Federal Register Notice 86 FR 46278, https://www.federalregister.gov/documents/2021/08/18/2021-17737/request-for-information-rfi-on-an-implementation-plan-for-a-national-artificial-intelligence, October 1, 2021.

Request for Information (RFI) on an Implementation Plan for a National Artificial Intelligence Research Resource: Responses

RFI Response

White House Office of Science and Technology Policy and National Science Foundation

Implementation Plan for a National Artificial Intelligence Research Resource

From the National Center for Atmospheric Research (NCAR)

Scientists and Software Engineers from NCAR have discussed the RFI on an Implementation Plan for a National Artificial Intelligence Research Resource, and are pleased to have the opportunity to respond.

Operations and Mission of the National Center for Atmospheric Research

NCAR's mission is "to conduct research that contributes to the depth of fundamental understanding of the atmosphere and its interaction with society and the environment and to develop and transfer knowledge and technology that expands the reach of atmospheric science." NCAR has a successful history of transferring technology and knowledge to U.S. government agencies, the private sector, and foreign governments. NCAR is eager to collaborate with other organizations on a shared computing and data infrastructure to provide AI researchers and students with access to a holistic, advanced computing ecosystem.

NCAR is operated by the University Corporation for Atmospheric Research (UCAR), a non-profit organization established in 1960 to oversee a wide range of programs and facilities that support its 120+ university affiliates and the national and international scientific community.

NCAR's current role in advanced computing, data ecosystem, and AI research

NCAR has a long-standing research program in artificial intelligence (AI) and machine learning (ML). That research began under externally funded projects and has helped lead the meteorology and climate community to recognize the efficacy of applying these methods to weather forecasting, hydrometeorology, climate simulation, and various applications, including aviation meteorology, renewable energy, surface transportation, forecasting in support of agriculture and food security, modeling for wildland fire management, and much more. Much like statistics, AI/ML provides a facility that can be used for understanding data, such as to describe the physics of the Earth system. In addition, these methods enable actionable science, such as by combining physical and social data to infer the likelihood of impacts. While care must be exercised in the training data sets to avoid bias, machine learning is well suited for tasks that are hard to describe formally or too expensive to compute at scale but where data are abundant or easily generated, a realm prevalent in Earth system science. Observation and remote sensing systems are collecting higher spatial- and temporal resolution data, and higher-resolution models are being run with more coupled processes to increase the realism of simulations. Our scientific ambitions overshadow the expected trajectory of our computing resources. Machine learning approaches promise a computationally efficient and scalable means to model the relationships between different data sets with more efficient and scalable computation.

More recently, NCAR has embarked on a major effort to expand the use of AI in environmental systems. For instance, the Computing and Information Systems Laboratory has teamed with other NCAR science labs, including the Research Applications Laboratory, High Altitude Observatory, and Climate and Global Dynamics Laboratory to build, demonstrate, and validate ways to replace certain physical models in our various modeling systems with AI emulators. The preliminary results of these initiatives are encouraging and are expected to lead to significant advances in model accuracy and efficiency.

Q1: What options should the Task Force consider for any of roadmap elements A through I above, and why?

Answers to roadmap topic D

i. **Develop an ecosystem of interoperable compute and data resources with easy and equitable access** - To provide Artificial Intelligence (AI) researchers and students across scientific disciplines with access to computational resources, access to the current national compute ecosystem needs to be simplified and hurdles that limit equitable access need to be reduced. Despite the efforts of XSEDE, the step of moving from on-campus compute resources to cloud and national resources is still nontrivial. A national effort can reduce this hurdle by providing consistent AI software stacks, user interfaces like Jupyterhub, and the integration and support of cloud approaches. Connecting, leveraging, and expanding the current national infrastructure, primarily NSF-funded resources, together with increased training and support, will democratize access for AI researchers. Domain-specific computational and data resources similar to the research data archive[1] at NCAR provide the weather and climate research community with connections to other NSF-funded centers integrated by cloud technologies; these could become the building blocks of a national AI compute data infrastructure. NCAR has played a crucial role in public-private partnerships, as shown in answer to question 5. NCAR is well-positioned to play a leading role in connecting the public, private and academic sectors to advance the state-of-the-science in AI research applied to Earth System predictability.

