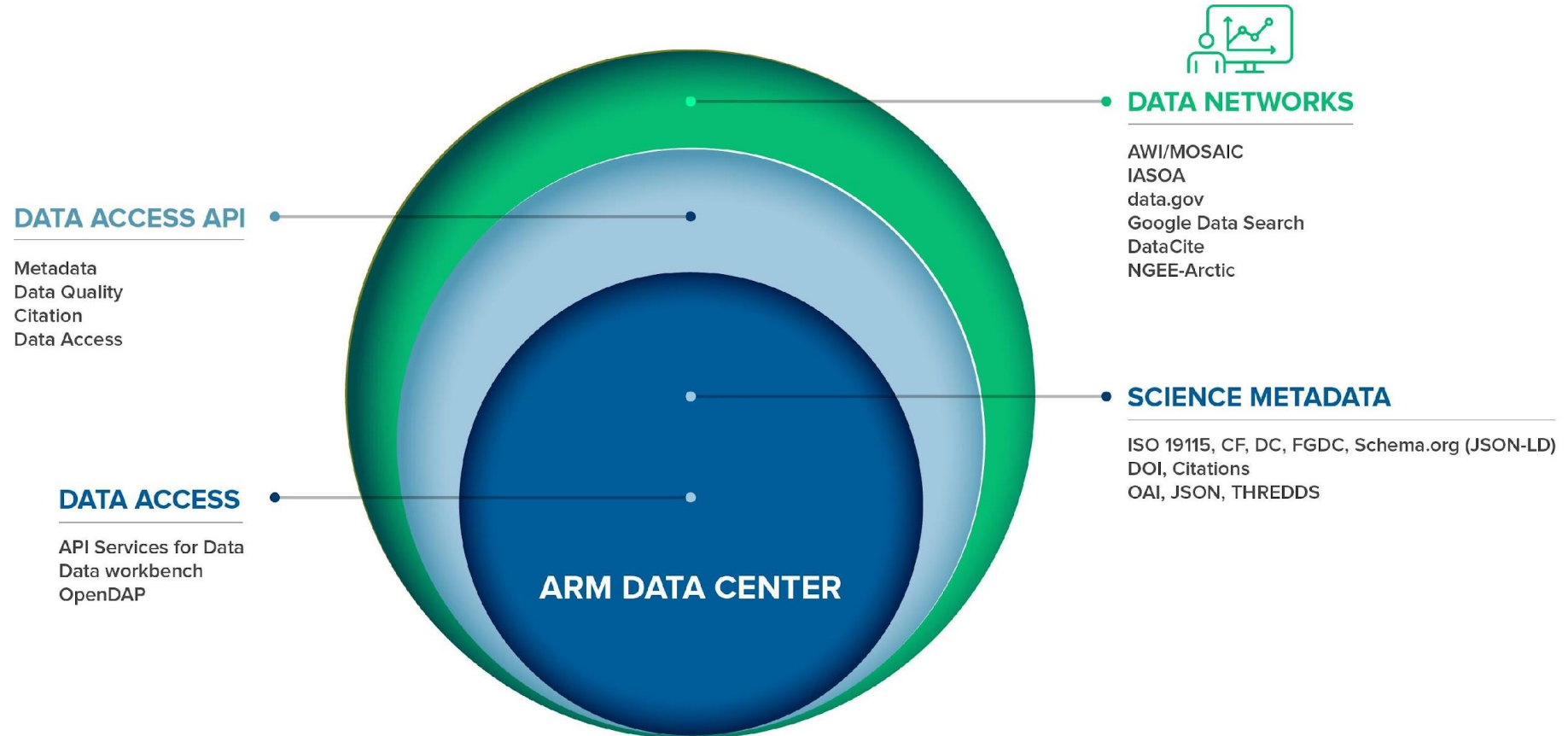- AI-based approach in data management
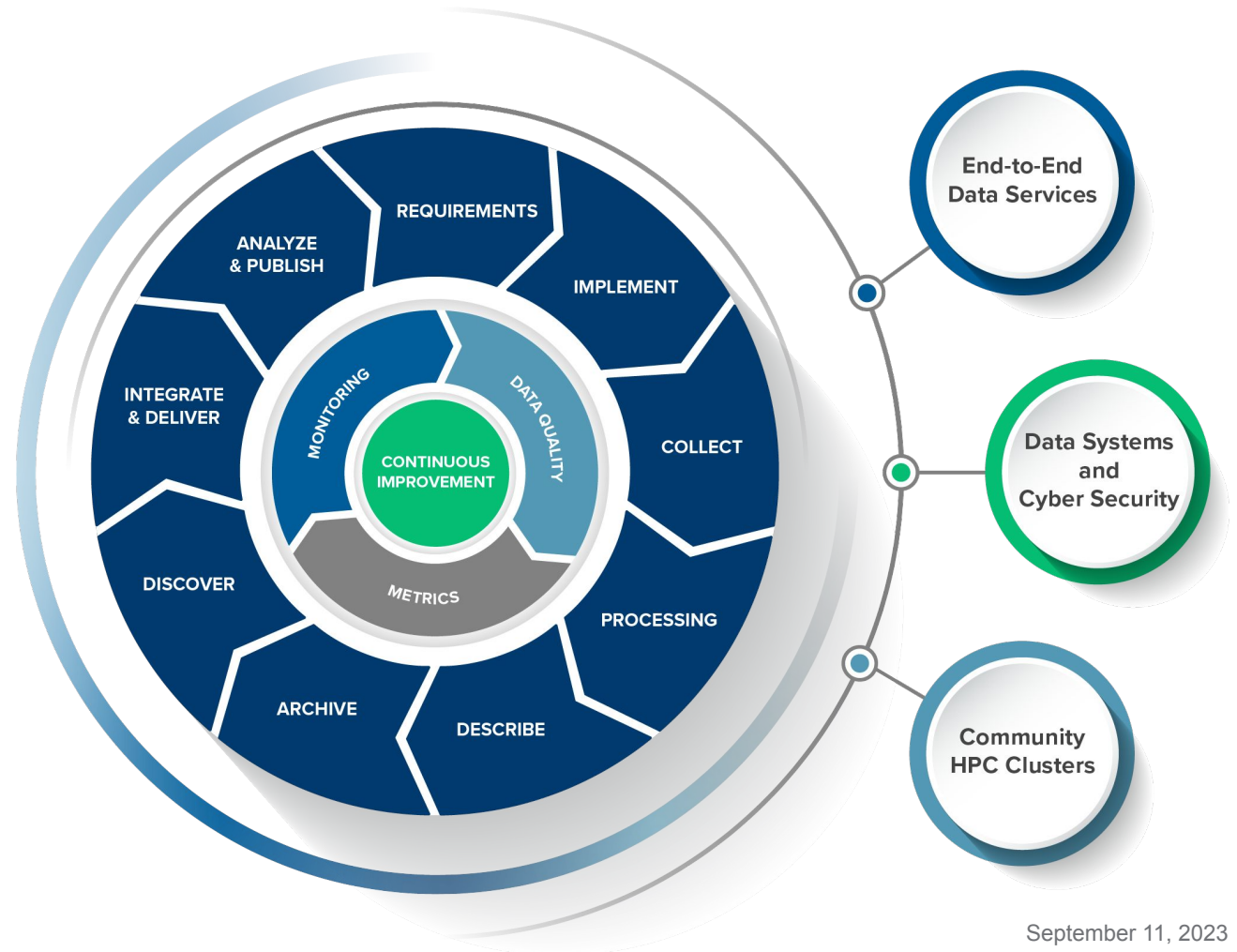
FAIRness Assessment and Community Engagement
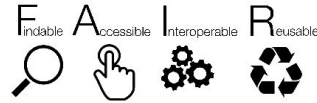
- Review of data management capabilities and obtaining certifications

