

Every Scientist Should Be Able to Collaborate with Every Scientific Facility and Every AI

A National Digital Backbone for Autonomous Scientific Discovery



CONNECT



SHARE



INTEGRATE



ANALYZE



DECIDE



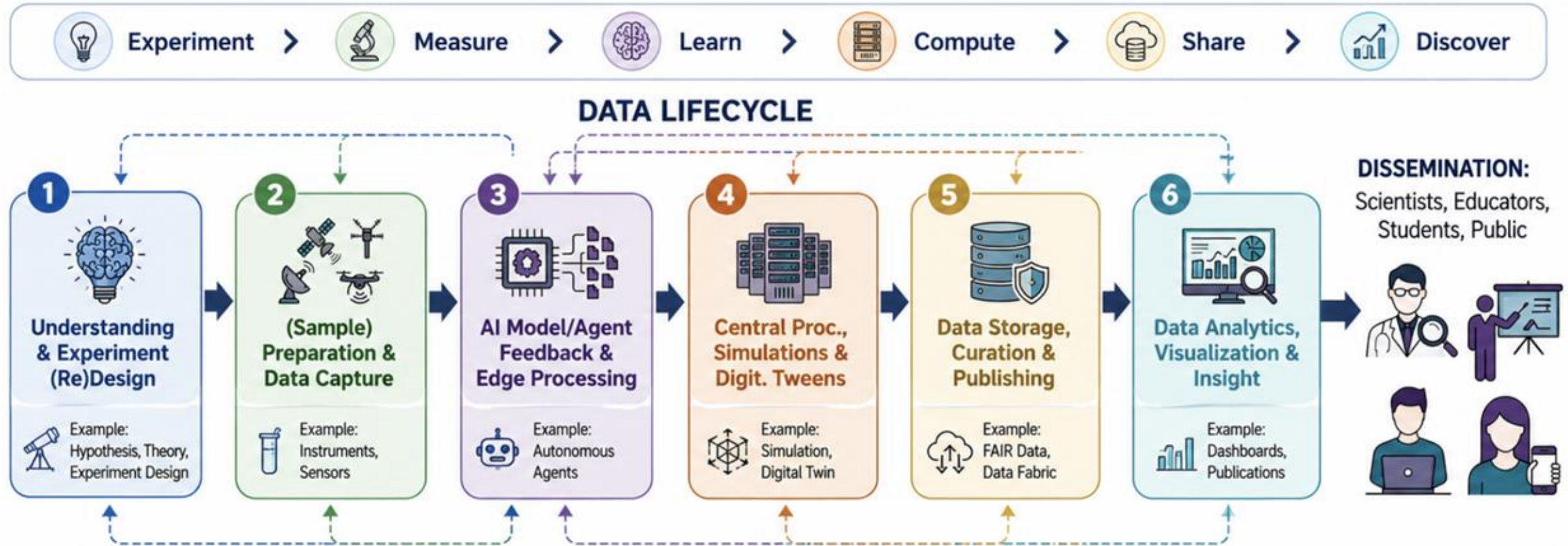
DISCOVER

Michela Taufer (**UTK**), Valerio Pascucci (**U Utah**), Alex Karargyris (**MLCommons**),
Carol Song (**Purdue**), Virginia Trueheart (**TACC**)



Science is Becoming **Autonomous**

AI is no longer limited to data analysis—it now guides experiments, orchestrates facilities, accelerates simulations, curates data, and disseminates scientific knowledge.



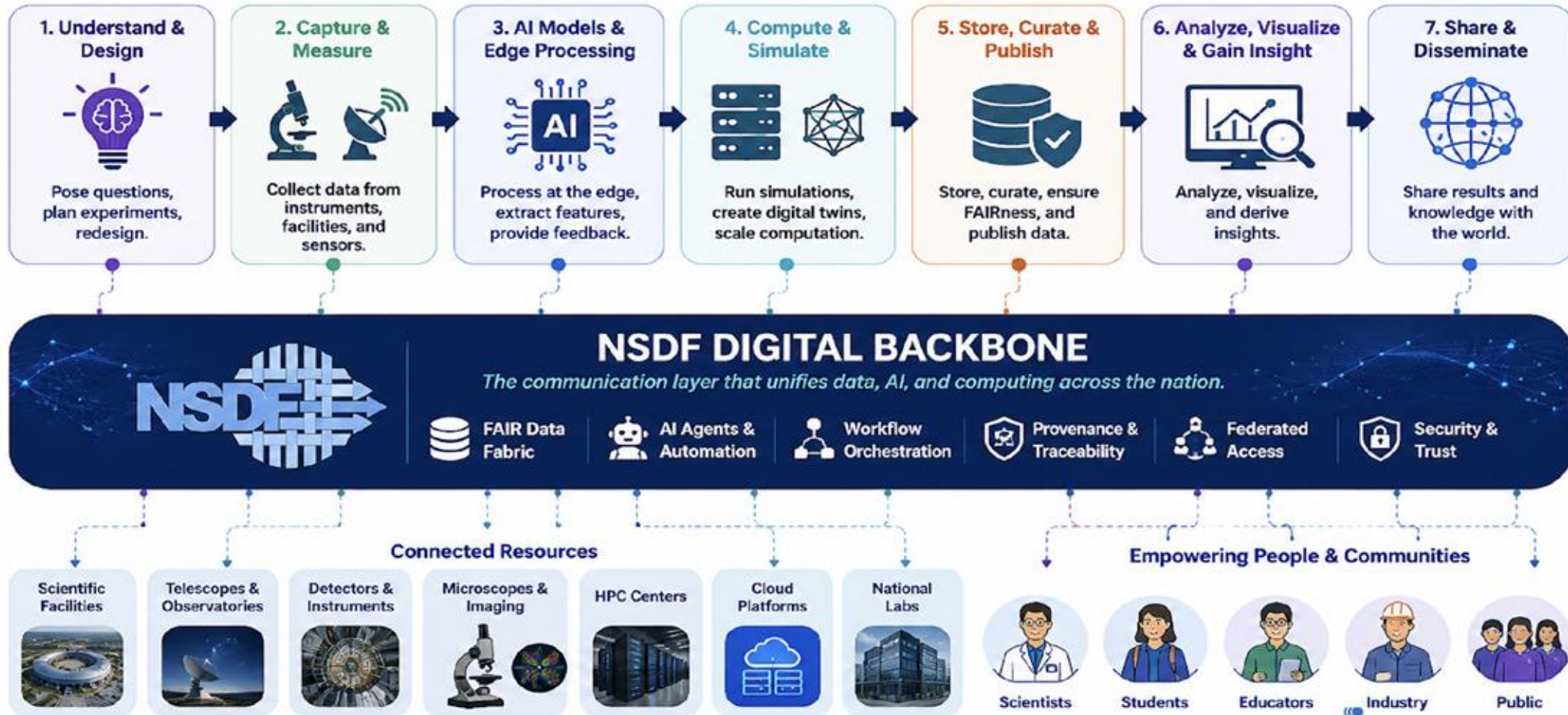
Every Scientist Should Be Able To Collaborate With Every Scientific Facility And Every AI

NSDF Connects the **Scientific Lifecycle**



2609465

One Federated Backbone. Every Facility. Every Scientist. Every AI.