ii. **Develop shared public datasets and environments for AI training, testing, and benchmarking** – This is undoubtedly happening in the weather and climate community and outside of it. Several organizations interested in advancing the use of AI within the weather and climate community are actively working to identify and deploy high-quality datasets for AI training and comparison to increase the use of AI. For example, the American Meteorological Society (AMS) has fostered a Committee on the Application of Artificial Intelligence in the Environmental Sciences since the late 1990s and began teaching short courses on the topic in 2001, eventually archiving the lectures in a book. In 2008, that committee began holding AI forecasting contests, providing a shared dataset that researchers could use to test their algorithms. More recently, the AMS AI Committee has been banding together with other organizations to provide a complex series of datasets that researchers can explore and compare their applications against peers. When the committee turned to the Kaggle competition website in 2014 to host their contest, they were surprised to find that some non-meteorology AI experts applied techniques that outperformed those currently being applied in the weather community. Those novel techniques were quickly adopted by many others doing parallel research. Thus, the weather community has evidence that such common datasets are effective in helping to advance the state-of-the-science. The data sets will also provide an opportunity for computational scientists and data and compute systems researchers to inform the procurement of the next generation of national HPC systems.

In FY2021, NCAR was selected to serve as the long-term data repository for observations made by the NSF-funded Community Instruments and Facilities (CIF) program. This repository, known as the Geoscience Data Exchange (GDEX), will be an expansion in scope and capacity of the existing DASH Repository. GDEX will provide archival storage, data discovery services, open access, and citable DOIs for CIF data. NCAR is well-positioned to play a leadership role in providing Earth System data sets for the AI research community. Additionally, NCAR has partnered with a commercial cloud vendor to host a subset[2] of the Community Earth System Model Large Ensemble[3] (CESM LENS) dataset together with example notebooks demonstrating how to use the data. This approach democratizes the access to access and use of this data.

iii. **Expand the capacity of the national compute and data ecosystem to support the demands of AI** - The current NSF national systems are over-allocated and will not be able to provide enough capacity for the demands of

[1] https://rda.ucar.edu/
[2] doi:10.26024/wt24-5j82
[3] doi:10.1175/BAMS-D-13-00255.1

the AI research community. Investments in compute capacity as well as in innovative allocation and access models are necessary. An increase of capacity of high-quality scientific themed data repositories connected to advanced HPC and AI cyberinfrastructure will be necessary to support high-resolution output of community model runs like the NCAR-led Community Earth System Model (CESM) that then can be used for the AI research process. Machine Learning meta-data and new data standardized to enable automatic consumption by the AI research process needs to be created as part of the data design process. The expansion of the ecosystem should be driven by a co-design process between the AI research community, the data archives, and the infrastructure engineers.

iv. **Transdisciplinary collaboration** - The creation of a national AI research resource will require close collaboration among modelers, theorists, experimentalists, engineers, humanists, social scientists, computer and data scientists to create the data sets, training, and compute and data infrastructure necessary for the advancement of AI research. For example, storm-resolving model ensemble runs on pre-exascale and exascale systems will generate massive amounts of data that are too large to store fully. In situ training or lossy compression tailored for AI are approaches that need to be explored as part of the infrastructure roadmap. Convening the Earth System Science community and bringing different communities together is one of the strengths of NCAR, who would be glad to participate in leadership roles in this domain.

v. **User interfaces** User interface design has produced alternatives to the command line interface such as Jupyter notebooks and science portals (eg. Open OnDemand and Science Gateways) that reduce the learning curve for first-time users, enhance collaboration, and simplify visualization, documentation, dissemination, and reproduction of scientific results. As these tools evolve, researchers need to navigate the perennial tradeoffs between ease of use vs. greater control over the underlying hardware and software. One of the goals of the national AI research resource should be to make AI at scale easy for researchers to conduct.

