



30 June 2022

To: Jeri Hessman
National Coordination Office for Networking and Information Technology Research and Development
2415 Eisenhower Avenue
Alexandria, VA 22314
NAIRR-responses@nitrd.gov

Re: RFI on Implementing Initial Findings and Recommendations of the National Artificial Intelligence Research Resource (87 FR 31914)

IEEE-USA is pleased to submit comments in response to the findings and recommendations of the interim report by the National AI Research Resource Task Force.

The input provided below represents the expertise of the IEEE volunteer members who are living and working in the US, and who are actively conducting research and development into artificial intelligence (AI), software engineering, cybersecurity, and advanced computing. As a community of researchers, and developers, IEEE-USA strongly supports the efforts of the White House Office of Science and Technology Policy (OSTP) to ensure diversity in the R&D community, and to create a roadmap that will enable expanded access to resources, data, testbeds and associated tools for all researchers, students, and developers of AI systems.

We applaud the efforts of the Task Force and thank them for their hard work in fulfilling your 2020 Congressional mandate to create this roadmap as part of the National AI Strategy. IEEE-USA has provided specific thoughts and recommendations below. However, our overall impression is that the report could benefit from streamlining and better integration of its ideas. The report captures quite a few excellent findings, but enhanced continuity would help stakeholders better understand the impact of the report's recommendations. IEEE-USA suggests employing planning practices to the strategic objectives to streamline the roadmap and capture metrics of success. We also found that the document seemed to be duplicative in areas; this is the case for some of the recommendations regarding creation and maintenance of the resource infrastructure in Chapter 4. We also believe the document could benefit from the key recommendations and associated actions being set out as a roadmap in chart form.

The concept of explainable AI (XAI) is missing from the document. In various publications, the US government addresses the challenge of ensuring that stakeholders - both users and those impacted by AI systems - understand and trust algorithmic outcomes; see for example.^{1, 2, 3, 4} IEEE-USA recommends that NAIRR address XAI as an important research element to ensure that research outcomes produce solutions to problems that are understood by humans. Ensuring that XAI is a goal helps to replace 'black box' solutions in machine learning, where even the designers of the AI application cannot explain why the AI application reached a specific decision.

IEEE-USA is strongly supportive of the Task Force's work and believes this roadmap represents needed guidance for our AI innovation ecosystem. We thank OSTP and NSF for considering these comments and welcome further discussions with the agency on these matters. If you have questions, please do not hesitate to contact Erica Wissolik at (202) 530-8347 or e.wissolik@ieee.org.

1 DARPA: Explainable Artificial Intelligence (XAI) <https://www.darpa.mil/program/explainable-artificial-intelligence>. Accessed 29 June 2022.

2 Congressional Research Service: Artificial Intelligence: Background, Selected Issues, and Policy Considerations <https://crsreports.congress.gov/product/pdf/R/R46795>. Accessed 29 June 2022.

3 From Ethics to Operations: Current Federal AI Policy <https://atarc.org/wp-content/uploads/2022/01/Current-Federal-AI-Policy-Assessment-FINAL.pdf>, see page 8. Accessed 29 June 2022.

4 NIST: AI Fundamental Research - Explainability; <https://www.nist.gov/artificial-intelligence/ai-fundamental-research-explainability>. Accessed 30 June 2022.

IEEE-USA's comments on the interim report:

a. A Vision for the NAIRR. (Chapter 2)

NAIRR User Base

IEEE-USA thanks the NAIRRRTF for specifically mentioning students at community colleges among the users who should receive NAIRR support. Community colleges are a necessary part of the innovation economy and specifically including these students encourages diversity in the future American workforce.