*NSDF Is The **Connective Tissue** That Enables Autonomous Science Through Collaboration, AI And FAIR Data Across Every Step Of The Scientific Lifecycle.*

NSDF Is a Digital Backbone for **Autonomous** Science

2609465



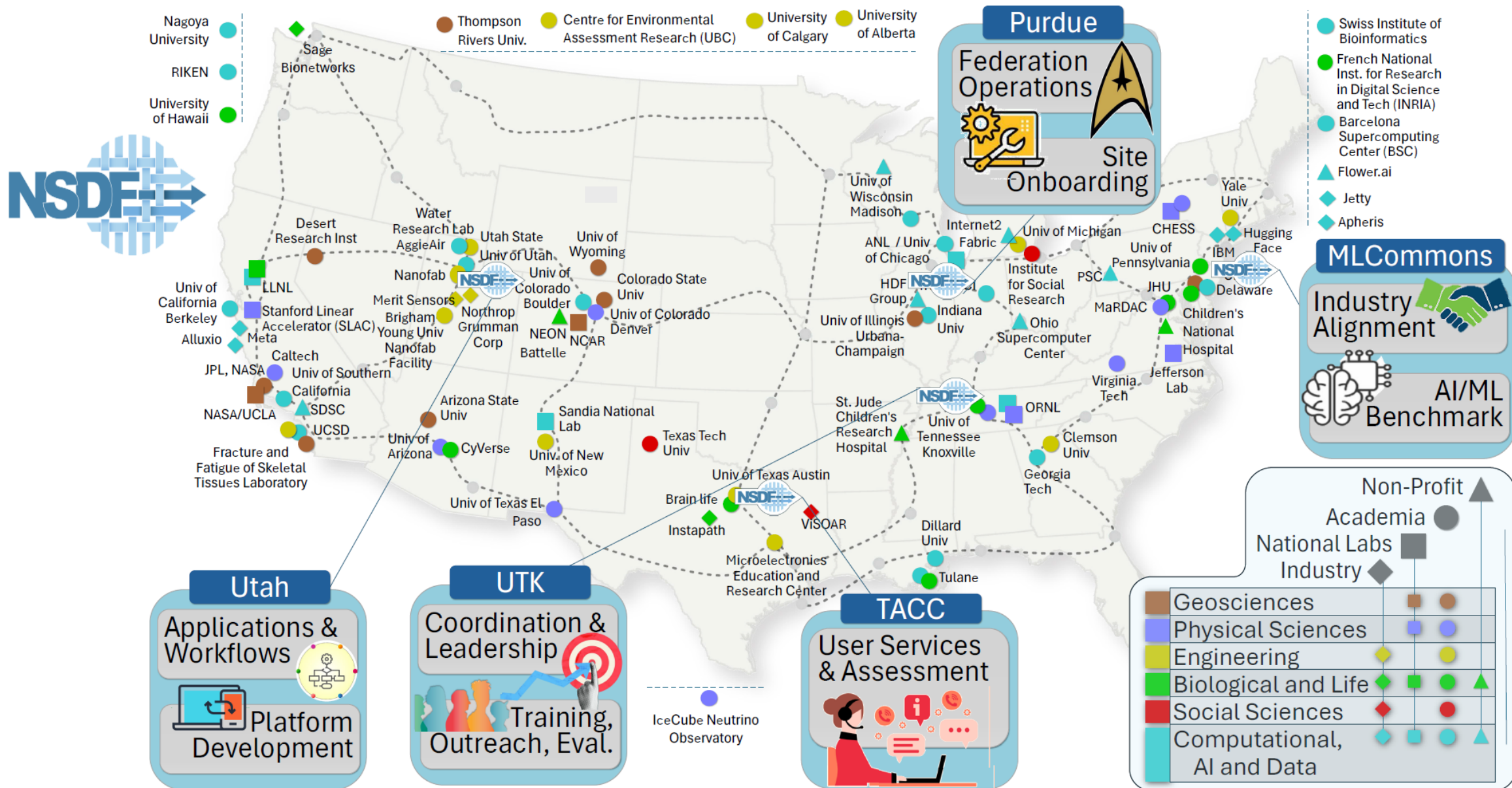
Multiple federated entry points. Secure. Intelligent. Interoperable.



*NSDF Builds The **Communication And Coordination Layer** That Connects Data, AI, People, And Computing Across The Nation.*

Support a **National Federation** of Data & Communities

2609465



Serving **Communities** one **Domain** at the Time

2609465



CHESS - Materials Science (NSF Facility)



- Real-time, high-flux X-ray analysis
- Uncovering materials properties
- Accelerating discovery of new compounds



ORNL - Neutron & Energy Sciences



- Unlocking structural secrets
- Materials for energy solutions
- Neutron scattering data management



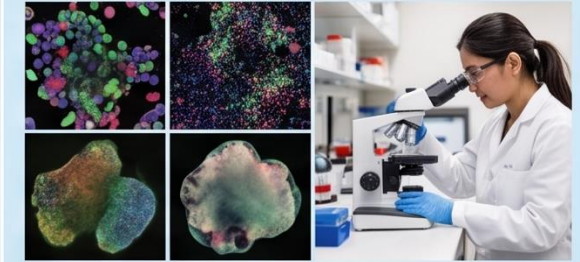
EARTH & ENVIRONMENTAL SCIENCE (NSF Support)



- Global environmental monitoring
- Climate modeling and prediction
- Ecosystem data integration



BIOMEDICAL & LIFE SCIENCES (NSF Innovation)



- Multi-omics data integration
- AI for drug discovery and genomics
- Enabling precision medicine
- Secure, privacy-preserving data sharing



INTERSECT

(AI & SYSTEMS)



Smart & Connected Labs Collaborative team on machine systems

- AI-human partnership in discovery
- Closed-loop experimentation
- Distributed computing for systems biology

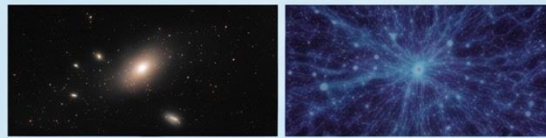


DARK MATTER PHYSICS

(NSF Programs)



- Direct detection of dark matter
- Distributed underground observatory data
- Federated analysis pipelines
- Global collaboration at scale

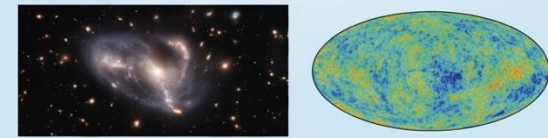


ASTROPHYSICS

(NSF Facilities)



- Petabyte-scale cosmic surveys
- AI for object classification
- Data from Vera Rubin & ALMA
- Unveiling the early universe



EDUCATION & WORKFORCE DEVELOPMENT

(NSF Grants)



- Training the next generation of scientists
- Accessible notebooks and cloud labs
- Workforce development for the digital economy
- Broadening participation in science



NSDF Learns From Each Domain, Identifies Reusable Services, And Strengthens The Common Backbone.

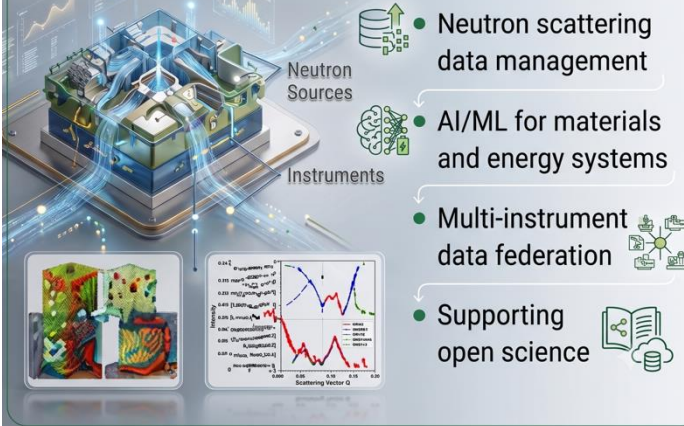
Science in Action: From Live Data Collection to **AI-Guided** Beamtime Decisions

2609465



ORNL – Neutron & Energy Sciences

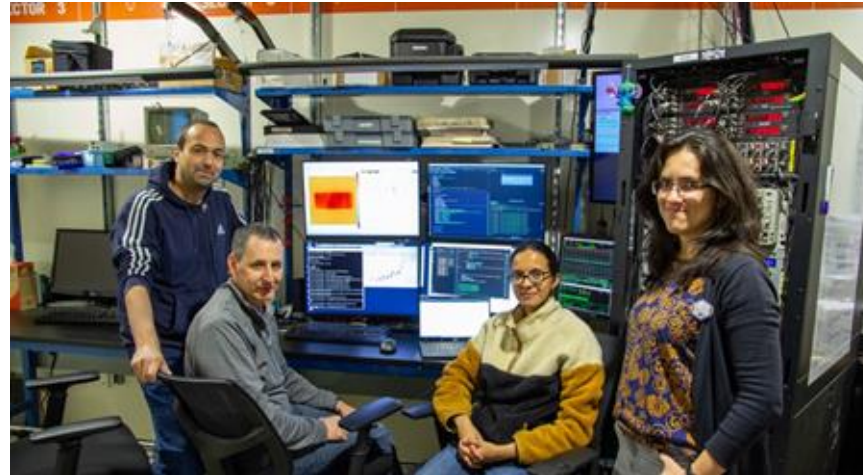
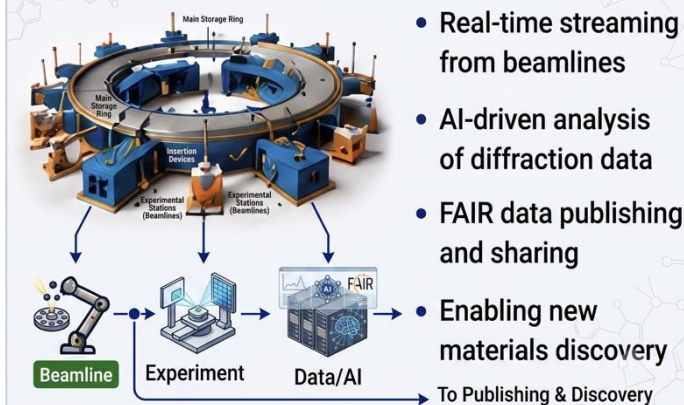
Driving Insights in Energy & Materials



