

## **Computing Community Consortium’s Response to RFI “[Implementing Initial Findings and Recommendations of the National Artificial Intelligence Research Resource Task Force](#)”**

**June 30, 2022**

**Written by:** Liz Bradley (University of Colorado, Boulder), David Danks (University of California, San Diego), Bill Gropp (University of Illinois Urbana-Champaign), Daniel Lopresti (Lehigh University), Ufuk Topcu (University of Texas at Austin), Ann Schwartz (Computing Community Consortium), and Cat Gill (Computing Community Consortium).

**This response is from the Computing Research Association (CRA)’s Computing Community Consortium (CCC). CRA is an association of nearly 250 North American computing research organizations, both academic and industrial, and partners from the professional societies. The mission of the CCC is to bring together the computing research community to enable the pursuit of innovative, high-impact computing research that aligns with pressing national and global challenges.**

The Computing Community Consortium commends the National Artificial Intelligence Research Resource (NAIRR) Task Force on the release of the interim report. Overall, we believe that the interim report lays the foundation for a national resource for the research community. In addition to our earlier comments, submitted as a response<sup>1</sup> to the initial RFI on the National Artificial Intelligence Research and Development Strategic Plan, we would like to share the following:

One issue we took with the NAIRR interim report is the narrow scope of ethical and societal considerations. The report listed only “privacy, civil rights, and civil liberties” as qualities that should be protected by the NAIRR resource. This limited definition neglects a plethora of considerations, such as the trustworthiness of the data and resulting models; degree of harmful or damaging biases; or entrenching structural inequities. We recognize that the NAIRR’s charge mentions only these three terms (rather than the broader “ethical and societal considerations”), which may explain their focus in the interim report. Nonetheless, we recommend that these three terms be interpreted as broadly as possible to ensure that the NAIRR, as well as the data and models that result from it, are sensitive to a wide array of ethical and societal concerns, not only those that fall under a narrow interpretation of “privacy, civil rights, and civil liberties.”

An issue we find absent from the interim report is a clear, explicit plan to collaborate with and take advantage of commercial providers. Many of these providers, such as Google, Microsoft, and Amazon, have already created AI research resources that are available to the public. The interim report states that NAIRR does not intend to rework any preexisting AI research resource

---

<sup>1</sup><https://cra.org/ccc/wp-content/uploads/sites/2/2022/03/CCCs-Response-to-RFI-Information-to-the-Update-of-the-National-Artificial-Intelligence-Research-and-Development-Strategic-Plan.pdf>

frameworks if they can be more-easily acquired and implemented using a third-party product. However, the report gives no details on how a third-party product would be implemented, whether the third party would be involved in regulating and/or overseeing the NAIRR resource, or any information at all regarding a joint effort between NAIRR regulating agencies and commercial providers.

Something that is not considered in the interim report, but that we feel ought to be addressed, is the inclusion of surrogate and synthetic datasets in the NAIRR resource. In many instances, the data required to train AI models may be too sensitive to share (e.g. health records, proprietary business data, data in support of national security and public safety, etc.). While the development of privacy-preserving techniques is undoubtedly an important research topic, current techniques are not sufficient to protect those kinds of data, which are too sensitive to share at scale. For AI models that are reliant on sensitive data, this would mean pausing research until protective-enough methods to disseminate sensitive data are invented. Since losing time to work on these matters would be damaging, we instead recommend that appropriate resources be included in NAIRR to allow these models to be trained on synthetic and/or surrogate datasets and then employ, for example, transfer learning to address the original motivating application.

Regarding the plan for the NAIRR Ownership, Administration, and Oversight, a “federated approach” to regulating the NAIRR resource is proposed many times, though guidance on how to achieve this federated approach is lacking. Four options for ownership and administration are proffered in the interim report; however, these options are not elaborated on, and it is left unclear as to which option the NAIRR Task Force recommends. It is also unclear to the CCC Council whether a resource or infrastructure of this scale has ever been so centrally managed as the interim report seems to propose. We strongly encourage the final report to elaborate on the options for ownership and administration enabling transparency between those institutions which will oversee the NAIRR resource and between the NAIRR oversight committee and the public.

Another factor we find missing from the report is the opportunity for the NAIRR user base to be used as an educational resource. Throughout the “Educational Tools and Services” portion of Section 4 of the interim report, many suggestions for educational tools, such as instructional text and video, and chat rooms with NAIRR employees, are recommended. The report does not, however, consider the NAIRR user base itself to be a valuable resource. This user base will be comprised of both novice and expert AI researchers, some of whom will have had extensive experience with multiple coding languages, software systems, and infrastructures similar to the NAIRR resource. As explained in the 20-Year Community Roadmap for Artificial Intelligence Research in the US<sup>2</sup>, the computing research community feels strongly that a community

---

<sup>2</sup> <https://cra.org/ccc/wp-content/uploads/sites/2/2019/08/Community-Roadmap-for-AI-Research.pdf>

resource, such as the NAIRR, can and should enable training for an AI workforce. The final report should outline how not only the NAIRR staff, but also the NAIRR users, will come together to create the infrastructure for a national AI workforce, in terms of both human and intellectual infrastructure, to better enable the broad research ecosystem.