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

DISCLAIMER: Please note that the RFI public responses received and posted do not represent the views and/or opinions of the U.S. Government nor those of the National AI Research Resource Task Force, and/or any other Federal agencies and/or government entities. We bear no responsibility for the accuracy, legality, or content of all external links included in this document.
September 30, 2021

Ms. Wendy Wigen
NCO
2415 Eisenhower Avenue
Alexandria, VA 22314

Via email

Re: ITI Response to Request for Information on an Implementation Plan for a National Artificial Intelligence Research Resource (NAIRR)

The Information Technology Industry Council (ITI) appreciates the opportunity to respond to the White House Office of Science and Technology Policy’s Request for Information on an Implementation Plan for the National Artificial Intelligence Research Resource.

ITI represents the world’s leading information and communications technology (ICT) companies. We promote innovation worldwide, serving as the ICT industry’s premier advocate and thought leader in the United States and around the globe. ITI’s membership comprises leading innovative companies from all corners of the technology sector, including hardware, software, digital services, semiconductor, network equipment, and other internet and technology-enabled companies that rely on ICT to evolve their businesses. Artificial Intelligence (AI) is a priority technology area for many of our members, who develop and use AI systems to improve technology, facilitate business, and solve problems big and small. ITI and its member companies believe that effective government approaches to AI clear barriers to innovation, provide predictable and sustainable environments for business, protect public safety, and build public trust in the technology.

We recognize that AI is an active area of research that is constantly evolving and improving. To harness this growth, it is vital to both utilize AI’s potential benefits while monitoring its impacts carefully, and research and development (R&D) is a critical enabler of those possibilities. Indeed, our Global AI Policy Recommendations, released earlier this year, include an entire section devoted to facilitating innovation and investment in AI that emphasizes the critical role R&D must play in that effort. As such, we welcome the opportunity to provide input on the Implementation Plan that the NAIRR Task Force is developing to guide the NAIRR.

We offer our response to several specific questions set forth in the RFI below.
Responses to Specific Questions

1. What options should the Task Force consider for any of 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;

Metrics will play an important role in establishing and sustaining the NAIRR. We propose three distinct sets of metrics that we believe will be helpful in measuring success:

**Researcher usage.** We believe that focusing first on the development and rollout of the NAIRR will be most useful. Measuring the use of hybrid-cloud technologies, as well as consumption of high-value datasets, will help provide insight into how much the NAIRR is being utilized. Once the NAIRR has had the opportunity to develop, metrics could then shift to the impact of the resource on researcher productivity. For example, it may be useful to measure the number of institutions utilizing the resource, the increase in the scale of research (including the amounts of data and compute used), the time it takes to conduct and reproduce experiments, the number of research publications facilitated using NAIRR access, and shifts in the citation scores of researchers using the resource.

**Community development.** As the technical infrastructure for the resource is developed, the Task Force should simultaneously seek to develop a community of users whose consistent utilization make the NAIRR a mainstream, go-to resource. Metrics could include the development of tools for the community, the creation of new benchmarks or standards adopted by the community, and the number of experiments shared within the community.

**Wider impacts.** The Task Force should also seek to measure the wider impacts of the NAIRR, as it seeks to spur research by increasing access to compute power and high-quality data. Once the NAIRR has been established, it would be helpful to assess new technologies or services that have been developed as a result of using the resource. These metrics could include increases in productivity and automation, new products and startups created, and the integration of AI by large companies. This will lead to improvements in the U.S. economy, including to the workforce. The Task Force could use metrics like number of jobs created or shifted or skills acquired to deepen understanding of the wider impact. Metrics could also be developed to determine how the NAIRR has influenced international competitiveness.

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 ii. A governance structure for the Research Resource, including oversight and decision-making authorities;
We encourage the Task Force to take a federated, or shared, infrastructure approach to implement, deploy, and administer the NAIRR. This will allow for a diversity of compute resources while enabling continued innovation. Indeed, such an approach will enable the NAIRR to rapidly upgrade/integrate new capabilities, while also enabling researchers to participate in the procurement and deployment of the resource. We encourage the federal government, in standing up the NAIRR, to develop a standard set of interfaces for federation of the resource—this could include how providers make computing and data resources available, for example, by specifying hybrid cloud computing architectures, application programming interfaces, and standards for data and metadata representation.

Initially, we recommend that the NAIRR be governed as a Federally Funded Research and Development Center (FFRDC), which is a private sector entity contracted by the government to undertake research and sponsored by a specific government agency. As FFRDCs bring together stakeholders from government, industry, and academia, they are often able to provide perspective that other, more traditional government resources cannot. Because the NAIRR will be a complex, first-of-its-kind resource, we believe this will be a valuable governance structure. Importantly, the NAIRR should not only be sponsored by one agency, as is typical of most FFRDCs, but should instead be sponsored by multiple agencies, so as to ensure partnership with a breadth of entities that can both contribute to and benefit from the NAIRR. Beyond that, allowing multiple government agencies to sponsor the FFRDC will help drive coordination and facilitate data-sharing between agencies, which is a necessary prerequisite to standing up the NAIRR.