CHESS – Materials Science



Accelerating Materials Discovery with Photon Science



University of Tennessee, Knoxville, Knoxville

Michela Taufer, Jack Marquez, Kin Hong NG

Oak Ridge National Laboratory, Oak Ridge

Marshall McDonnell, Ben Mintz, Rob Moore, Lance Drane, Gregory R. Watson, Kaz Gofron, Bogdan Vacaliuc, Zach Thurman, Gregory Cage, Gavin Wiggins, Cody Stiner, Jesse McGaha, Andrew Ayres, Robert Smith, Greg Watson, Addi Malviya Thakur, Yuanpeng Zhang, Jue Liu, Stephen DeWitt, Ankit Shrivastava, Paul Laliu, Craig Bridges, Mathieu Doucet, Matt Tucker, Emily R. Van Auken, Luke Daemen, Marie Backman, Darsh Dinger, Melanie Kirkham, Thomas Proffen

National Institute of Standards and Technology (NIST)

Austin McDannald, A. Gilad Kusne, William Ratcliff

University of Utah, Salt Lake City

Amy Gooch, ViSOAR LLC, Salt Lake City

Giorgio Scorzelli, Valerio Pascucci

NSDF & Facilities: Scientific Challenge & Objectives



2609465



THE CHALLENGE

Scientific data, models, instruments, and facilities are fragmented and operate in isolation. We collect data, but we cannot understand it in time to act.

OUR OBJECTIVES

Enable AI-guided, closed-loop experimentation Make real-time decisions during experiments to accelerate discovery.

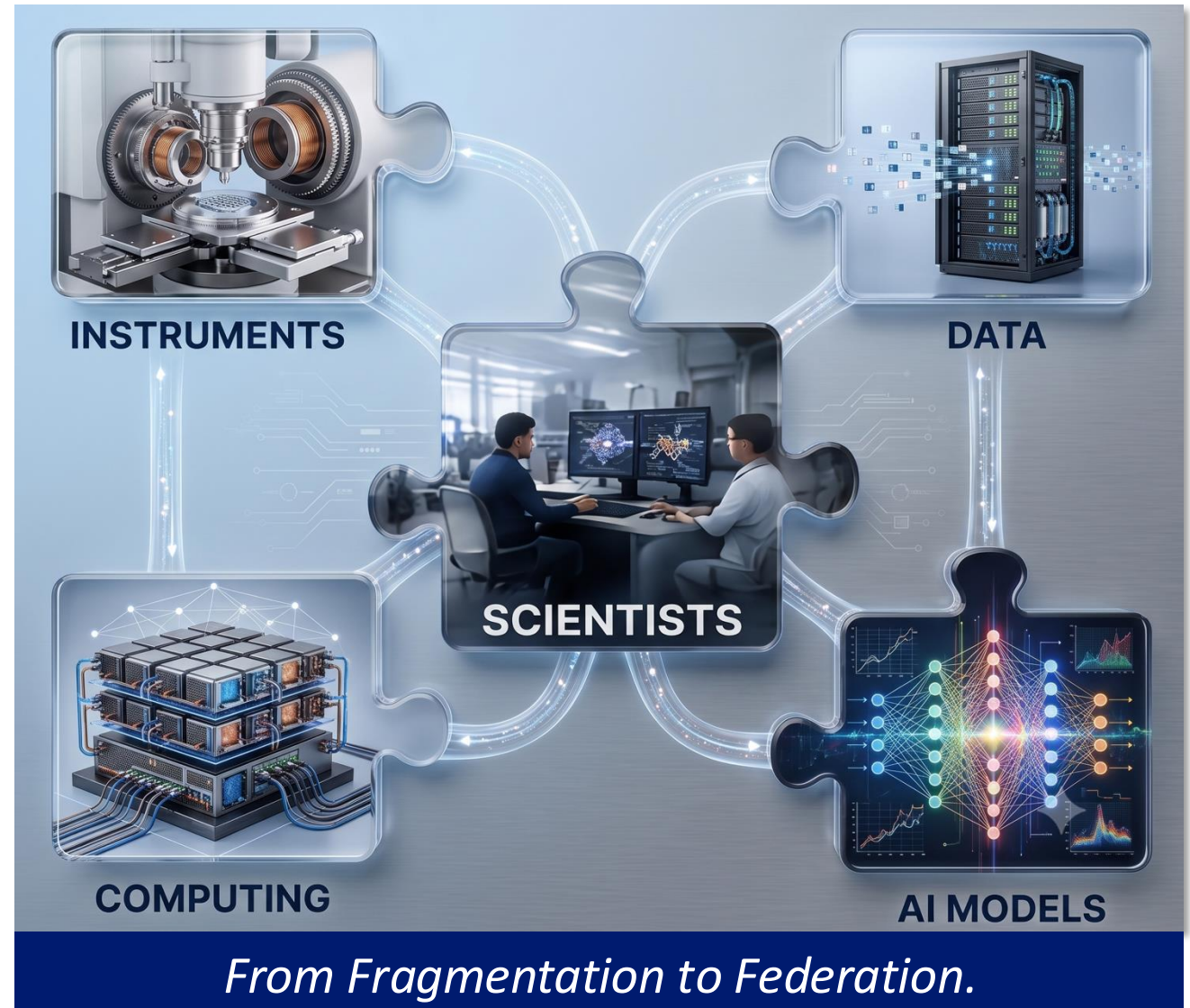


INTEGRATE INSTRUMENTS, DATA, AI, AND HPC

Create the missing integration layer across experiments, models, facilities, and prior data.

DEMOCRATIZE AUTONOMOUS SCIENCE

Provide open, scalable infrastructure so any researcher can leverage AI for discovery.



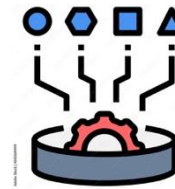
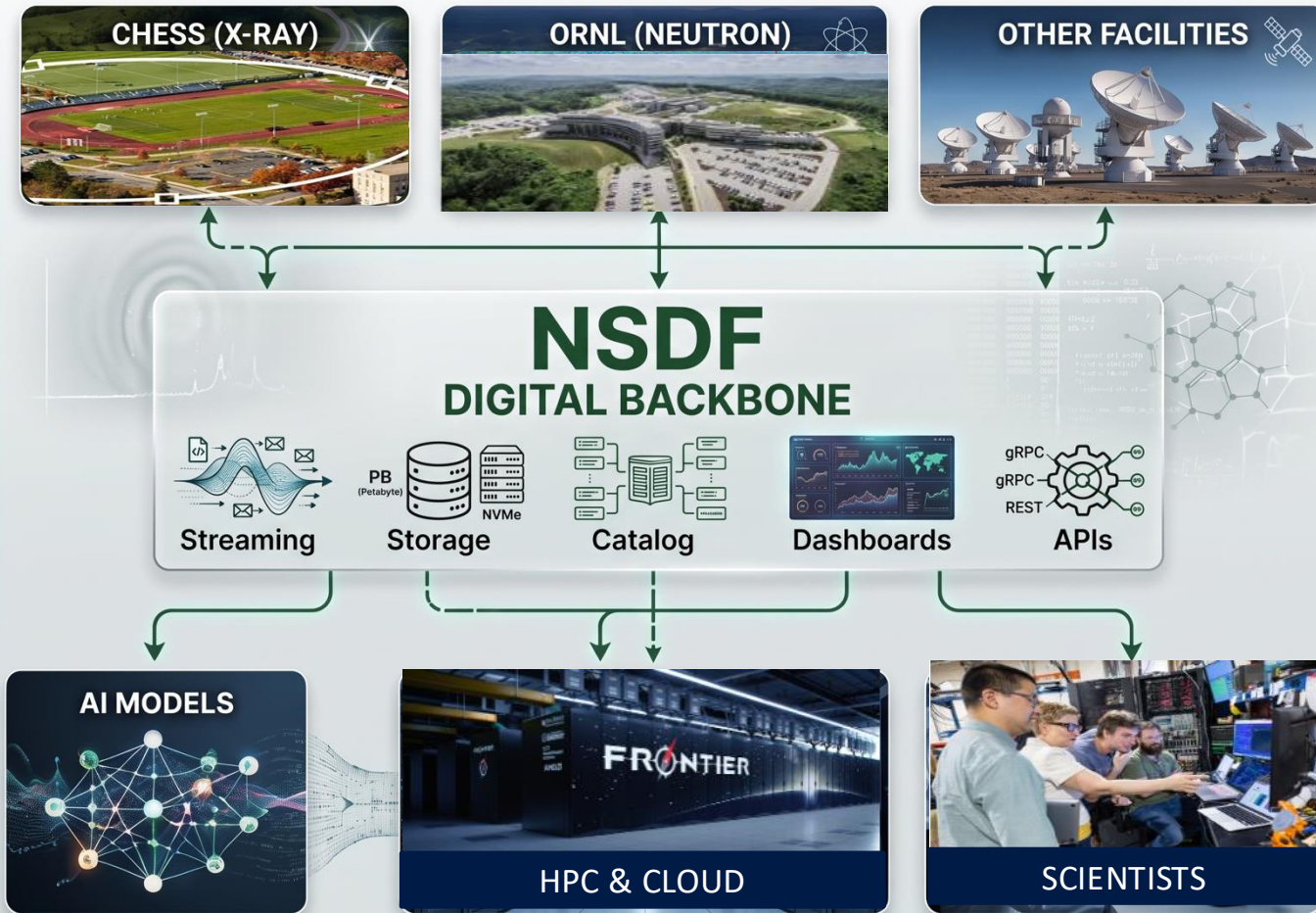
NSDF & Facilities: Scientific & Technical **Approach**

