Update on Opportunities in the Office of Advanced Cyberinfrastructure and the NAIRR

Oklahoma Supercomputing Symposium Sept. 25, 2024





National Science Foundation's Mission

"To promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense..."



Programmatic directorates and offices supporting the NSF Mission

Engineering

Biological Sciences

> Mathematical & Physical Sciences

Integrative Activities International Science & Engineering

> Social, Behavioral & Economic Sciences

Technology, Innovation and Partnerships (including Polar Programs) Education & Human Resources

Geosciences

Compu Inform Science

Computer & Information Science & Engineering Division of Computer & Network Systems

Division of Computing & Communications Foundations

Division of Information & Intelligent Systems

Office of Advanced Cyberinfrastructure (OAC)

OAC – Transform Science and Engineering Through an Integrated Cyberinfrastructure Ecosystem



Defining, advancing and interconnecting broad computing and data ecosystem



Growing and developing communities and workforce



Enabling discovery through integrations of data, software and infrastructure

NAIRR

Infrastructure for AI

Investing in and transitioning to new technologies



Developing partnerships for long-term US leadership in research Cl

OAC investment areas

Advanced Computing Systems and Services	Production and operational level advanced computing and data capabilities and services
Networking & Cybersecurity	Advanced networking and security infrastructure, research and communities of practice capabilities
Learning & Workforce Development	Foster a national research workforce for creating, utilizing, and supporting advanced CI
Software & Data	Supports development and deployment of robust, reliable and sustainable data and software
Strategic Investments	Special opportunities, cross-cutting and national initiatives, open science and public access



A Few FY24 OAC Highlights and FY25 OAC Opportunities





21-065

1M



Cyber-Campus infrastructure Cyberfor Sustained infrastructure Scientific Innovation (CSSI)

Elements

Instrumentation

Major Research

CISE **CSSI** Community Frameworks Research Infrastructure

Advanced Computing Systems & Services

Mid-scale Research

Research



Infrastructure 2 Infrastructure 1

Campus Cyberinfrastructure (CC*) NSF 24-530

Area 1: Data Driven Networking Infrastructure

> Campus Region

Campus network upgrades Area 2: Computing and the Computing Continuum

> Campus Region

Campus compute resources; 20% is shared, typically through PATh Area 3: Network Integration and Applied Innovation

> Small Large

Networking R&D applied to the campus network with graduate student involvement Area 4: Data Storage and Digital Archives

> Campus Region

Campus data storage platforms; 20% is shared, typically through OSDF Area 5: Strategy

> Campus Region

A grant to help teams plan for a full proposal! No CI plan; Funds community building activities; No hardware

Region proposals must focus on the science requirements of small and under resourced institutions; next Deadline is October 15, 2024 Amy Apon, <u>awapon@nsf.gov</u> and Kevin Thompson, <u>kthompso@nsf.gov</u>

CC* campus and regional investment examples



Support for Campus Connections to Regional Research and Education Network



Lawrence 2.0 Advancing Multidisciplinary Research and Education in South Dakota, OAC 2346643



- Support for computationally intensive research and education;
- Outreach to state and regional collaborators and SD tribal institutions



Advanced Computing Systems & Services: Adapting to the Rapid Evolution of Science and Engineering Research 2.0

View guidelines



The intent of this solicitation is to request proposals from organizations who are willing to serve as resource providers within the NSF *Advanced Computing Systems and Services* (ACSS) program.

Deadlines: October 29, 2024, Category I Submissions June 24, 2025, Category II Submissions



Leadership Class Computing Facility Awarded

LCCF provides the computational and data analytics ecosystem that provisions large-scale capabilities for S&E research to enable discoveries that would not be possible otherwise.



The LCCF, led by the Texas Advanced Computing Center (TACC)

- **5** Distributed Sites
- 27 academic partners, including
- **10** MSI partners

Current Status

- July 2024 award with expected operational deployment in FY 2027
- Phase 1 Vista announced to set the stage for Horizon, flagship of LCCF

11

CI resources and services that the research community can use now

Democratized access to advanced computing









Commercial cloud

Science Gateways expertise

CI services for NSF major and mid-scale RI





Facility data lifecycle



Security Operations



Regulated Research

Community and workforce development



Minority Serving Cl Consortium (MS-CC)



CI Workforce Development

Portals: • ACCESS: <u>https://access-ci.org/</u>

- LCCF: <u>https://lccf.tacc.utexas.edu/</u>
- PaTh: <u>https://path-cc.io/</u>
- SGX3: <u>https://sciencegateways.org/</u>
- MSCC: <u>https://www.ms-cc.org/</u>
- RCD Nexus: https://rcd-nexus.org/
- Trusted CI: <u>https://www.trustedci.org/</u>
- Research SOC: https://omnisoc.iu.edu/services/researchsoc/
- CI Compass: <u>https://ci-compass.org/</u>
- NAIRR Pilot: <u>https://nairrpilot.org/</u>