vi. **Training, education, and workforce development** - The nation should leverage existing educational tools at different academic and research institutions to create a central repository/resource for ML education and training. Resources should focus on training upcoming scientists and researchers in ML (i.e., students, early-career), as well as training, experienced scientists who are interested in applying ML in their respective areas of expertise. Examples of educational resources can include tutorials and summer schools hosted by universities and research laboratories and educational modules (following the UCAR COMET program). The AI4ESS summer school in 2020[4] and the TAI4ES summer school in 2021[5] are examples of educational activities that are necessary to bring participants up-to-speed on how to develop trustworthy AI for the Earth & environmental sciences.

vii. **Better understand the national AI R&D workforce needs** – AI is becoming critical to our scientific approaches. We need people capable of building and

[4] https://www2.cisl.ucar.edu/events/summer-school/ai4ess/2020/artificial-intelligence-earth-system-science-ai4ess-summer-school

[5] https://www2.cisl.ucar.edu/tai4es

maintaining the computational infrastructure necessary to continue to push the state-of-the-science. These people must know how to deal with “Big Data” at all levels, including how to manage datasets. When AI is used in production involving the types of big datasets that we often use in meteorological applications (e.g., blending in situ and remote observations with output from many models smartly to improve forecasts), this aspect becomes as important as producing the new AI algorithms. This personnel trained in AI and Big Data must be prevalent in public positions as well as in private enterprise. In the meteorology community, several universities are providing training in AI at the undergraduate as well as graduate levels. National laboratories must emphasize applications of AI. At NCAR, we have re-invigorated a cross-laboratory effort in AI, both in basic research and applications that we have fostered over the years. We firmly believe that applying these principles more broadly across a range of topics is needed to assure continued advances in our understanding and in our applied technology.

Q2: **Which capabilities and services (see, for example, item D above) provided through the NAIRR should be prioritized?**

Prioritization in the following order (see details of the buildings block in the previous answer):

- Priority 1: (iv) Transdisciplinary collaboration: Only close collaboration between a diverse set of domains of expertise will enable a democratized and equitable AI research resource.
- Priority 2: (ii) Develop shared public datasets and environments for AI training, testing, and benchmarking: To advance the use of AI, it is essential to identify and deploy high-quality datasets that are sufficiently described for automatic model training and testing.
- Priority 3: (vi) Training, education, and workforce development: A central training resource is important to provide equitable access to training materials connected to the high-quality data sets to train upcoming scientists and researchers in AI, as well as training experienced scientists who are interested in applying AI in their respective areas of expertise.
- Priority 4: (i) Public and private partnerships: The compute and data ecosystem necessary is only achievable by synergistic partnerships between the private and public sectors.
- Priority 5: (iii) Expand capacity: The current NSF national systems are over-allocated and will not be able to provide enough compute capacity for the demands of the AI research community. Investments in capacity and integration with commercial cloud providers and innovative allocation and access models are necessary. The increase of capacity of high-quality scientific themed data repositories connected to advanced HPC and AI cyberinfrastructure will be necessary to advance AI research.

Priority 6: (v) User interfaces: Evolve the existing JupyterHub ecosystem, to support AI research at scale through an easy-to-use interface.

***Q3:* How can the NAIRR and its components reinforce principles of ethical and responsible research and development of AI, such as those concerning issues of racial and gender equity, fairness, bias, civil rights, transparency, and accountability?**

Our organization has successful examples of approaching this challenge. Specifically, NCAR is a key partner in the NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES). As part of this institute, we are developing novel methods of evaluating the trustworthiness of AI systems for a diverse array of geoscience use cases. Evaluating the trustworthiness of AI systems for environmental science is important because these tools if adopted will be used for high-stakes decision-making that could greatly impact lives and property. A key aspect of the work of this center is involving social scientists to better understand how the stakeholders understand trustworthiness and working directly with end-users to help develop ways to meet their trust. Another critical aspect of this work is clearly defining the components of trustworthiness for a given set of problems and consistently evaluating each system to ensure that trust is warranted. The three focus areas for trustworthy AI in AI2ES are ensuring that AI is explainable, physics-based, and robust.