2609465



INTEGRATED SCIENTIFIC RESEARCH ECOSYSTEM

FRAMEWORK ECOSYSTEM OVERVIEW



FEDERATED DATA FABRIC:

Connect instruments, facilities, cloud, and HPC with FAIR data, metadata, and provenance.

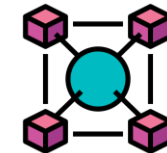


REAL-TIME INTEGRATION: Stream, harmonize, and deliver data in real time to AI and scientists.



AI-READY INFRASTRUCTURE:

Provide trusted services for AI workflows, models, and scientific reasoning.



SECURE & INTEROPERABLE:

Respect facility boundaries and security while enabling cross-facility collaboration.



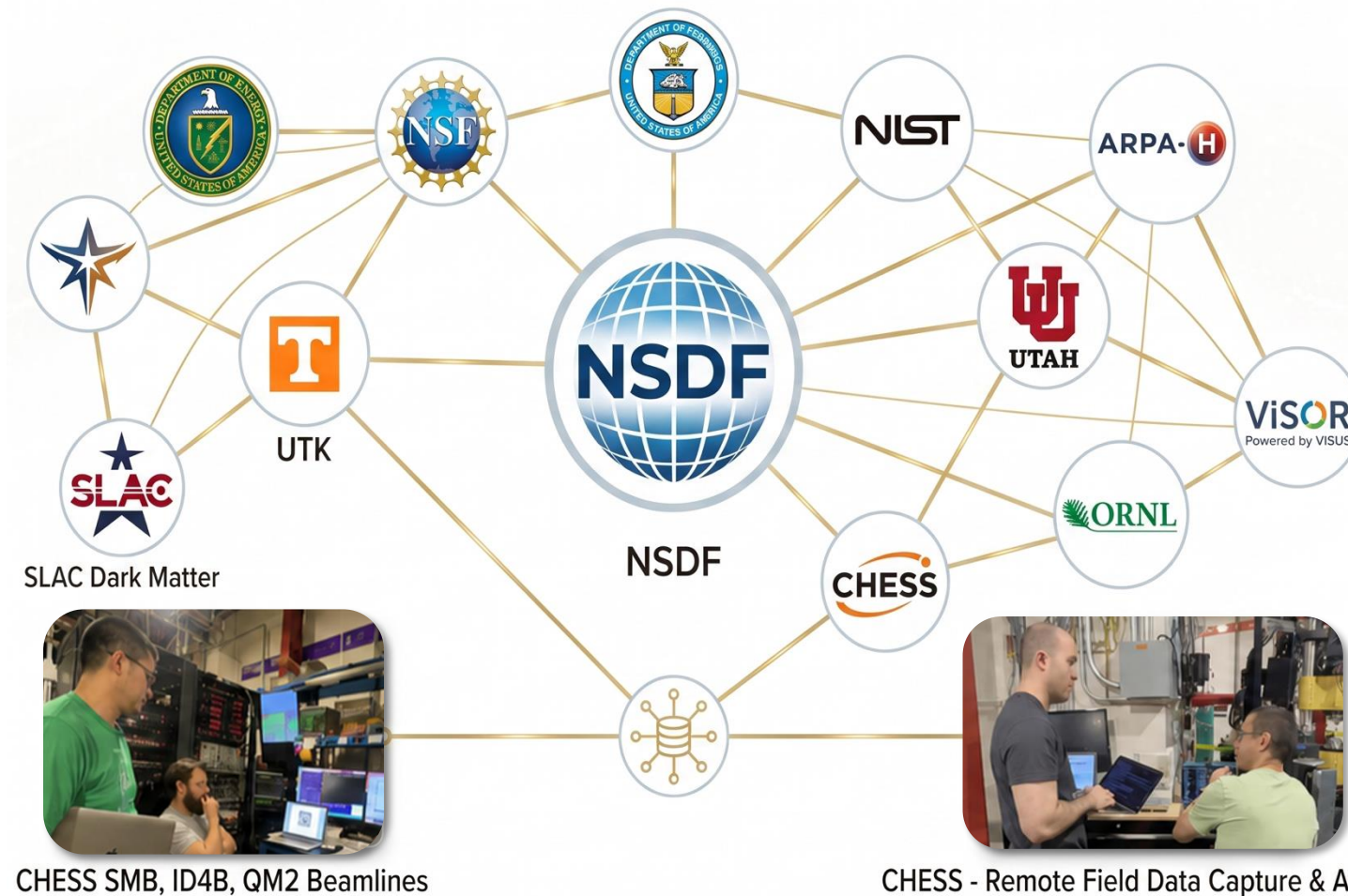
REUSABLE & EXTENSIBLE:

Modular services that can be deployed across facilities and scientific domains.

NSDF is the digital backbone that connects facilities, AI, and scientists.

NSDF & Facilities: Community & Scientific Collaboration

2609465



Strategic Institutional Partnership

Key collaborators include UTK, University of Utah, ORNL, CHESS, NIST, and ViSOAR, with support from NSF, DOE, and ARPA-H.

Live Facility Integration

The framework is operationally applied to the CHESS SMB, ID4B, and QM2 beamlines for real-time experimental monitoring.



Multi-Agency Ecosystem

Primary funding and oversight from NSF and DOE, with alignment across the Air Force, Army, and Navy, and planned expansion to industry partners.



Broad Scientific Scope

Beyond beamlines, the framework is applied to Dark Matter research at SLAC, Power Grid resilience, and ARPA-H cancer surgery initiatives.



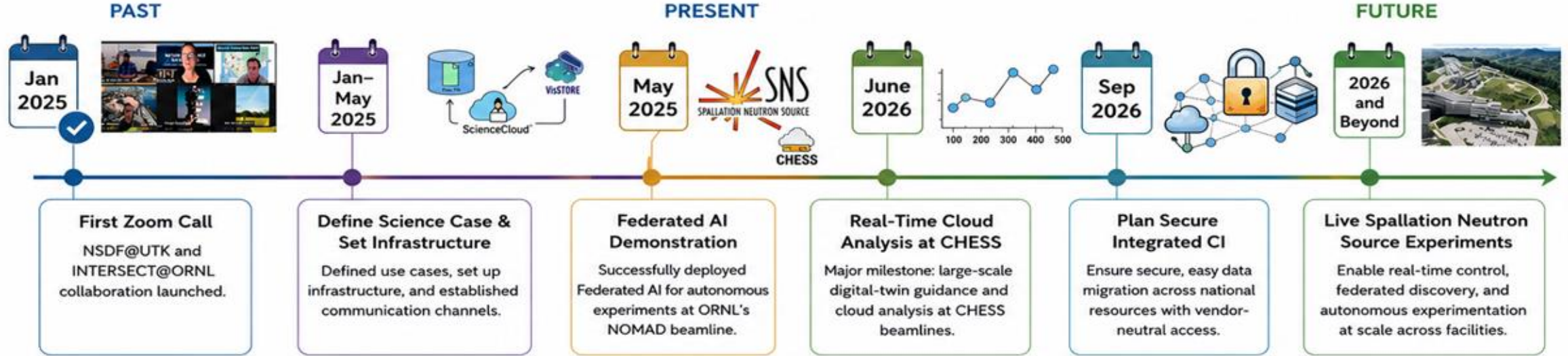
Community & Scientific Impact

Enabling shared infrastructure, open tools, and collaborative discovery across institutions and disciplines.



