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

NASA Feedback RE: National Artificial Intelligence Research Resource (NAIRR) Request for Information (RFI).

I. Strategic Level Inputs:

- NASA sees AI as a powerful capability for the United States, and therefore encourages the NAIRR as it seeks to bolster AI research and partnerships
- NASA's resources are constrained on priority missions directed by Congress and no spare NASA resources are available to contribute to the NAIRR
- NASA flagship missions such as the Artemis campaign rely on technology that has been substantially matured and tested by researchers and technical experts.

II. Detailed Inputs, based on the NAIRR RFI. NASA has embedded answers under each RFI question.

1. What options should the Task Force consider for any of the roadmap elements A through I above, and why? [Please take care to annotate your responses to this question by indicating the letter(s) of the item (A through I in the list above) for which you are identifying options.]

A. Goals for establishment and sustainment of a National Artificial Intelligence Research Resource and metrics for success.

NASA: The options should include establishing:

- *A Center of Excellence to oversee implementation*
- *Training and education program to develop workforce*
- *Framework and processes for reuse and reproducibility of AI experiments*
- *Incentivization for open sharing and reproducible results*
- *A method to rate capability maturity to guide adoption when ready*

B. A plan for ownership and administration of the National Artificial Intelligence Research Resource, including:

i. An appropriate agency or organization responsible for the implementation, deployment and administration of the Research Resource; and

NASA:

One possible model is for industry and academia to largely run and resource the NAIRR, with government advocacy, participation, and inspiration

ii. A governance structure for the Research Resource, including oversight and decision-making authorities;

NASA:
Based on B.i. (above), A possible model is for the government to encourage industry and academia to run the NAIRR, to include business & academia providing most oversight and decision-making functions. The government may not have to provide program management,

but could instead focus on NAIRR advocacy, participation, and inspiration. This is one option; not all of NASA's answers to other questions assume this option.

Another option could be the NIST model. See https://www.nist.gov/artificial-intelligence.

Regardless of the mechanism for governance, an independent oversight function would be necessary; an organization to ensure the public interest is respected as technologies develop.

C. A model for governance and oversight to establish strategic direction, make programmatic decisions, and manage the allocation of resources;

NASA: The options should include:

- *Long-term funding model to support and manage allocation of resources*
- *Collaborative governance with representation from industry, academia and governments*
- *Strategic planning meetings, reports, success stories, and new challenges*
- *The Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) solicitations framework is an option that could, at least in part, be used a model for governance and managing the allocation of resources*
- *The formation of tightly compartmentalized teams focused on well-scoped and discrete tasks, coming together to discuss and disperse innovations and insights*

D. Capabilities required to create and maintain a shared computing infrastructure to facilitate access to advanced computing resources for researchers across the country, including provision of curated data sets, compute resources, educational tools and services, a user-interface portal, secure access control, resident expertise, and scalability of such infrastructure;

NASA: Many options to address items in this element are being advanced by industry. We should collaborate with the industry to adapt their solutions to meet our needs recognizing regulatory and legal requirements may require unique investment from the government to realize.

E. An assessment of, and recommended solutions to, barriers to the dissemination and use of high-quality government data sets as part of the National Artificial Intelligence Research Resource;

NASA Science Mission Directorate's (SMD's) Strategic Data Management Working Group's Report on Groundbreaking Science is an option to refer to which addresses this element.

Also, consider creating policies, procedures and sandbox environments to make it easier for industry, academia, and government to team on shared problems. For example, guidance on sanitization of large sensitive datasets to allow for collaboration outside of government.

NAIRR should iteratively poll the user, participant, and customer community to ensure it is providing what is needed. Deliver what the people want by asking the people what they need.

F. An assessment of security requirements associated with the National Artificial Intelligence Research Resource and its management of access controls;

As some NAIRR capabilities mature, methods would be needed to on-board them into selected sensitive government programs; mechanisms for traversing sensitivity levels would be necessary.