NSF Software and Data Infrastructure programs



Cyberinfrastructure for Sustained Scientific Innovation (CSSI)

- Supports the <u>development and deployment</u> of robust, reliable and sustainable <u>data and software</u>
 <u>cyberinfrastructure</u>
- Brings <u>innovative</u> capabilities towards sustained scientific innovation and discovery
- Provides a <u>cross-directorate</u> opportunity to advance common approaches to sustain and innovate research cyberinfrastructures

Webinar 10/17/2024 Register here:

https://new.nsf.gov/events /cssi-nsf-22-632-programwebinar/2024-10-17

Submission Deadline: 12/1/2024



https://www.nsf.gov/pubs/2022/nsf22632/nsf22632.htm



Programs to Support People

SCIPE: Strengthening the Cyberinfrastructure Professionals Ecosystem (new solicitation pending)

SCIPE: WildWEST: Wild-area-networks Wireless Enabling Science Team

Addresses best practices & gaps in wireless+WAN engineering, field systems design, and science systems integration for field science applications

NSF Award # 2416870, (OAC). **PI: Scotty Strachan**, Nevada System of Higher Education **Jim Stewart**, Utah Education & Telehealth Network, **Joe Breen**, University of Utah, **Derek Masseth**, Arizona Sun Corridor Network

Building a Computational and Data-Intensive Research Workforce & Network in the Mid-Atlantic Region

Creates a talent pipeline of Research Software Engineers to accelerate and enable domain sciences, especially targeting social and behavioral and economic as well as coastal sciences.

NSF Award # 2417814, Sunita Chandrasekaran, Rudolf Eigenmann, John Huffman, Benjamin Bagozzi, Tian-Jian (Tom) Hsu, U of Delaware



CyberTraining: Training-based Workforce Development for the Cyberinfrastructure Workforce

Vision: Prepare, nurture, grow scientific research workforce

- Ensure broad adoption of CI tools, research with training, develop instructional materials, undergrad/grad curricula
- Three levels: Pilot (\$200K), Small (\$500K), and Medium (\$1M)
- NSF 23-520, deadline January 16, 2025
- Webinar, 10/7/24 3pm Eastern, Register here: <u>https://new.nsf.gov/events/cybertraining-nsf-23-520-</u> program-webinar/2024-10-07

arctic science * geoscience * climate water environment * construction science * bioengineerin g * dark matter *discovery* * earth science * molecular science * cybersecurity * physics ...



National Discovery Cloud for Climate (NDC-C)

The National Discovery Cloud for Climate (NDC-C) initiative enables access to advanced computing, data, software and networking resources in support of climate research and its applications to other research fields. NDC-C supports OAC training and research programs through co-funding.



https://new.nsf.gov/cise/national-discovery-cloud-climate



National Al Research Resource (NAIRR) Pilot Update



Vision for the National AI Research Resource

A widely-accessible, national research infrastructure that will advance the U.S. AI R&D environment, discovery, and innovation by empowering a diverse set of users through access to:



Secure, high-performance, privacy-preserving **computing**



High-quality datasets

Catalogs of **testbeds** and **educational materials**

-

Training tools and user support mechanisms



NAIRR <u>Pilot</u> Goals

- 1. Demonstrate the value & impact of the NAIRR concept.
- 2. Support novel & transformative AI research and education with participation from broad communities.
- 3. Gain initial experience to advance and refine the NAIRR design in preparation for a full NAIRR implementation.



The pilot will demonstrate or investigate all major elements envisioned in the NAIRR, but at a **limited** scale: A best effort, **proof of concept** approach leveraging agencysupported resources and in-kind contributions from industry and non-profits.



Two NAIRR Pilot Opportunities Available



* Apply to use NAIRR Pilot resources

Access to computing, data, model, collaboration resources – nairrpilot.org



* Train, educate and/or build community around NAIRR pilot resources

NAIRR Pilot Expansion NSF 24-109

https://www.nsf.gov/pubs/2024/nsf24109/nsf24109.jsp

EducateAl initiative NSF 24-025



Apply for NAIRR Pilot Resources

NAIRR Pilot

National Artificial Intelligence Research Resource Pilot

Go to https://nairrpilot.org/ and scroll down

CLASSROOM AND EDUCATOR RESOURCES

Request access to educational platforms (such as Juypter notebooks).

RESEARCHER RESOURCES

Request access to advanced computing, cloud computing, models, software, platforms, and collaborations.

DATA, MODELS, AND MORE

Additional government and government-funded resources that do not require an application.

Apply

Apply

View



Agency and Non-govt Resources Available to Community



Application Process

- Project Description (3-pages)
 - A. Scientific/Technical Goal
 - B. Estimate of Compute, Storage, and other Resources (Justify this!)
 - C. Support Needs
 - D. Team and Team preparedness
- Supporting Documents
 - References
 - CV / Bio sketch material for team lead

FAQ at <u>https://nairrpilot.org/help/faq</u>

Help sessions for submitting to nairrpilot.org are coming!