From Institutions to a Connected Scientific Ecosystem.

NSDF & Facilities: **Impact**, Deliverables & Scientific Outcomes



Deliverables & Impact



Operational Framework

A working model for federated AI and digital-twin experimentation across multi-institutional facilities.



Data & Infrastructure Integration

Secure, scalable, and interoperable infrastructure connecting national resources and cloud services.



Scientific Outcomes

Improved additive manufacturing quality control, residual strain mapping, and new materials discovery.



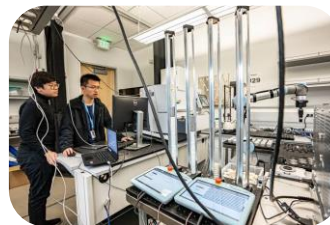
Community Impact

Open tools, shared workflows, and reproducible science accelerating discovery across disciplines

Vision of the Future

Autonomous Science

AI-driven systems propose, execute, and adapt experiments in real time.



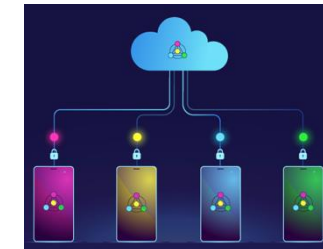
Open & FAIR Ecosystem

FAIR data, open tools, and shared standards enable trust and reuse.



Federated Discovery

Seamless collaboration across institutions, facilities, and disciplines.



Transformative Impact

Accelerated discovery, innovation, and solutions for national and global challenges.

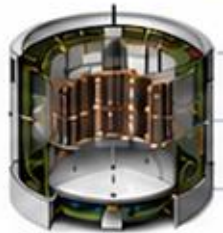


From Three Cross-Facilities Demonstrations to a Reusable National Capability.

Science in Action: NSDF-Enabled **Dark Matter** Data Commons

CDMS – Dark Matter Physics

Uncovering the Nature of Dark Matter



Detector Array
Cryostat
Shielding

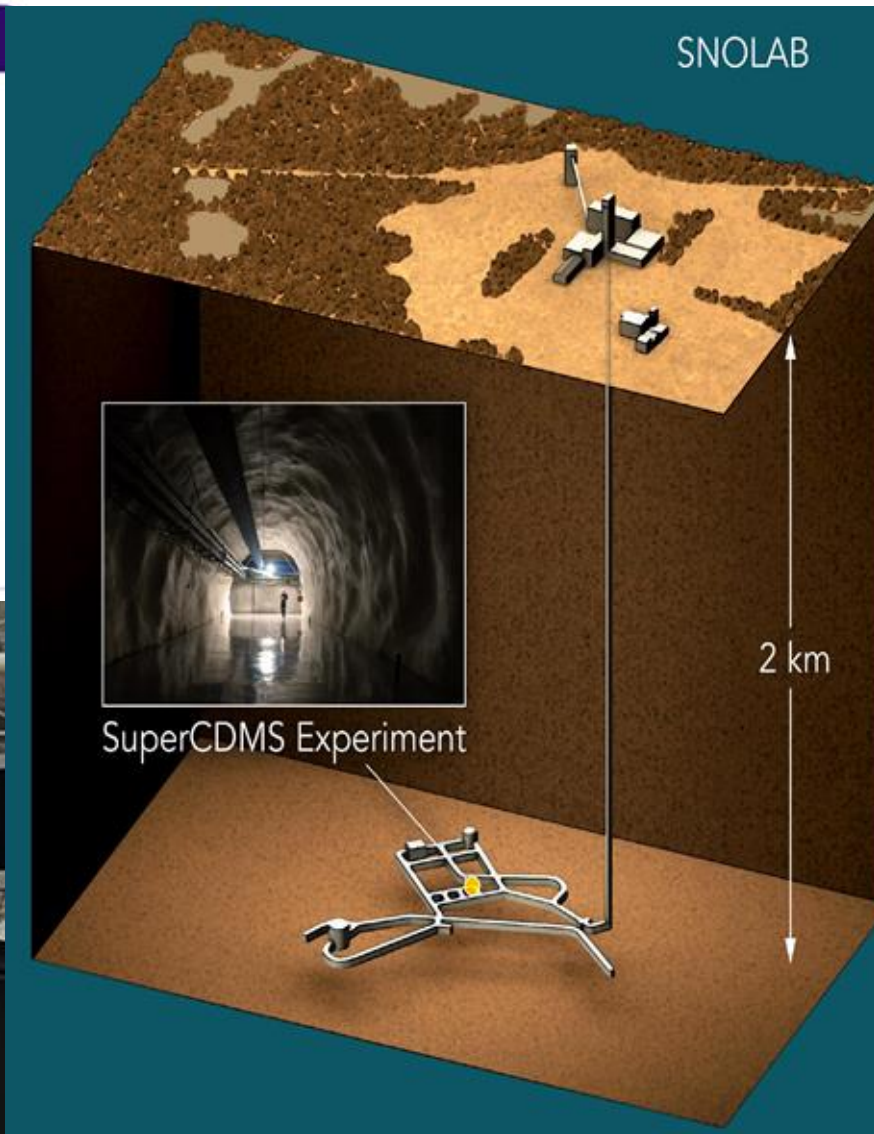
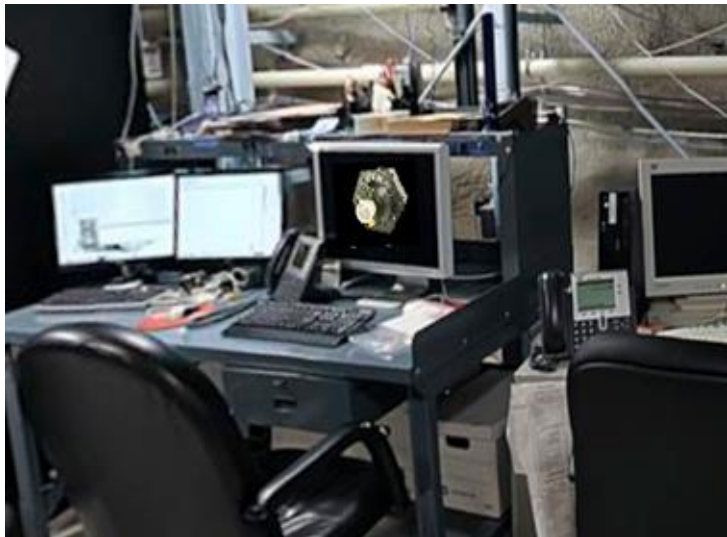
- Distributed data ingest from underground detectors
- Federated analysis pipelines
- Reproducible workflows
- Global collaboration at scale



Galaxy



Collision Event



PETABYTES of multiple-channel tie series (1D)



University of Colorado Denver, Denver

Amy Roberts, Kitty Mickelson,

University of Tennessee, Knoxville

Jack Marquez, Kin Hong NG, Michela Taufer

ViSOAR LLC, Salt Lake City

Amy Gooch

University of Utah, Salt Lake City

Aashish Panta, Giorgio Scorzelli, Valerio Pascucci,

University of Minnesota, Minneapolis

Prisca Cushman, Himangshu Neog,

Washington University, Saint Louis

Matthew Fritts,

NSDF & Dark Matter: Scientific Challenges & Objectives



CDMS Detectors
(Cryogenic Dark Matter Search)



**Proprietary
MIDAS Files &
Custom Software**



**Limited Access
(Collaboration Only)**



**Scientific Barriers
& Lost Opportunities**

LOCATION:

SNOLAB
2 km underground
in Canada
(at left)

LOCATION:

Local systems
within the
collaboration

LOCATION:

Restricted to
authorized
collaborators

LOCATION:

Community-wide
impact

Why It Matters

Dark matter searches increasingly depend on collaboration with other experiments and disciplines. Data wrangling is the biggest obstacle.