Additionally, furthering the strategic objective to democratize the US AI innovation ecosystem, we believe that the NAIRR should emphasize its support for AI research at Minority Serving Institutions that serve historically underrepresented populations. For example, The University of Texas at San Antonio - which offers an excellent multidisciplinary studies degree in Artificial Intelligence - is a Minority Serving Institution due to its high number of Hispanic students.⁵ Supporting its programs, and those of similar institutions, is among the most efficient ways of pulling underserved populations into the AI workforce.

b. Establishment and sustainment of the NAIRR. (Chapter 3)

Ownership and Administration

IEEE-USA recommends that the final report include a section that clarifies the NAIRR's management and governance structure. While the interim report alludes to the creation of a management structure and offers several alternatives, we ask that the NAIRR recommend the establishment of a shared AI resources national center management team. The structure could include a combination of the proposed alternatives to ensure that all stakeholders, in both the public and private sectors, can participate. This acknowledges the finding that no single current agency should form a new division for the reasons stated, while giving clear guidance to Congress and the Administration.

Resource Allocation and Sustainment

Recommendation 3-10

The requirement that access to NAIRR resources be contingent on research project proposal review increases the barrier of entry for students attempting to enter the field. Instead, NAIRR should look at ways to reach students in community colleges and perhaps even training programs like RAMTECH - a robotics and advanced manufacturing training center in Ohio. These stakeholders likely lack the know-how and resources to write grants.

To ensure that the research community attracts a diversity of talent, access should not be limited to only those practicing an advanced state of the art. IEEE-USA recommends that access to NAIRR resources be accessible via a tiered criteria based on level of understanding. Tiered access, where larger projects requiring more resources would have higher proposal submission requirements while smaller projects would be subjected to less stringent proposal requirements, would invite new researchers into the field, opening opportunities to less experienced professionals.

Additionally, a free "on-demand" cloud instance which serves as a sandbox (a safe environment where students can engage and learn with no financial grant) for those who may have an interest in AI systems. This could be similar to Google's Colaboratory.

⁵ IEEE-USA Position Statement, *Artificial Intelligence: Jobs, Education, Workforce, and Diversity*, November 2021.

Ultimately, tiered access would make the NAIRR available to high level research scientists, university and community college students, and those who are just beginning to be interested in AI technologies. Since innovation can come from anywhere, the NAIRR must be open to participation from anywhere.

NAIRR Performance Indicators and Metrics

IEEE-USA recommends establishing specific measures in the roadmap that clearly explain how it will be implemented, including timelines and metrics for success. We agree with the interim report recommendation that determining value and impact of successful research programs is facilitated by clearly defined and measurable goals at the outset.

The simplest starting point for a rollout could be divided into four phases:

- Phase one - assembling the cyberinfrastructure (i.e., hardware, servers, user interface).
- Phase two - roll out platform at the community college level for debugging / beta testing. We suggest this because incomplete products at higher levels could result in loss of interest in the platform.
- Phase three - university level adoption of a platform.
- Phase four - industry level adoption.

The interim report does not address educational outcomes associated with Figure 1, Chapter 1. To ensure that the NAIRR reaches the target audience, characteristics such as gender, age, ethnicity, and program of study (i.e., engineering, physics, mathematics, biology) should be reported to demonstrate diversity in the population. These metrics could be valuable in identifying and targeting areas where the program performs weakly. This is possibly discussed in Recommendation 3-18, but the wording is unclear.

As the interim report states in the strategic objective, the NAIRR should strengthen and democratize the American AI innovation ecosystem. To successfully achieve this objective, the NAIRR must clearly measure success. We recommend using specific measures such as:

- resources sharing utilization rate,
- government branch efficiencies and operation efficiencies,
- number and type of users,
- time spent on data resources or testbeds, and
- measures of impact, both social (e.g., developed and then actually used) and academic (e.g., number of citations and influence, additional research, and other foundational work)

While the document does address strategic objectives, there is no clear strategic planning to achieve these objectives. In general, SMART (Specific, Measurable, Achievable, Relevant, and Time-bound) goals are commonly used constraints for strategic planning. These goals would clearly explain and define the strategic objective, associated measures of success, probability of success, importance of the goal being sought (the document does a great job of this), and the estimated time for completion. A source of confusion was the lack of semantic consistency when discussing the strategic objectives. Recommendations should be clearly linked to each strategic objective. SMART goals could be distributed across several sections, but all of them should be present and addressed thoroughly.