- Not an NSF Proposal!
 - Not a request for funding
- Resource availability is first come first serve
 - No guarantee of future availability during pilot
- Review process:
 - Expectation: Requests reviewed within ~6 weeks of submission
 - Requests are batched to 15th day of each month (i.e. requests submitted by September 15th are currently being reviewed).

NAIRR Pilot Resource Request Review Process

How the pilot handles ~40-50 submissions that come in per month



NAIRR Pilot Resource Geographic Impact so far As of 9/3/24



Researchers supported by numerous agencies

- NSF DOE
- NIH DARPA
- NIJ DOD
- NOAA ARO
- USDA ONR
 - VA

Allocated projects

Demo/Education/Training projects

O Resources

Train, educate and/or build community: NSF 24-109 NAIRR Pilot Expansion

- Promote and diversify AI education, workforce development, and broaden access to AI resources
- Two focus areas:
 - *Researchers* new or beginning to develop, use or integrate AI into research investigations
 - Supplements, EAGER, Conference, workshop
 proposals
 - Educators
 - Conference and workshop proposals
 - Research (or education) Coordination Networks

Communities of interest:

- Rural institutions and institutions from EPSCOR states
- Minority serving institutions (MSI)
- Non-R1 institutions including primarily undergraduate institutions & community colleges
- Emerging research institutions
- Intersection of above (E.g. Non-R1, MSI)
- Regional communities





Train, educate and/or build community: NAIRR Pilot support through EducateAI initiative NSF 24-025

CUE-P: Integrating Artificial Intelligence Literacy into Community College Programs PI: Joyce Malyn-Smith, Education Development Center

CUE-P: Artificial Intelligence Entry Pathways PI: Antonio Delgado, Miami Dade College

CUE-T: Mobilizing CAHSI Institutions to Infuse Ethical Reasoning in CS Curricula through Contextualized AI Scenarios PI: Ann Q. Gates, University of Texas, El Paso

Collaborative Research: CUE-T: Expanding the undergraduate AI talent pipeline by democratizing access to instructional capabilities across institutions of higher education (IHEs) PI: Judith A. Spitz, UIUC, Hofstra, Cornell.

Community Colleges

Train, educate and/or build community: First Four NAIRR Classroom Awards



Course: Machine Learning in Cybersecurity Institution: U of Southern Mississippi Educator: Nick Rahimi Students: 30



Course: Deep Learning Institution: North Carolina State Educator: Edgar Lobaton Students: 800



Course: Machine Learning for CS Undergrads Institution: Southern Oregon University Educator: Bernadette Boscoe Students: 25



Course: Al-enabled data visualization interfaces Institution: University of California, Berkeley Educator: Carl Boettiger Students: 122

Goal: understand classroom requirements for the full NAIRR

What role does the NAIRR have in supporting data for AI research?

- Set guidelines and criteria for dataset inclusion in the NAIRR
- Support a data search and discovery service
- Encourage and incentivize communities to contribute 'analysis-ready' datasets
- Facilitate an Al data commons

NAIRR Pilot

- Provide access to integrated computing and data platforms
- Provide access to restricted datasets
- Provide technical expertise and user support for data communities and community driven curation efforts



- NAIRR will not define dataset standards, which continue to evolve and best defined by communities
- NAIRR will not fund the collection or creation of specific community datasets

A number of agency datasets are available through the pilot



National COVID Cohort Collaborative





Tropical Cyclone Dataset



USPTO Research Dataset



Lake Michigan Substrate Prediction Dataset



US Census of Agriculture

Note: NAIRR Steering subcommittee is working on a process for non-govt datasets to be contributed to the NAIRR Pilot

Plan to evolve the NAIRR Pilot in next 18 months

- Support additional workshops and community activities: portal, data, software, education, training and outreach
- Write an interim report with year 1 experiences
- Complete integration of partner resources
- Sponsor a NAIRR Pilot Annual Meeting in early CY 2025
- Support budding NAIRR Pilot user community and continue outreach
- Mature NAIRR Pilot operations, develop metrics and assess gaps
- Integrate additional resources or demonstrations projects

We have discussed opportunities in the Office of Advanced Cyberinfrastructure and the NAIRR





Extra slides



Many NAIRR Pilot Activities Underway



NAIRR Pilot

Lessons learned so far

- NAIRR Task Force Report is a strong guiding document
- Industry has remained engaged and is actively participating and contributing
- Similarly, agency engagement is high and new agencies continue to reach out (DHS, FDA, DoT, USAID)
- Onboarding users new to AI or to a new platform is resource intensive and requires outreach and training
- Demand from research community is high
 - Agency resources are at capacity
 - Industry contributed resources are filling up