BREAK DATA SILOS

Data locked in proprietary formats limits reuse, collaboration, and innovation.

LOCATION:

Across experiments
and institutions



LOWER THE LEARNING CURVE

Reduce years of training and complex software dependencies to analyze and interpret data.

LOCATION:

New researchers
and collaborators



ENABLE OPEN & REPRODUCIBLE SCIENCE

Make calibration datasets FAIR, transparent, and accessible to the broader community.

LOCATION:

Global scientific
community

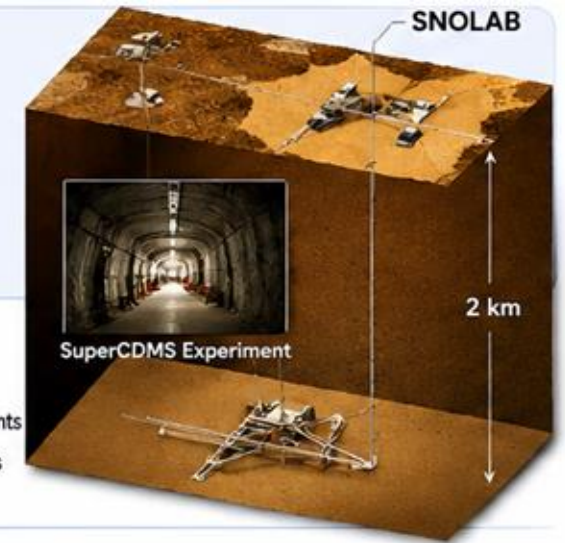


ACCELERATE AI-READY DISCOVERY

Prepare high-quality datasets for AI/ML methods and cross-disciplinary research.

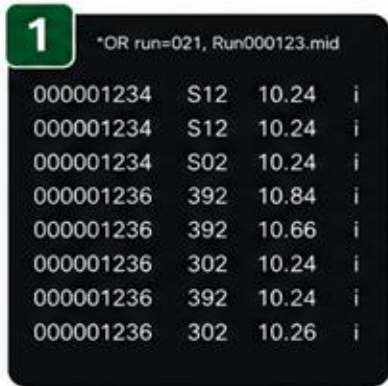
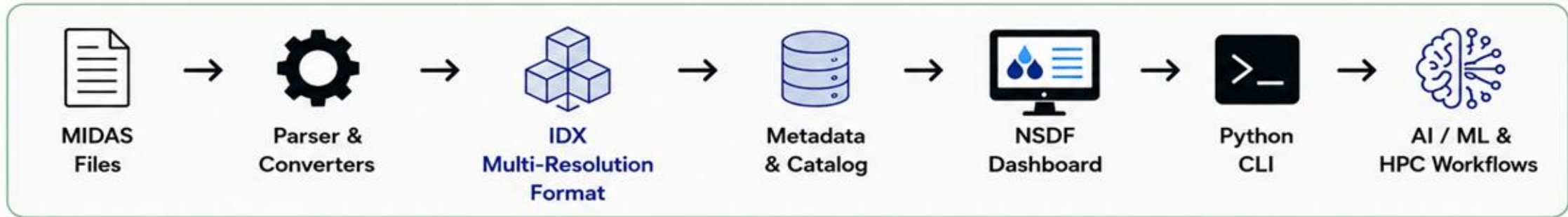
LOCATION:

AI/ML researchers
and data scientists

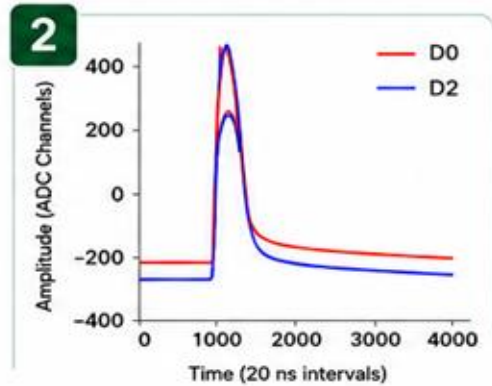


*Transforming Proprietary Dark Matter Data into
an Open, FAIR, and Reusable Resource for Scientific Discovery.*

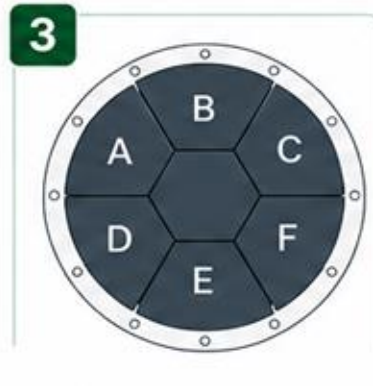
NSDF & Dark Matter: NSDF Scientific & Technical **Approach**



MIDAS raw data excerpt



Multi-channel time series (CDMS event)



Detector layout (Phonon sensors A-F)



NSDF Dashboard
Real-time visualization & exploration



Python CLI
Programmatic access & automation



Built on FAIR Principles:

Findable

Accessible

Interoperable

Reusable

NSDF Provides The Cyberinfrastructure, Services, And Tools That Transform Raw Experimental Data Into An Open, Ai-ready Ecosystem.

NSDF & Facilities: Community & Scientific Collaboration



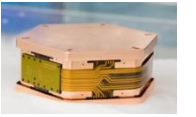
FAIROS Dark Matter Data Commons

Building a FAIR and open infrastructure to share experimental dark matter data and workflows.



CDMS Collaboration

Releasing the rich R76 calibration dataset for broad community use.



NEXUS-DM & Community Partners

Aligning with national efforts to unify tools, data, and standards across experiments.



Open-Source Software Ecosystem

NSDF tools, dashboards, and CLI released as open source for maximum impact.



Workforce Development

Training, tutorials, and reproducible workflows to grow the next generation of researchers.



Community Metadata Standards

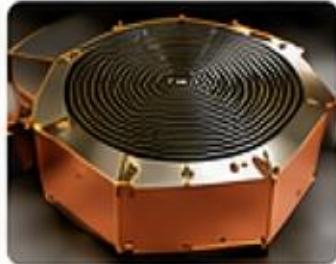
Establishing shared practices for rare-event physics data and documentation.



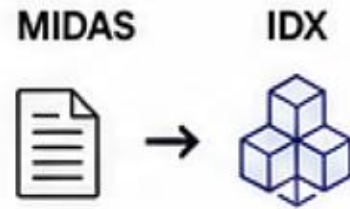
From Proprietary Data to Shared Discovery.

NSDF & Dark Matter: **Impact**, Deliverables & Scientific Outcomes

Key Deliverables



R76 Calibration Dataset
Rich dataset for broad community use.



MIDAS → IDX Conversion
Transform proprietary data to open, multi-resolution IDX format.



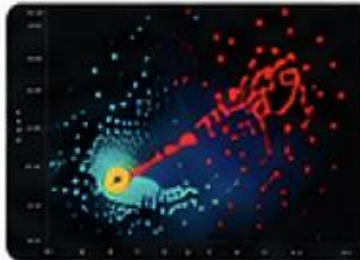
NSDF Dashboard
Interactive exploration of events, channels, and waveforms.



Python CLI
Programmatic access and scalable workflows.

Dataset	Size	Resolution	Access
1. R76 Calibration Dataset	100 GB	100 MHz	Open
2. R76 Calibration Dataset	100 GB	100 MHz	Open
3. R76 Calibration Dataset	100 GB	100 MHz	Open
4. R76 Calibration Dataset	100 GB	100 MHz	Open
5. R76 Calibration Dataset	100 GB	100 MHz	Open
6. R76 Calibration Dataset	100 GB	100 MHz	Open
7. R76 Calibration Dataset	100 GB	100 MHz	Open
8. R76 Calibration Dataset	100 GB	100 MHz	Open
9. R76 Calibration Dataset	100 GB	100 MHz	Open
10. R76 Calibration Dataset	100 GB	100 MHz	Open

FAIR Metadata & Catalog
Metadata-rich search and dataset discovery.



AI-Ready Workflows
Enable ML-based analysis and new discoveries.

Scientific Impact



Democratize Dark Matter Data

Open access to petabytes of calibration data.



Accelerate Reproducible Science

Standardized formats and tools ensure transparency and repeatability.



Enable AI-Driven Discovery

High-quality data empowers machine learning and advanced analysis.



Lower Barriers, Grow Community

Empower researchers, students, and new collaborators.



Strengthen U.S. Leadership

Advances open science and rare-event physics worldwide.

