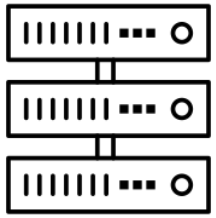
Broadening Access and Accelerating Responsible AI Innovation Through the National AI Research Resource (NAIRR) Pilot

Katie Antypas
Director, Office of Advanced Cyberinfrastructure

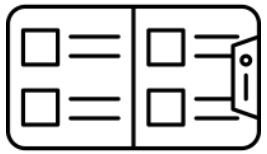
Vision for the National AI Research Resource

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

Why do we need a NAIRR?

- Many potential contributors lack access to requisite resources which can be costly ***as well as*** hard to navigate
- Researchers investigating AI to serve the public good require access to resources
- To train the next generation of researchers and AI leaders

Urgent national goals we are aiming to fulfill with NAIRR

Spur **innovation**

Increase the **diversity** of talent in AI

Improve U.S. **capacity** for AI R&D

Advance **trustworthy AI**

- Facilitate *national, coordinated* access to AI resources for the broad research and education community to accelerate discovery and innovation.
- Assure that public interest is strongly represented in AI and drives a responsible and trustworthy AI ecosystem.
- Combine forces to increase AI resource capacity and expertise.

Full NAIRR Vision vs NAIRR Pilot Goals

NAIRR Task Force Report Budget Estimate

* Note no funding has yet been appropriated

Table 1. NAIRR Six-Year Budget Summary

Year	Resource Providers	Operating Entity	Evaluation	Total
1	$375M	$70M	$5M	$450M
2	$375M	$60M	$5M	$440M
3	$375M	$60M	$5M	$440M
4	$375M	$60M	$5M	$440M
5	$375M	$60M	$5M	$440M
6	$375M	$60M	$5M	$440M
6-year total	$2.25B	$370M	$30M	~ $2.6B

**** NOTE: only an estimate from the Task Force. No funding has yet been appropriated for NAIRR.***


THE WHITE HOUSE

OCTOBER 30, 2023

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

BRIEFING ROOM • PRESIDENTIAL ACTIONS

By the authority vested in me as President by the Constitution and the laws of the United States of America, it is hereby ordered as follows:

Section 1. Purpose. Artificial intelligence (AI) holds extraordinary potential for both promise and peril. Responsible AI use has the potential to help solve urgent challenges while making our world more prosperous, productive, innovative, and secure. At the same time, irresponsible use could exacerbate societal harms such as fraud, discrimination, bias, and disinformation; displace and disempower workers; stifle competition; and pose risks to national security. Harnessing AI for good and realizing its myriad benefits requires mitigating its substantial risks. This endeavor demands a society-wide effort that includes government, the private sector, academia, and civil society.

Full NAIRR Vision vs NAIRR Pilot Goals

NAIRR Task Force Report Budget Estimate

* Note no funding has yet been appropriated

Table 1. NAIRR Six-Year Budget Summary

Year	Resource Providers	Operating Entity	Evaluation	Total
1	$375M	$70M	$5M	$450M
2	$375M	$60M	$5M	$440M
3	$375M	$60M	$5M	$440M
4	$375M	$60M	$5M	$440M
5	$375M	$60M	$5M	$440M
6	$375M	$60M	$5M	$440M
6-year total	$2.25B	$370M	$30M	~ $2.6B

**** NOTE: only an estimate from the Task Force. No funding has yet been appropriated for NAIRR.***

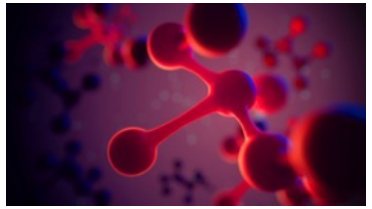
NAIRR Pilot Goals

Using existing and in-kind resources:

1. Demonstrate the value & impact of the NAIRR concept.
2. Support novel & transformative AI while reaching broad communities.
3. Gain experience to advance and refine the NAIRR design

NAIRR Pilot Users

AI Researchers



Domain Scientists Applying AI

Students and Educators

US-Based Institutions including:

- Academic institutions
- Non-profits
- Federal agencies or federally-funded R&D centers
- State, local, or tribal agencies
- Startups and small businesses with Federal grants

NAIRR Pilot Research Thrusts

- Accelerate societally-relevant research on **AI safety, reliability, security, and privacy**.
- Empower advances in **cancer treatment and individual health outcomes.**
- Support resilience and optimization of **agricultural, water, and grid infrastructure.**
- Improve design, control, and quality of **advanced manufacturing systems.**
- Address **earth, environmental, and climate challenges** via integration of diverse data and models.

Photo credits: NSF, iStock

NAIRR Pilot Organization

User Journey

US-based Researchers, Educators & Students

NAIRR Pilot Portal

https://nairrpilot.org

Pilot Resources and Opportunities

The NAIRR Pilot provides infrastructure and resources; it does not fund end-user research.