c. Resource elements and capabilities. (Chapter 4)

IEEE-USA suggests establishing a national AI resources clearinghouse as a central technical network. The National Science Foundation Network (NSFnet) could serve as a model, one where anyone who is interested can access data as well as provide information about what is available. NSFnet, which linked major universities' resources and made them available across the country, was initially restricted to government and academia. Ultimately, NSF opened access to anyone interested in and working with research or learning about computer networking in the US. The result was rapid growth of internet service providers. Establishing a clearinghouse could allow users to learn of others undertaking specific research and training utilizing NAIRR resources, with the intent of facilitating connections among disparate users in the NAIRR community.

Data

Recommendation 4-1

Recognizing that it would be impossible for the NAIRR to curate a central repository of resources, IEEE-USA suggests the design and implementation of specific metrics for quality that could be referenced by users when contributing to the central repository.

Key to implementing this recommendation is the need for a governance structure to vet candidate providers of both data and compute resources, coordinate contributions and their utilization as well as evaluate over time (e.g., annually) with suitable metrics regarding the use and scale of data sets and the computing plus network connectivity performance of contributors and any associated NAIRR resources.

Recommendation 4-5

This will be an implementation challenge. Perhaps a tiered approach may work. For example, Tier 1 - quality data suppliers exceeding X gigabytes/month or terabytes/month may access compute resources at 80 percent discount from list price services. Quality data implies that it is 'clean'.⁶ Tier 2 - quality data suppliers that supply clean data streams below the Tier 1 threshold may access compute resources at a 40 percent discount. Tier 3 - users that do not supply data pay list prices for services, with an option to reduce cost based on the type of user (e.g., K-12, college, university, not-for-profit research organizations, and collaborative research consortia involving educational institutions).

Recommendation 4-7

It is unclear where the training programs will be located. We recommend centralization and implementation through a 'school' or 'training center' operated by NAIRR staff, who would be the focal point for ensuring a qualified user community. These training centers should issue certificates of qualification for user organizations as well as individuals, on a scale based on the sophistication of the applications the user community seeks. IEEE-USA recommends that the report clarify that training programs be made available both in-person and virtually to help ensure widespread accessibility.

Recommendation 4-8

IEEE-USA recommends defining "value ecosystem" as used in this instance.

Recommendation 4-9

⁶ https://en.wikipedia.org/wiki/Data_cleansing

Without a common framework, if data providers are permitted to set the security category of data, there could be many different thresholds. This may become a management challenge. We recommend that NAIRR set the standards that data providers must use to classify their data sets. This would help eliminate the potential for varying degrees of data security access for similar data from different providers.

Compute Resources

Recommendation 4-12

This recommendation requires considerable effort to organize and execute with multiple anticipated contractual relationships between NAIRR and compute resource suppliers in government, industry, and the private sectors. When implementing, we recommend looking at similar efforts to create a federated mix of resources such as the Digital Research Alliance of Canada, an organizational vehicle that facilitates access to a variety of independently operating HPC resources.⁷ The NSFnet was an early model of sharing discreetly operated HPC resources via a shared WAN network.

NAIRR might also consider establishing relationships with exoscale compute resources for advanced AI research in partnership with National Laboratories such as Los Alamos.

Recommendation 4-13

We recommend that the NAIRR specifically characterize the three levels, e.g., beginner, intermediate and advanced, if this is what is being described.

Recommendation 4-16

We recommend that NAIRR specify the architecture of edge computing resources and establish working relationships with edge computing resource suppliers to ensure continuity of resource availability over time.

Testbeds

Recommendation 4-19

IEEE-USA recommends requiring, not simply asking “when possible,” that this function be charged to the NAIRR staff who will be responsible for the implementation, management, and evolution of the testbed environment NAIRR chooses to utilize. This may guarantee a complete and widely accessible catalog, thus providing the needed consistency and avoiding duplication of efforts.

e. Privacy, civil rights, and civil liberties requirements. (Chapter 6)

Advanced analytics and artificial intelligence are powerful technologies that, along with their clear societal benefits, create new threats to privacy, equality, fairness, and transparency. Existing law does not yet protect sufficiently against these threats. The NAIRR does not mention ethics and appears to conflate privacy and security. IEEE-USA recommends adding clarifying language and reflecting on how NAIRR can manage the ethical use of these resources. The NAIRR could look to the recent work of both the Administration and Congress to ensure that AI systems adhere to accepted democratic standards of protections.