Transforming Dark Matter into the Blueprint for Future AI-Ready FAIR Scientific Data Commons.

Science in Action: NSDF Powers Earth Science at NASA Ames



BEFORE:

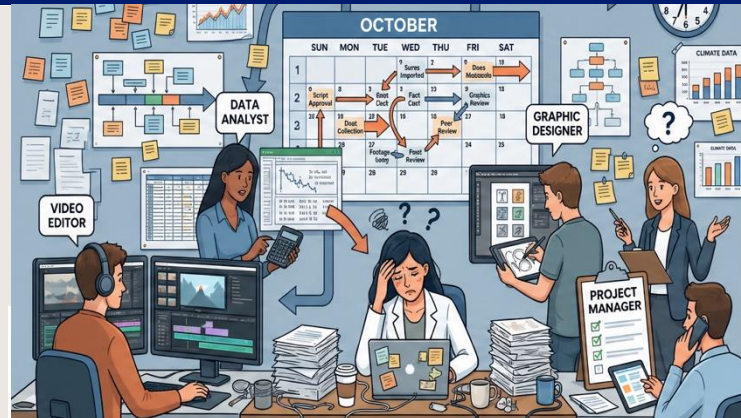
1-2 Weeks

Turnaround Time

Dedicated Support Team

Static Output

**Each Revision Requires
a Complete Rerun**



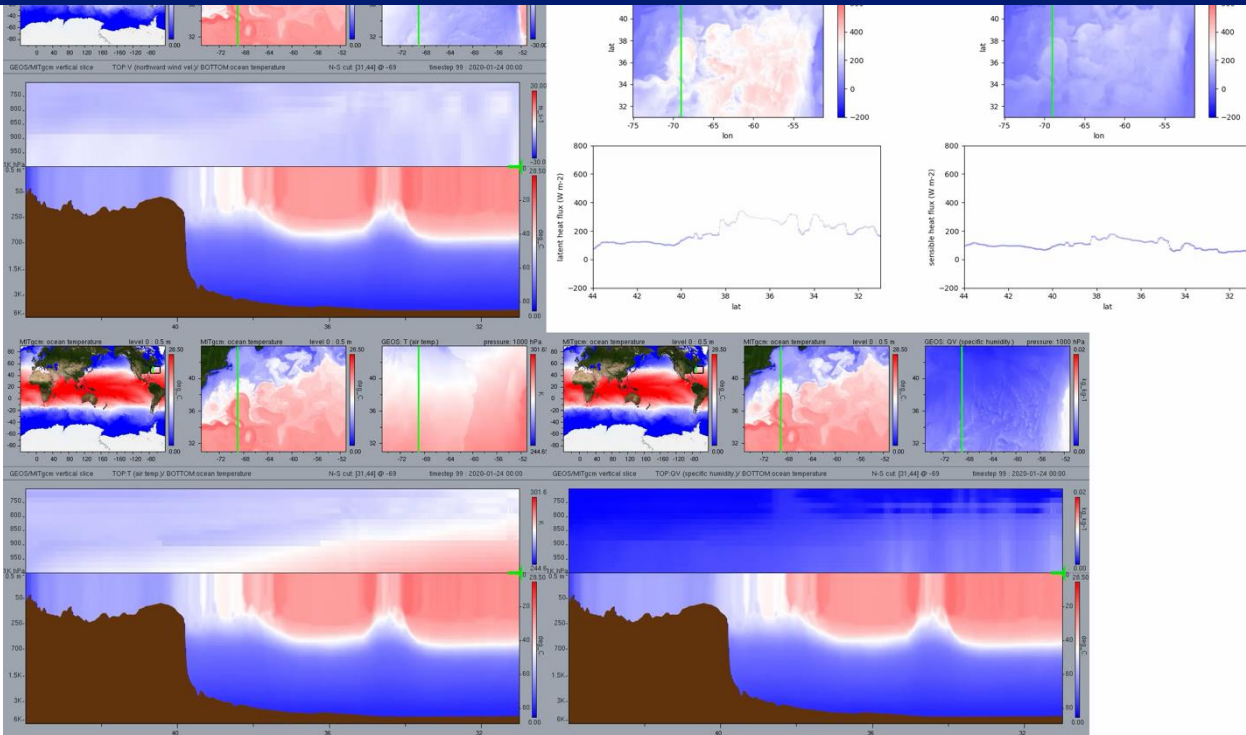
AFTER:

**Real-Time Exploration with
Interactive Dashboard**

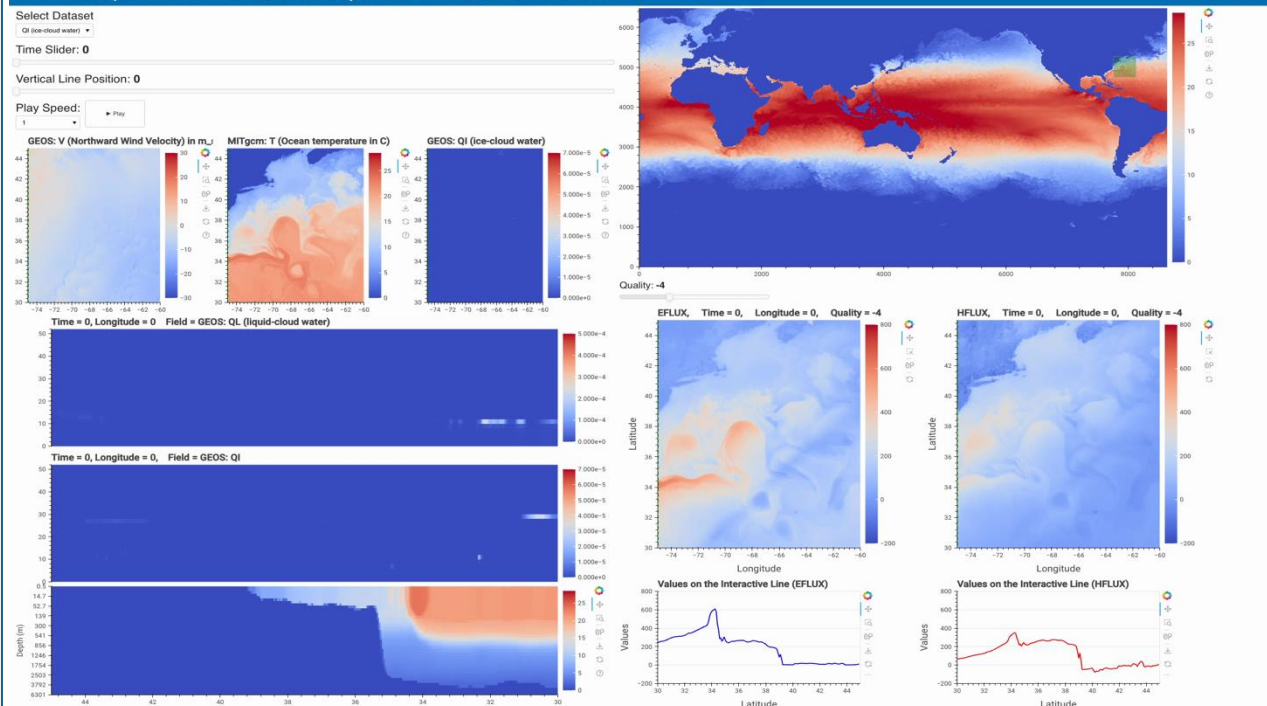
Global Parameter Flexibility

**Instant Visualization and
Analytics**

With NSDF, Scientists Moved From Waiting For Answers To Exploring Data In Real Time.