Operations

NAIRR Open

Enable open AI research and access to diverse AI resources via a central portal and coordinated allocations

NAIRR Secure

Enable AI research needing privacy and security-preserving resources. Assemble exemplar privacy preserving resources.

NAIRR Software

Facilitate use of AI software, platforms, tools and services across platforms

NAIRR Classroom

Reach new communities through education, training, user support and outreach

Governance

Community Design Process

Bringing together the strengths of government, private industry and non-profit partners

Contributing Partners

Agencies

- National Science Foundation
- Defense Advanced Research Projects Agency
- Department of Agriculture
- Department of Defense
- Department of Energy
- Department of Veterans Affairs
- National Aeronautics and Space Administration
- National Institutes of Health
- National Institute of Standards and Technology
- National Oceanic and Atmospheric Administration
- US Patent and Trademark Office (USPTO)

Non-governmental orgs

- AI2: Allen Institute for AI
- AMD
- Amazon Web Services
- Anthropic
- Cerebras
- Databricks
- Datavant
- EleutherAI
- Google
- Groq
- Hewlett Packard Enterprise
- Hugging Face
- IBM
- Intel
- Meta
- Microsoft
- MLCommons
- NVIDIA
- Omidyar Networks
- OpenAI
- OpenMined
- Palantir
- Regenstrief Institute
- SambaNova Systems
- Vocareum
- Weights & Biases

Contributions to the pilot go far beyond compute

Contributed Resources

- Access to computing hardware, systems and testbeds
- Cloud computing credits and access to associated models, data and software platforms
- Software and platforms
- Open models, datasets and PETs
- API access to closed models
- Educational platforms online notebooks for students
- Enhanced training, expertise and user support.

Forging a new kind of partnership between govt and non-govt

Computing (+ associated software, models platforms, training and user support)

Government contributions

DOE 10% Summit system
NSF % of multiple systems

Private and non-profit Contributions*

Amazon
Cerebras
Microsoft
NVIDIA
SambaNova Systems

≈

30-40M A100 GPU hours – split about evenly between gov and non-gov contributors

* Discussions are on-going with other partners

Primary contributions of agency and non-govt partners

	Computing (+associated software, models, user support)	Datasets	Integrated Platforms	Models	Software, Tools, Benchmarks	Training and User Support, education	Secure platforms, safety, PETs
Agency contributions	Amazon Cerebras Microsoft NVIDIA SambaNova Systems	NIH NASA NOAA USPTO USDA VA	NIH	NASA	DARPA DOD NIH	DOE NASA NSF VA	DARPA DOE NSF NIST VA
Industry & Non-profit contributions	DOE NSF	AI2 Google IBM Regenstreif Institute	Google Hugging Face Palantir Vocareum	AI2 Anthropic Groq IBM Meta OpenAI	Databricks EleutherAI HPE MLCommons Weights and Biases	AMD Intel HPE NVIDIA Vocareum	Datavant OpenMined

Omidyar Networks will sponsor workshops and expertise from portfolio partners

** Contributions are not exclusive to categories listed

Pilot Launched in Jan with 10 agency and 25 non-govt partners

Extended to March 31st

Initial call closed March 1st
> 150 submissions
Next open call mid-April

Data Challenges and Opportunities

Data growing in size and complexity

Data pipelines, staging and wrangling often dominate researcher time

Data quality and fairness is of top concern

Data often needs to be transferred to reside close to compute

Edge computing for sensors and detectors add new use cases

Developing a data discovery services that provides incentives for community datasets

Data policies that enable trustworthy AI

Data Challenges and Opportunities

- Data growing in size and complexity
- Data pipelines, staging and wrangling often dominate researcher time
- Data
- Data
- Edge computing for sensors and detectors add new use cases
- Developing a data discovery services that provides incentives for community datasets
- Data policies that enable trustworthy AI