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

As referenced above, we believe the FFRDC model will be most appropriate for governing the NAIRR. However, to the extent possible, we encourage the Task Force to consider how to automate day-to-day governance of the NAIRR. Indeed, this will be essential to ensuring researchers can easily access and utilize the NAIRR, helping to democratize access. One way this may be possible is to utilize a software-defined governance framework, which could be configured/reconfigured as necessary.

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;

Data access/secure access control. Given the immense amounts of data that will be collated within the NAIRR, capabilities around data access are necessary. When devising guidelines for the NAIRR, the Task Force should keep in mind the FAIR Principles for data management, ensuring that data is findable, accessible, interoperable, and reusable. Indeed, in order to appropriately leverage and democratize the NAIRR, data needs to be both identified and consumed by humans. As such, the Task Force should ensure that there are directions set forth for how data housed in the NAIRR can be accessed. Because data brought into the NAIRR will
be diverse, it is also important that the Task Force consider how to make the data interoperable and also how to ensure replicability and portability. Without this, it will be difficult, if not impossible, for researchers to utilize the NAIRR.

**Compute resources/scalability.** Compute resources will also bring important capabilities to the NAIRR. Beyond housing data, the Task Force should consider how to facilitate access to computing resources. We recommend taking an approach based on shared infrastructure, in which the NAIRR leverages existing cloud infrastructure to enable more rapid deployment. Starting or building the computing resources from scratch would not be cost effective or efficient. A hybrid or multi-cloud approach would also allow for public cloud to be integrated into the NAIRR.

**Educational tools.** Educational tools are also necessary for ensuring the success of the NAIRR and democratizing access. The Task Force needs to develop educational materials that demonstrate how to utilize the NAIRR, including how AI models might be developed using the NAIRR in a responsible and ethical fashion. The Task Force should consider developing educational materials applicable to all levels of students, researchers, etc. to ensure the broadest reach. The Task Force should also consider how to set up mechanisms to share best practices around use of the NAIRR, bringing together users to share their experiences to facilitate broader uptake. This could take the form of an annual conference, or some other sort of symposium.

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

Data is fundamental to innovation in AI. Indeed, one of our key recommendations in our *Global AI Policy Recommendations* is for policymakers to consider how to increase access to government sources of publicly available data in machine-readable formats and across borders to enable access to a foundational building block of AI. While the federal government is already making vast quantities of data available through resources like data.gov, and both NOAA and NASA provide public access to petabytes of earth and space data, the United States does not have a national strategy on data access and data-sharing for public interest applications. It is not always clear who owns data or how much data belongs in the public space. Beyond that, because of inconsistent data management practices, data set size, compute costs, bandwidth limitations, and inconsistent use licenses, government data can be difficult and costly to use or combine with other datasets.

Additionally, there are challenges with making available large data sets that contain personally identifiable information. Although there are constructs like HIPAA which protect health data, similar controls do not exist for data collected in other settings, like faces and license plates captured by street-view mapping processes or video doorbell monitors. There may also be security challenges associated with the NAIRR, as a consolidated data resource may be a target...
for malicious actors seeking to access and use the data for nefarious purposes. These challenges limit the innovation economy and can hinder academic research.

In order to address these challenges, we encourage the NAIRR Task Force to consider how to facilitate the adoption of AI by encouraging data sharing with meaningful stakeholders and determining how to best make datasets available to the broader AI research community. One way in which the NAIRR could do this is by taking a “data-as-a-service” (DaaS) approach, which sets forth clear policies and practices related to data access, security, retention, and use, with specified processes for updating those policies over time as technologies and data evolve. Although this approach has historically been used to monetize datasets, using this model for research could help to enable access to more data for AI modeling and research, while ensuring that safeguards are already built in. This model allows data to be located anywhere and searched from anywhere, while simultaneously reflecting an access structure that is based on graduated security policies that reflect the risks associated with each dataset. Data is “fluid” but controlled. The NAIRR Task Force should define overarching policies for dataset inclusion in the NAIRR, with dataset owners maintaining control of cost, access, and retention of their data within the scope of those overarching policies.

Additionally, the NAIRR should create opportunities to collect and distribute data responsibly, allowing U.S. citizens to opt-in to data collection by providing meaningful consent through services they already use, such as: healthcare exchanges, tax payments, social security, and Medicare. Another way to achieve this goal is to broker more data-sharing agreements, similar to the UK Open Banking Initiative or the IBM Mastercard Partnership.