⁷ <https://alliancecan.ca/en/services/advanced-research-computing>

Additionally, NAIRR could refer to the work of scholars and practitioners from law, engineering, sociology, and statistics communities. For example, in a 2021 Ohio State University legal studies research paper, the authors interviewed corporate privacy managers, lawyers, and consultants, and surveyed a wide range of privacy managers to answer fundamental ethics questions about business data ethics management.⁸

Users of AI resources and tools must ensure that outcomes do not result in disparate treatment, disparate impacts, or other algorithmic harms and violations of democratic principles, and consequently undermine public confidence in and acceptance of AI. When AI systems are developed and deployed, objectives of accuracy and lack of algorithmic and other biases towards different groups can conflict. To mitigate these issues, it is imperative to ensure that researchers and students have access to established metrics and standards that will enable their operators to comply with standards that will enable their operators to comply with applicable legal and other standards for fairness, privacy, safety, and security. Transparency mechanisms for stakeholders that require third-party access to data in standardized, machine-readable formats are also needed.

Recommendation 6-1

NAIRR has noted that the protection of individuals' privacy, civil rights, and civil liberties are paramount to creating trustworthy AI and, therefore, the innovation, economic, and societal benefits that AI technologies and uses can deliver. Given the paramount importance of these protections under law, as well as under ethical systems, and given the devastatingly abundant evidence that government and private sector use of AI systems is falling short of the mark, Recommendation 6-1 should be strengthened to provide more substantive protections than are currently expressed in the accompanying text to take "efforts to ensure" transparency and "appropriate oversight" for NAIRR's operations, research, and governance.

Reaching beyond Recommendation 6-1, attorneys knowledgeable in privacy, civil rights, civil liberties, and consumer protection within the contexts of AI, specifically, and information technology, generally, should be integrated within NAIRR's operations, research, and governance teams.

IEEE-USA recommends incorporating references to laws and legal compliance. Specifically, transparent compliance with the Constitutional requirements for equality under the Fourteenth Amendment, the Civil Rights Act of 1964, and other civil rights laws should be specifically addressed and demonstrated. In addition, transparent compliance with the Freedom of Information Act (FOIA) should be highlighted, and, where exceptions under FOIA may apply, NAIRR should invoke and apply those exceptions only to the extent legally necessary to maximize transparency.

As to privacy, civil rights, and civil liberty protections within user agreements, the inclusion of terms is customary and generally assigned as an obligation to comply with the law. Mere agreements to comply, however, are inadequate. The law is not sufficiently evolved in its interpretation and application within AI contexts to make it clear to signatories how compliance is achieved. Therefore, we recommend that such user agreement terms provide specific guidance and include specific audit, reporting, and enforcement provisions.

Furthermore, user agreements should bar signatories from asserting or attempting to assert, including by registration or recordation, any property or other proprietary rights in resources made available by NAIRR thereunder. This bar should encompass a bar on attempts or assertions by those signatories to incorporate

⁸ Business Data Ethics: Emerging Trends in the Governance of Advanced Analytics and AI, *Ohio State Legal Studies Research Paper No. 628*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3828239. Accessed 27 June 2022.

those resources into their intellectual property assets, trade secrets, confidential information, or proprietary information, including as they may subsequently attempt to broadly define those terms.

Lastly, user agreements should incorporate compliance, audit, reporting, and enforcement provisions to protect data about individual human beings, irrespective of whether such data have been de-identified. Such privacy terms should not rely upon the now-outdated “reasonably linkable” standard articulated by the U.S. Federal Trade Commission in its 2012 Report to Congress.⁹

Recommendation 6-2

IEEE-USA recommends that ethics reviews under the recommended ethics review process be made iteratively over successive periods, for example, annually. Complete reports, findings, and corrective actions should be timely prepared and made publicly available.

