**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

**Aishik Ghosh**
[redacted]

**To**
*Attn: Wendy Wigen, NCO, 2415 Eisenhower Avenue,*
*Alexandria, VA 22314, USA.*

August 31, 2021

I am a researcher at the Department of Physics and Astronomy, University of California, Irvine and Physics Division, Lawrence Berkeley National Laboratory, developing AI models that help with our broad spectrum of physics research. I collaborate with the Organisation for Economic Co-operation and Development (OECD) in Paris on topics such as the role of AI in current and future scientific productivity and AI policy, and I am also involved in field of AI Ethics. During my time in France I also informally advised the French computing centre CCIN2P3 (https://cc.in2p3.fr/en/) on how to make their AI resources more accessible to researchers in France.

The National Artificial Intelligence Research Resource (NAIRR) is a welcome step to democratize the access to AI research and development infrastructure. Thank you for offering the opportunity to provide comments to the task force. I am responding to the Request for Information (RFI) on an individual capacity. My response to the topics mentioned in the RFI document are given below alongside the topic numbers and letters.

1.

A: In additional to the usual metrics, it is important to measure the diversity of users of the resources. For example the number of research groups (of various academic disciplines) that are embarking on their first AI related project with the help of NAIRR, the fraction of users from smaller research institutions compared to large institutions that already have excellent resources, and the number of individuals without advanced academic degrees who are able to take advantage of the educational resources to start a project. Users from institutions that already have excellent resources would indicate that NAIRR is truly competitive in terms of user-friendliness.

While several small research groups currently maintain their own GPU mini-clusters. This is not efficient because of two reasons. First, it takes significant human time to maintain them for a part-time system administrator. Second, the resources are under-utilized, often free more than 50% of the time. A reliable common national resource may reduce the need to maintain separate mini-clusters. A survey of how many researchers have switched from maintaining their own mini-clusters to using NAIRR would certainly be a useful metric to monitor.

C: While maintaining high utilization for such a large resource requires automated protocols, it must also include the flexibility to request/book for period of guaranteed access to certain resources (such as a single GPU node). This service is offered by the French computing centre CCIN2P3, which is excellent for PhD students at a critical stage of their thesis, before critical conference deadlines and so on.

E: Datasets curated for publicly funded research could be made accessible, in certain cases after an appropriate embargo period.

2.

Educational resources and a software start-up kit targeted at students and researchers with an existing scientific and programing background should be a top priority. There are low-hanging fruits in nearly every scientific domain when it comes to the application of AI. These solutions could improve the state-of-the-art or to tackle a research question in a new way that was computationally infeasible using traditional methods. For many graduate students, the barrier to trying an AI based solution is the large upfront investment of time required to get started. This involves find appropriate learning resources targeted towards researchers

in their particular field and setting up a software pipeline. The time taken for both of these activities are minimized in larger institutions (such as Berkeley Lab) where tutorials and a software start-up kit are readily available. However, for best results, the NAIRR should also host an annual workshop for various researchers to meet in person, exchange ideas, expertise and forge multi-disciplinary collaborations. A small fund (for travel or secondments) could be specifically allocated for collaborations that are forged through the NAIRR workshop. With these measures the NAIRR could boost research productivity and increase the number of innovative ideas that come from smaller research institutions all over the nation.

In addition, educational resources could be made available for people without advanced degrees but with prior programming experience. There are currently far too many application cases for AI compared to the number of individuals with technical know-how. The problems for small communities may not be of immediate interest to large researcher groups or companies but they could be addressed by individuals of that community or organizations working for a social cause, if they have easy access to such resources.

3.

It is absolutely essential for NAIRR to insist that users consciously consider the unintended uses or consequences of their innovations. Warnings may be given along with the publication of material. Educational resources must include components on AI Ethics as well as interpretability and tests for robustness of models. These will be particularly useful for smaller organizations that do not have the resources to employ an in-house AI Ethics expert. A common resource for users who want to understand the ethical implications of their projects may also be provided, either in terms of access to experts or informal discussion forums comprising of other NAIRR users.

6.

AI research evolves much faster than research in traditional academic fields and it is difficult for even the top researchers to keep up with the state-of-the-art. Smaller research groups, particularly those working on applications of AI in their domain of expertise, are unlikely to be able to keep up, despite the support of NAIRR. One possible solution is to host an annual workshop for all NAIRR users with the explicit goal of facilitating the exchange of ideas and forging of inter-institutional, multi-disciplinary collaborations. A small fund (for travel or secondments) could be dedicated to collaborations forged during this workshop.

In closing, I believe the NAIRR is potentially an invaluable initiative to democratize infrastructure that fuels AI innovation. It could help accelerate both small and large research projects and have positive impacts at various levels of society. However, it is essential not only to keep the ethical implications in mind while forming these policies but also to keep them relevant with periodic updates in the future. A successful NAIRR could become a template for other countries and political unions to follow.

Thank you again for providing this opportunity.

Yours Sincerely,

**Aishik Ghosh**
Postdoctoral Scholar,
University of California, Irvine
CA 92697, USA
&
Affiliate,
Lawrence Berkeley National Laboratory,
1 Cyclotron Rd, Berkeley, CA 94720, USA