We also encourage the NAIRR Task Force to develop flexible guidelines that are based on security and access control standards and best practices. Indeed, flexibility will be key in ensuring that such guidelines remain up to date as both cyber and privacy risks evolve. We believe the NAIRR should take a risk-based approach to security and access control, where data that poses minimal risk is made widely available and easily accessible. In doing so, the NAIRR should also consider developing guidelines that will help to protect data, intellectual property, and computing resources commensurate with the level of risk associated with those resources.

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

We are encouraged that the NAIRR is considering privacy and civil liberties requirements in the context of responsible R&D, as we believe that maintaining appropriate privacy protections is imperative to fostering trust in AI technology.

Like all technologies, AI operates in an existing policy and regulatory framework, and accordingly, personal data and related privacy concerns must be taken into consideration. We face important questions around striking the right balance between various objectives in the responsible development of AI, such as ensuring accountability, which requires some level of visibility into an AI system, while also protecting privacy. Adding to the complexity of a dynamic international environment of laws that are not always in alignment, the US currently has a patchwork of privacy policies and regulations that could become more complex and
fragmented as additional states follow California’s lead in establishing state-level comprehensive consumer privacy laws (already, Virginia and Colorado have followed suit, and many other state legislatures are currently considering such bills). Conflicts across these laws could have a chilling effect on AI advancement, as well as other data-driven technologies. To maximize the use of AI, we need strong, globally accepted privacy standards to enable trust and interoperability, and to incentivize investment in research to develop new techniques for even stronger privacy and security guarantees. To achieve this, we recommend the development of a national privacy law in the United States, consistent with ITI’s Framework to Advance Interoperable Rules on Privacy.¹

In lieu of such a law, however, we encourage the NAIRR to play a leading role in evaluating and defining research data practices, AI use, and implementation guidelines that protect individuals’ privacy and ensure equity. One way in which the NAIRR could do this is to create a working group or advisory committee that includes academic and government researchers, computer scientists, industry representatives, non-profit organizations such as the Center for Democracy and Technology, and legal and ethics scholars with an understanding of modern computing technologies. Such a group of diverse, expert stakeholders could help the NAIRR to identify the privacy, security, civil rights, and civil liberties risks associated with the aggregation of vast amounts of data and the application and advancement of AI, and develop recommendations to address those risks. It could also lead and engage in public discussion around the risks that may be created by data and AI. Upon developing an understanding of what the specific risks might be, the NAIRR should seek to implement the recommendations of the working group, leveraging a similar combination of subject matter experts and stakeholders. Finally, the group should seek to develop and publish educational tools and guidelines for broad public understanding of the benefits of and risks inherent to data and advancing AI.

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

The government plays an important role as an essential source of funding for long-term, high-risk research initiatives, and we recommend investment in diverse fields of AI research including cyber-defense, data analytics, fraud detection, robotics, human augmentation, natural language processing, and visualization and perception technology. Advanced algorithms, specialized computing hardware, high-quality data, and, most importantly, skilled human expertise are essential to enabling machine learning and the success of AI. To remain the leader in AI R&D, the United States must continue to promote an entrepreneurial environment, research network, and openness to talent. One of the reasons the United States has succeeded in this space is that it has invested heavily in R&D. In 2019, for example, the United States led the world in total R&D expenditures, with combined public and private sector spending totaling $657 billion.² Maintaining American leadership in AI related R&D efforts will not only require continued government R&D investment, but in promoting scientific

¹ https://www.itic.org/public-policy/FINALFrameworktoAdvanceInteroperableRules%28FAIR%29onPrivacyFinal_NoWatermark.pdf
collaboration among like-minded nations. Additionally, maintaining leadership in AI R&D will require continued strong political support from Congress and the Executive Branch, as well as active participation from the private sector and society.

The NAIRR should be primarily funded using public sources, as the NAIRR is intended to be a public resource. We encourage the U.S. government, and the NAIRR in particular, to map out a clear framework including funding commitments and timelines. One promising example of a framework the Administration could look to as a model, and develop in partnership with industry, is Canada’s Pan-Canadian Artificial Intelligence Strategy, which is delivered through the Canadian Institute for Advanced Research (CIFAR). CIFAR received $125 million to launch the Strategy in 2017, and the Canadian government has approved $349 million (USD) to fund the Strategy over the next ten years. Through the Digital Europe and Horizon Europe programs, the European Commission plans to invest €1 billion per year in AI. It plans to seek additional investments from the private sector and Member States to reach an annual investment volume of €20 billion over the course of ten years. If the U.S. is to remain the global market leader in AI, it needs a strong government framework to integrate resources and set goals to enable AI to grow and prosper. Partnerships with the private sector have an important role to play as well – we encourage the NAIRR to look to such partnerships to supplement public funding, either in cash or in kind, for specific projects.