G. An assessment of privacy and civil rights and civil liberties requirements associated with the National Artificial Intelligence Research Resource and its research;

H. A plan for sustaining the National Artificial Intelligence Research Resource, including through Federal funding and partnerships with the private sector;

NASA: NASA's Mission organizations leverage industry and academia via a variety of engagement mechanisms, some of which might allow NASA to become a customer of NAIRR participants. As one example of many, NASA SMD's open source science program includes targeted solicitations, incentives, and collaborations with private sectors.

I. Parameters for the establishment and sustainment of the National Artificial Intelligence Research Resource, including agency roles and responsibilities.

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

NASA priority list of capabilities and services are as follows (order is not implied):

- *Curated training datasets and metadata registries*
- *Computing resources*
- *Educational tools and services*
- *Resident expertise, to include AI, ML, scientific domains, cybersecurity, social bias and IT*
- *Scalability of infrastructure*
- *A taxonomy of levels of AI capability*
- *Standards to support interpretability and interoperability*
- *Free open source libraries and APIs for relevant capabilities (*Note: if government investment helps propel the NAIRR, this makes a good case for broad, free sharing)*
- *Secure access control (where necessary)*
- *A user interface portal for accessing and finding resources*

3. 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?

NASA Recommends:

- *To develop practical guidance for ethically applying AI and forming mechanisms to encourage & ensure adherence as needed*
- *To develop a checklist for best practices*

- *To conduct independent ethics review of research*
- *To develop systems to examine unintended consequences of AI research*
- *To include additional ethics experts into discussions where decisions are made*
- *Develop a portfolio of federal entities looking to be responsive to actions such as the executive order related to Advancing Racial Equity and Support for Underserved Communities Through the Federal Government, targeting inclusive AI can leverage best practices, industry resources and community contacts for strategic federal mission focused initiatives. Link the NAIRR with other initiatives for complementary effect.*
- *Participate in global and national ethical AI discussion and debate; this area is a work-in progress; NAIRR can both contribute-to and benefit-from the larger ethical AI discussion*
- *The NAIRR and its components should establish a set of standards and guidelines for usage of data sets. These guidelines should define the boundaries for the research using NAIRR, taking issues of ethics and transparency into account. In addition, standards should be established to mitigate impacts of biases inherent within data sets when used by machine learning algorithms*
- *Provide for resident personnel with expertise and knowledge in how AI and ML systems have exhibited bias that has led to social harm – personnel that work integrally with educational tools and services*
- *Recognize that ethics are subjective and vary by culture. Since there is no canonical definition of "ethical", the goal for any software system should be responsible and transparent behavior. This holds whether that system incorporates AI or not*

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

NASA: The following building blocks exist across the technical community:

- *High end computing platforms for AI*
- *Procurement of cloud computing platforms to scale AI in production*
- *Open data and open source policies*
- *Competitive programs targeted towards AI*
- *Public and private partnerships with relevant partners*
- *The European Union's coordinated AI plan, which can serve as a benchmark*
- *The Small Business Innovative Research (SBIR) and Small Business Technology Transfer (STTR) programs could provide channels for pursuing some NAIRR work*
- *Federal programs and partnerships, such as the NSF National AI Research Institutes, could serve as existing building blocks for the NAIRR*
- *DARPA's Explainable AI work could contribute to overall NAIRR approaches*
- *NASA's Space Apps Challenge is a good example of global collaboration and crowdsourcing in emerging technology spaces, to include AI*

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

NASA:

Most of the innovations around AI are happening in the private sector and academia. The role of the partnership is invaluable in bringing the experts and tools and services to solve problems within the government. The partnership can also provide avenues to fill in necessary AI workforce within the government.

NASA's Earth Science Data Systems (ESDS) Program has established strategic partnerships with public and private companies to help further its data management and data development efforts through non-reimbursable Space Act Agreements. Some examples of our success and outcomes can be found at https://earthdata.nasa.gov/collaborate/esds-public-private-partnerships.