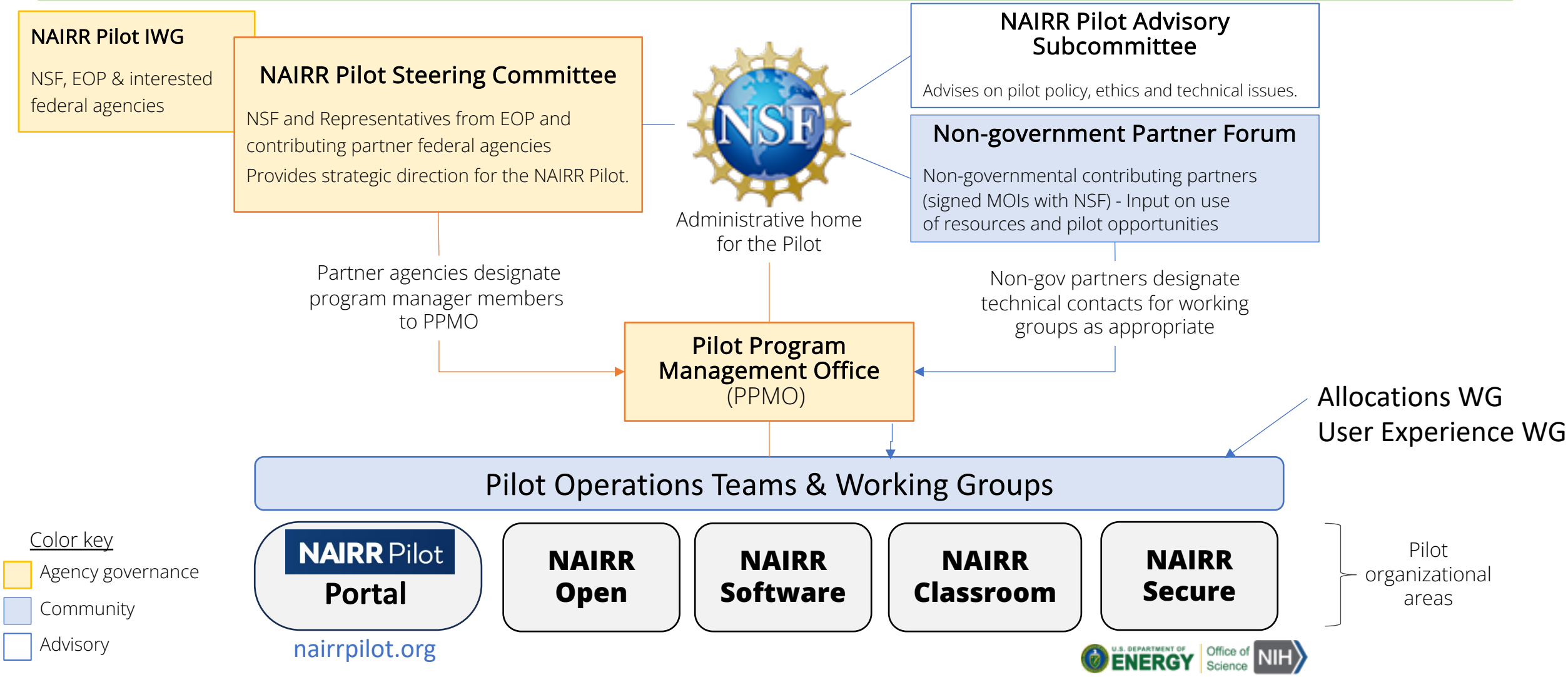
Our strategy in the pilot is to address data challenges through specific use cases and demonstration projects

Transparent and responsible AI will be a key focus of pilot

Advance
trustworthy AI

- Goal is to be transparent with NAIRR pilot processes
- Stand up an advisory sub-committee to aid with guidance on:
  - Evaluation of proposals and standards for NAIRR pilot contributions
  - Community outreach
  - Transparent operational policies
  - Training and User support
- Workshop with NIST and NIH on how pilot can support Trustworthy AI

NAIRR Pilot governance and operations organization



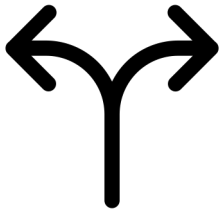
Many anticipated challenges…

Democratization: reaching broad communities

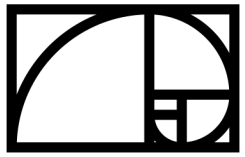
Assuring trustworthy & responsible AI in research space

Interoperability of resources

Data access, quality, curation, pipelines



Divergent software stacks



Applying design patterns across domains, NAIRR Open and Secure

On-boarding & user support

Questions and Discussion

Acknowledgements

- NAIRR Pilot Interagency Steering Committee
- NSF NAIRR Pilot Team
- Amy Walton
- Bill Miller
- Tess deBlanc-Knowles
- Sharon Geva
- Alejandro Suarez
- Daniel Bullock
- Varun Chandola
- Marlon Pierce
- Maria Fernanda Pembleton
- JD Kunda
- Dilma Da Silva
- Ellen Zegura
- Michael Littman
- Wendy Nilson
- Jeff Forbes
- Jim Donlan
- Sheikh Gafoor
- Juan Li
- Vivica Brooks
- Christine Christy
- Alice Kamens
- Kerstin Mukerji
- Gabby Cates
- Josh Chamot

- Allocations WG
  - Stephen Deems – PSC
  - Dave Hart – NCAR
  - Chris Keeley – UIUC
  - Bronson Messer – ORNL
  - Mike Norman – SDSC
  - Katherine Riley – ANL
  - Shava Smallen – SDSC
  - John Towns – UIUC
  - Veronica Vergara – ORNL

NAIRR pilot survey of researcher and educator use cases