Weather and climate observation and prediction data are also subject to racial and gender equity biases and AI systems trained with this data could also risk propagating these biases. Globally, weather observations are far denser in the United States and Europe, resulting in poorer characterization of current weather and forecasts in other parts of the world. Satellite measurements have reduced this gap when forecasting larger spatial scales, but there are still systematic biases in observing local weather patterns. Poorer urban and rural areas are less likely to have dense weather observations and may be less likely to report extreme weather events. These data issues could result in AI systems that also perform more poorly in these areas. AI model predictions should be evaluated for these kinds of systematic biases, and targeted observation campaigns may be necessary to collect more verification data in these data-sparse but heavily-populated regions.

Q4: What building blocks already exist for the NAIRR, in terms of government, academic, or private-sector activities, resources, and services?

Many of these aspects have been mentioned in the answer to Q1. Specifically, some existing building blocks include:

1. XSEDE service providers, such as the NCAR HPC ecosystem that support the atmospheric sciences community.
2. Training efforts at Universities and NCAR. As noted above, many excellent educational resources already exist that could be advertised widely and built upon. Many universities offer on-line classes and prior summer schools such as those offered by NCAR in collaboration with our university members and our partners in the public and private sectors, with archived content for use by future students.
3. Part of NCAR's role in the Earth System Science community is as a convener of the Community - we bring people and resources together to accomplish goals that are important to society.
4. In addition to convening the community, it is essential to provide continuing support. Our own approach in this role for modeling and observational realms is well poised to be scalable to broader communities.
5. As mentioned above, NCAR already provides leadership in the geoscience community for archiving Earth System data. We envision potential to grow services to include preparing digital notebooks that enable accessing the data and assessing the results of ML applications relative to a baseline, i.e., testing and verification frameworks.

Q5: What role should public-private partnerships play in the NAIRR? What exemplars could be used as a model?

Partnerships, including the academic sector in addition to public and private, are imperative to advance NAIRR. Each sector brings important, unique expertise and considerations relevant to their own needs.

As an example, NCAR formerly led a Public-Private-Academic Partnership to Advance Solar Power Forecasting (funded by DOE) that included three other national laboratories, six universities, and 11 private sector partners as well as several international affiliate partners. All sectors participated in crafting the initial vision to ensure that the end product was useful to the end-users and that all research was accomplished in a robust, open, and repeatable manner. This team found that the key to success was communication and that employing methods from the social sciences facilitated conversation and helped the team members to better understand each other. The final outcome included new models for solar power forecasting, both those based on physical principles and those based on machine learning. It required all sectors working together to adequately design and test their usefulness for the intended application.

Above, we described the current AI2ES partnership with academic, private, and public sectors that is working toward advancing trustworthy ML use in the environmental sciences. In similar ways to these two projects, we believe that the NAIRR will require working across sectors to provide the hardware, software, data, tools, and secure access needed to make NAIRR successful.

Q6: Where do you see limitations in the ability of the NAIRR to democratize access to AI R&D? And how could these limitations be overcome?

The security aspect of NAIRR will be critical to success if the goal of easy and democratic access is to be achieved without compromising software and data integrity. It is difficult to make large systems open to use by students and the public without opening access to bad actors. The right organizations must be involved to ensure the security of necessary data, user information, and applications that are archived.

Secondly, as we wish to democratize AI and data, it will be necessary to train some users in basic computer usage, programming, and database access. The planners should think through how to help users get started without taking an inordinate amount of time from those charged to help users. One opportunity could follow from the user interface perspective highlighted above. Specifically, tools like Jupyter notebooks can lower the barrier to accessing data, algorithms, and compute environments, and these have been successfully demonstrated in training programs for novice users.