NASA SMD and NOAA's NESDIS division have formed a working group to address cloud architecture data access and discovery challenges. The group meets quarterly to share knowledge, and status on joint initiatives focused on developing data governance (provenance) guidelines and common practices through the data life cycle in the cloud. Some examples include a joint cloud-based data expedition pilot to demonstrate interoperability between NASA and NOAA data.

The Frontier Development Lab, an industry-academia NASA-focused external innovation lab, is another good example. The Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programs is another example that could be used as a model for public-private partnerships in the NAIRR.

6. 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?

NASA:

Limitations include: seamless access to computing platforms to accelerate AI, availability of high quality training datasets, and reproducibility of AI experiments. To overcome these limitations, we need to prioritize AI across agencies and develop an AI strategy that addresses these limitations. Furthermore, we need better collaborations with academia and the private sector to develop workforce for the future and exchange knowledge. We also need to create an environment where domain experts and AI experts can work together. To accelerate adoption of AI we need to foster the culture of open sharing and collaboration. Some NASA missions include high-cost, low sample size data; techniques for adapting AI to these data sets are needed. The AI community must also find ways to communicate and show the value of AI-enabled solutions to mid-late career workers who are comfortable with more traditional techniques.

Scientific research depends upon collaboration and sharing of results, but heightened security barriers around a system cause obstacles that limit access by the scientific community. While security is critical in certain areas, easily accessible tools and data that can be easily shared to enhance collaboration is a critical aspect to accessible science. One solution to this problem is providing different tiers of access. These would have similar capabilities, but having a "public" tier with public data and compute that is accessible to anyone through standard, community based login methods would make these tools more accessible. There can then be tiers with

more increasing level of security, but with more resources, access to confidential information, and even restricted tools.

Supporting the open science community: Providing publicly accessible data and software under open and permissible license significantly democratizes the access to data. The fact that people anywhere can download the data and software for AI/ML research is probably the greatest democratization process in place. As these are available without any gateways, these are incredibly accessible. However, the teams that are producing the data need the resources to produce data in AI/ML accessible formats with metadata, that is well labeled, and has accessible documentation. Teams that are producing the open source software that underlies much of AI/ML are often not well supported, especially the libraries that are underpinning the basic structure. Sustainable funding to support both maintenance and innovation is necessary for the software development. One further gateway to lower is that the documentation and training is available to make these resources accessible.

Limitations may exist with respect to democratizing access to communities, institutions, and regions that have been traditionally underserved and underrepresented with regard to AI research. Limited access to resources/hardware/infrastructure (i.e, digital divide) necessary to leverage NAIRR capabilities and services may present additional challenges. These limitations could be addressed through targeted outreach to centers, historically-underserved universities, programs, and conferences that place emphasis on engaging underrepresented communities in AI research and educational activities.

Failure to elevate AI as a profession could be a barrier to progress. The NAIRR can bolster National AI progress by fostering AI as an emerging top-tier profession. With academia participation, there is a great opportunity to craft a variety of AI-related professional development materials & activities to foster broad and deep AI knowledge and skills across all elements of the United States. With industry, academia, and government participation in the NAIRR, there is also great opportunity to foster other aspects of AI as a profession, such as ethical AI debate & guidance, fostering powerful & useful standards, self-policing AI-related activities, and hosting workshops / conventions, etc.

Question 6 also generated some discussion regarding whether it is the right question. Should the NAIRR actually focus on democratizing access to AI R&D? Can access to these capabilities even be controlled or limited? Perhaps part of this function would be matching NAIRR resources to customers by expertise, problem set and data type. Another possibility would be for the NAIRR roadmap to focus on overcoming limitations by eliminating economic biases such as those present in our public education system in addition to systemic biases of fielded systems (e.g., Tech-company branded credit cards, facial recognition, COMPAS). Solutions could include exposing researchers, students, and the general public to AI resources such as educational tools and data sets as well as application-appropriate benchmarks for fielded systems.