- Continuous collaboration with broader data networks

- Active contribution to national and international working groups

Putting FAIR Principles into Practice: Standards and Protocols in Data Interoperability

DATA NETWORKS

AWI/MOSAIC
IASOA
data.gov
Google Data Search
DataCite
NGEE-Arctic

DATA ACCESS API

Metadata
Data Quality
Citation
Data Access

SCIENCE METADATA

ISO 19115, CF, DC, FGDC, Schema.org (JSON-LD)
DOI, Citations
OAI, JSON, THREDDS

DATA ACCESS

API Services for Data
Data workbench
OpenDAP

ARM DATA CENTER

Internal Persistent Identifier

- Instruments are assigned ARM Asset IDs
  - Asset IDs are generated at purchase, can be used in proposals
  - Data is linked specific instruments through Asset IDs
  - Plan to increase speed and efficiency through barcodes

- Data is identified through "Datastream" ID
  - Combination of site code, instrument code, facility code and qc level
  - Asset IDs are tied to Datastreams which are tied to ORCIDs
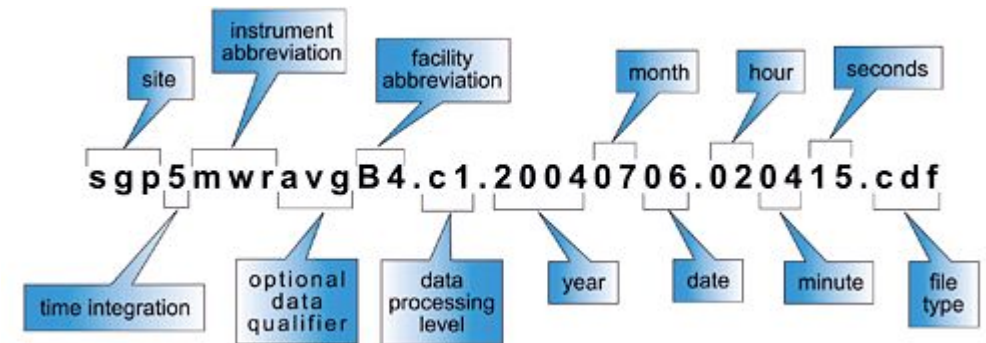
ARM Data Citation using DOIs

- Based on DOIs and citation guidance

- Enable metrics to quantify science impact and ensure data reproducibility

- Still evolving
  - Enabling time-based author credits
  - Additional citation formats
  - Nested citations
  - Data mashups

- https://www.arm.gov/working-with-arm/acknowledging-arm/doi-guidance-for-datastreams

https://www.mdpi.com/154276

Prakash, G, B Shrestha, K Younkin, R Jundt, M Martin, and J Elliott. 2016. "Data Always Getting Bigger—A Scalable DOI Architecture for Big and Expanding Scientific Data." *Data* 1:11.

Gathering ORCID

- Users are encouraged during ARM account creation to link ARM account to ORCID iD
- Linking is as easy as clicking a button in the user registration form and signing into ORCID iD
- Steps to create ARM account and link ORCID iD

  1. User begins ARM account creation process by completing fields in registration form
  2. User clicks "CREATE OR CONNECT YOUR ORCID ID" button in registration form
  3. User is redirected to ORCID login, where they sign into their ORCID iD
  4. Upon successful ORCID login, they are redirected back to the ARM registration form and have the option to auto-fill some fields using information from ORCID iD.

USER INFORMATION

Link ORCID record

ORCID provides a persistent digital identifier that distinguishes you from other researchers. Please link your ORCID record if you have one.

iD CREATE OR CONNECT YOUR ORCID ID

First name *

Middle name

Sign in

Email or 16-digit ORCID iD

example@email.com or 0000-0001-2345-6789

Password

••••••••••

SIGN IN

Forgot your password or ORCID ID?

Don't have an ORCID iD yet? Register now

or

Access through your institution

Sign in with Google

Sign in with Facebook

7479 Active user accounts.
2522 Active user accounts with an ORCID iD.

U.S. DEPARTMENT OF
ENERGY

Maximizing User Management and Data Citations using PIDs

- Integrating ORCID with other user metrics improves the program's ability to manage the quality of user details and metrics preparation

- Opportunities exist to improve user experience using AI/ML techniques
  - Discover relationships between ORCID identifiers, users, publications, data, metadata etc. Then use these relationships to improve the user experience with finding and using ARM Data