In addition to an ethics review process, the NAIRR should establish a legal compliance review process. This legal review process should be carried out prior to resources being included within the system and then periodically thereafter. The legal review process should operate in tandem with the ethics review process. As a best practice, the earlier in the resource development process that legal and ethics experts can be engaged and integrated into the development teams, the sooner and more comprehensively complications can be successfully addressed.

As to third-party resources that may be targeted to be made available via the NAIRR, a thorough legal review should be carried out to document the provenance of those data and ensure that third parties are legally entitled to make those resources available to and via the NAIRR without breach of contract, privacy, or copyright infringement, and other legal violations, including moral rights that may apply.

As to “higher risk data,” embedded privacy protections alone may be inadequate to protect the privacy, civil rights, and civil liberties of the human individuals about whom the data relate. If these embedded protections were defective, inadequate, outdated, or defeated by cyberattacks, the individuals’ privacy, civil rights, and liberties protections would be potentially irretrievably harmed and unrecoverable. Therefore, the NAIRR should strongly consider a multi-layered protective approach to incorporate zero-trust models and homomorphic encryption, for example, in addition to these embedded protections.

The NAIRR should incorporate detailed mechanisms to track access to and use of these higher risk data.

The NAIRR has stated in the text accompanying Recommendation 6-2 that it may use institutional review boards (IRBs) for its research. The protective principles underlying IRBs should be adopted and as deemed appropriate, incorporated within the NAIRR’s recommended ethics review process. Under no circumstances, however, should ethics and IRB reviews be substituted for the legal compliance reviews and review process suggested, *supra*.

In the text accompanying Recommendation 6-2, the NAIRR expresses that, to the extent feasible, the outcomes of research that it enables should be regularly and over the long-term vetted to ensure that the subject research is not causing or contributing to violations of privacy or civil rights or infringements of civil liberties. The auditing, reporting, and enforcement mechanisms discussed in the commentary provided, *supra*, as to Recommendation 6-1 and user agreement terms will enable this essential vetting by the NAIRR and render that vetting feasible. We strongly recommend including those mechanisms.

⁹ Federal Trade Commission, *Protecting Consumer Privacy in an Era of Rapid Change* <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>. Accessed 30 June 2022.

Any research, however, that is carried out, without such user agreements in place must also be covered by vetting and protective mechanisms. Impact assessments such as those proposed in the 2022 Algorithmic Accountability Act¹⁰ or modeled in the Ethics and Algorithm Toolkit from The Johns Hopkins University Center for Government Excellence, for example, may be useful for the NAIRR to consider in this regard or more broadly.¹¹

Recommendation 6-3

We support Recommendation 6-3. Training, however, should be iterated on a quarterly basis as a best practice. All training should include a non-trivial assessment to confirm that trainees have a demonstrable and full understanding of their obligations, including their legal, fiduciary, and other duties of care to protect individuals' privacy and civil rights and their civil liberties.

We also recommend considering the establishment of NAIRR certifications that go beyond the expected scope of training. More rigorous certifications could more effectively and expeditiously recover and build greater public and market trust in AI systems and uses.

Recommendation 6-4

The NAIRR should ensure the completeness, currency, easy searchability, and free and public availability of the full listing and description of each of the planned inventory of provided data sets and dated details as to the history and provenance of each. As to retired and deprecated datasets, those should be maintained, within the inventory and its listing, while being designated as inactive.

¹⁰ See Algorithmic Accountability Act of 2022, S. 3572, § 4(a)(11)(b), 117th Cong., 2nd Sess. (introduced Feb. 3, 2022). A companion bill was introduced in the House of Representatives. See Algorithmic Accountability Act of 2022, H.R. 3580, 117th Cong., 2nd Sess. (introduced Feb. 3, 2022).

¹¹ Ethics & Algorithms Toolkit: A risk management framework for governments (and other people too!), <https://ethicstoolkit.ai/>, Accessed 29 June 2022.