NASA Coupled Dashboard for Atmospheric and Oceanic Variables



From Research to Training: Build Foundational Skills

2609465



Tutorials & Workshops



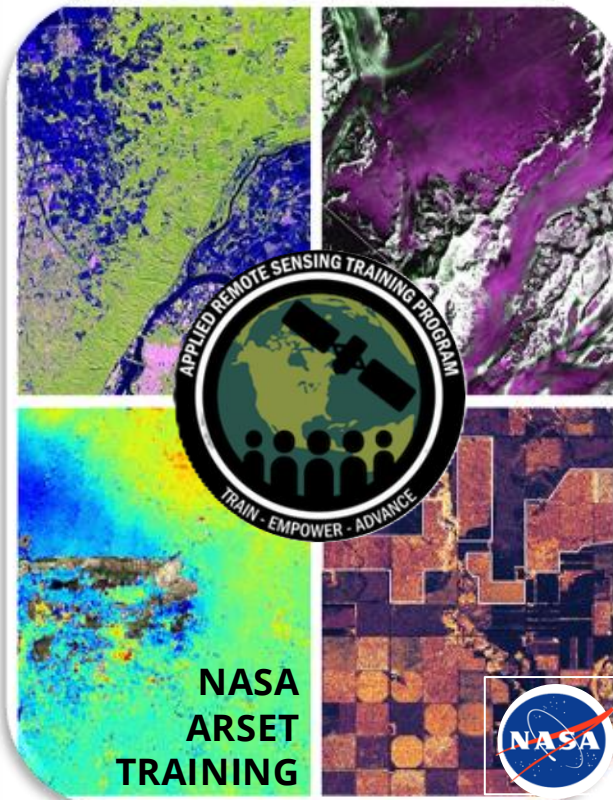
Hands-on Labs



Open Workflows



Docs & Training Materials



Cyberinfrastructure Training & Education for Synchrotron X-ray Science



50+ Tutorials



40+ Workshops & Training Events



2,000+ Participants Trained



100+ Institutions Reached



Open Educational Resources

From Research to Training: **Inspire** Future Leaders

2609465



NSDF FELLOWS



Fellows from universities across the U.S.

TRAVEL AWARD WINNERS



Supporting students and early-career researchers

STUDENT PRESENTERS



Opportunities to share research and build professional networks.

MENTORING & STUDENT PROJECTS



One-on-one mentoring



Collaborative research



Real-world experience



Building next-gen leaders

15

NSDF Fellows
(SC24 Cohort)

40+

Travel Award
Winners

100+

Universities
Represented

Hundreds

of Students
Impacted

From Research to Training: Engage the Community



All Hands
Meetings



Community
Panels &
Workshops



Partnerships &
Cross-Disciplinary
Coates
(SciVis Contest & More)

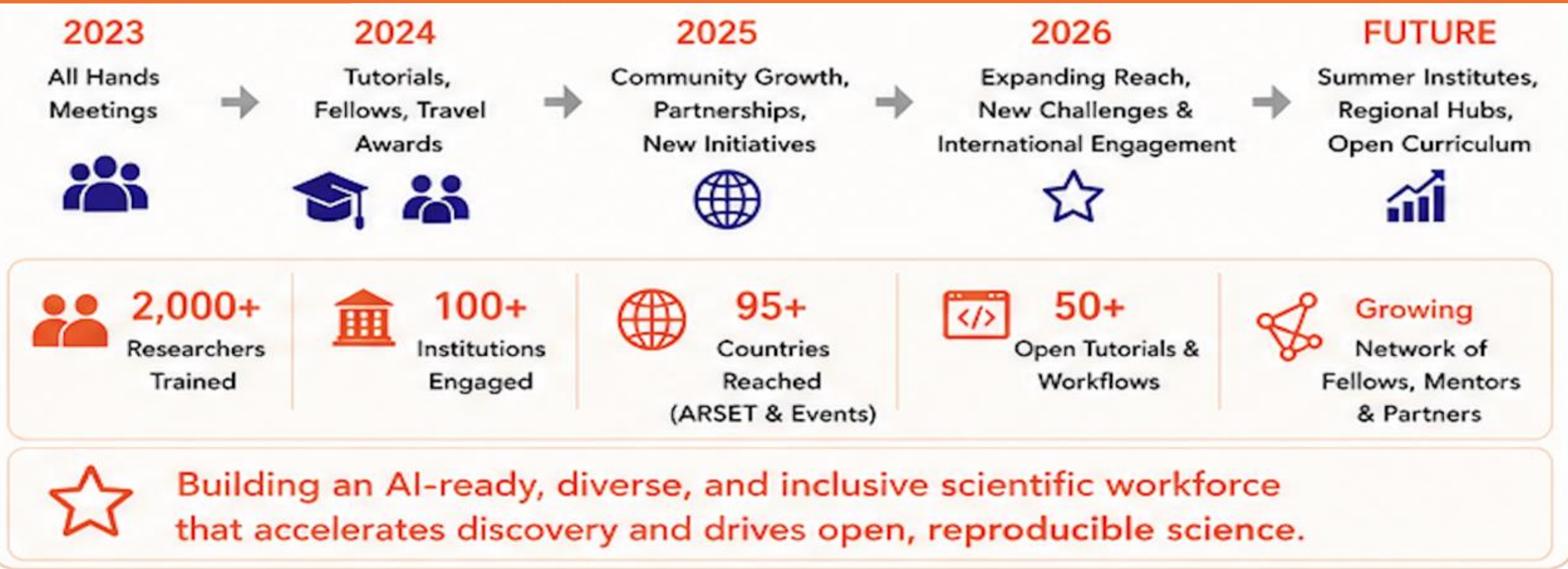


Outreach &
Broadening
Participation



*NSDF Builds Tomorrow's Scientific Leaders Through
Open Education, Hands-On Training, and
Community Engagement.*

From Research to Training: **Grow** the National Workforce



People

Education

Community

Collaboration

Impact

Supported by the National Science Foundation
2609465

METRICS OF SUCCESS

YEAR 1
2026 – 2027

YEAR 2
2027 – 2028

YEAR 3
2028 – 2029



Operational
Facilities



3+
facilities
connected



6+
facilities
connected



10+
facilities
connected



Discoverable
FAIR Data



5+
petabytes
cataloged



20+
petabytes
cataloged



50+
petabytes
cataloged



Service
Availability



95%
service
availability



99%
service
availability



99.9%
service
availability



Cross-Site
Interoperability



Baseline
interoperability



90%+
cross-site
interoperability



95%+
cross-site
interoperability



AI & Autonomous
Operations



Prototype
AI-assisted support
and workflows



Integrated
AI services across
use cases



Autonomous
AI-driven experiments
at scale



Community
Adoption



8
new institutions
engaged



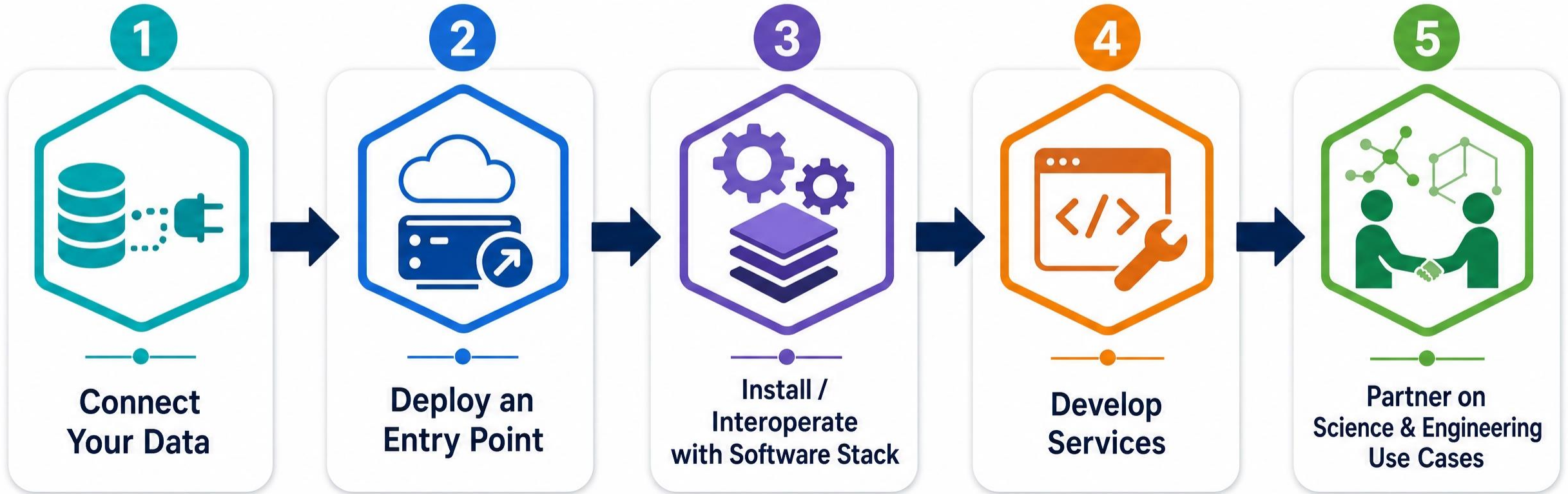
10
new institutions
engaged



12+
new institutions
engaged

How Can You Collaborate with NSDF?

<https://nationalsciencedatafabric.org/contact.html>



Contact us: info@nationalsciencedatafabric.org