As a further means to supplement public funding, the NAIRR could consider offering access to data as a subscription-based service and access to compute resources based on a tiered pricing model. It could also consider rewarding researcher contributions to open resources (e.g., algorithms, research data sets) and providing technical consultants to optimize projects for computation. Importantly, though, the NAIRR should offer fee waivers or otherwise consider a structure that would take into account institution size, need, etc. so as not to inadvertently hinder access.

1. 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?

We believe that NAIRR should prioritize developing and maintaining a shared computing infrastructure, which requires compute resources. In order to do so, we encourage the NAIRR to leverage existing governmental, academic, and commercial cloud computing and data resources sourced from multiple vendors, which will enable the NAIRR to offer the latest computing resources to users. The NAIRR should develop a common abstraction layer that enables users to develop AI systems in the same way across all vendors and seeks to adopt open standards for both data and compute resources, allowing for seamless transition between vendors.

---

3 https://vectorinstitute.ai/2021/05/03/federal-government-renews-pan-canadian-ai-strategy/
Education is also a key capability deserving the NAIRR’s focus, given it is critical to the ethical and responsible deployment and use of AI. Developing cross-disciplinary educational materials that demonstrate how AI models can be developed using the NAIRR in a responsible and ethical manner would be a helpful first step. Such materials should be accessible to students at all levels (ranging from primary to graduate students and researchers) as well as other researchers.

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?

Government, and the NAIRR in particular, has an important role to play in advancing ethical and responsible R&D of AI. In our *Global AI Policy Recommendations*, we emphasize the importance of facilitating public trust and understanding of AI technology, and we believe that the NAIRR can help to do so. We believe that the NAIRR can play an important role in facilitating interactions between AI companies and the communities they impact, with a view to better aligning stakeholders. More specifically, the NAIRR should invest further in research that supports the responsible development of AI, including in areas that improve the accountability, safety, fairness, and privacy of AI systems.

Beyond that, as the Task Force considers the funding and budget for the NAIRR R&D, we recommend that the NAIRR devote funds specifically to translational work – or the work it takes to translate basic research into industry-relevant insights or applications -- as a separate budget item. Too often in industry-academic engagements, it is assumed that research speaks for itself, which can lead to overburdened researchers doing significant amounts of work to ensure that industrial partners benefit from research, or only focusing on applied research where more basic, fundamental research is called for.

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

There are important building blocks that the NAIRR can leverage. For example, as governments and industry actively consider how to approach governing AI, standards play a key role in forming a bridge between written rules and practical implementation. We believe that the NAIRR should seek to support industry engagement in rules-based, consensus-based international standards organizations that are developing AI standards, and should work with industry stakeholders to consider how to leverage those standards.

Governments should maintain technology neutral policies that limit mandatory implementation requirements (e.g., for public safety considerations) in favor of voluntary implementation and self-attestation. To the extent compliance requirements are established, they should adhere to international best practices of conformity assessment. ITI and the InterNational Committee for Information Technology Standards (INCITS) recently launched an initiative, “Standards as a Tool
for Achieving Public Policy and Regulatory Goals (SPUR),” and developed a list of commonly used and developing international standards for AI for U.S. policymaker to reference so they don’t have to start from scratch. For example, ISO/IEC 24028 and ISO/IEC 24027 provide information related to trustworthiness in AI systems including transparency, explainability and bias, which are great tools considering the many policy debates around AI. There are also various developing standards around the use of biometrics, focusing on facial recognition and fairness, such as ISO/IEC 22116 and 19795.

NIST has also been undertaking significant work to cultivate trustworthy AI systems, which we believe is a fundamental building block for the efforts the NAIRR Task Force will undertake. Indeed, the benchmarks, metrics, and standards that NIST is developing – including its AI Risk Management Framework -- should inform specific guidelines the NAIRR Task Force develops to govern the administration of the NAIRR itself. As such, we believe the NAIRR should seek to heavily leverage this work.

In response to question 2, we also highlighted the importance of leveraging existing cloud infrastructure to develop and maintain a shared infrastructure. We reiterate that much of this infrastructure already exists – there are high-performance computing clusters on several university campuses (there are HPCs at Yale University, Kettering University, Clark University, Columbia University, as a few examples), the private sector already has cloud computing infrastructure, and organizations like Folding@home aggregate compute resources. The Task Force should consider how it can leverage this existing infrastructure to administer the NAIRR. In particular, the NSF-funded CloudBank model, which provides individual researchers access to commercial clouds for NSF-funded research, could be especially relevant and applicable here.

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

Public-private partnerships should play a critical role in the NAIRR. The National AI Advisory Committee and Subcommittee on AI and Law Enforcement, as well as the many public comment opportunities that have been provided to stakeholders as the US government seeks to implement the National AI Initiative Act of 2020, have been positive steps toward forging better collaboration between industry, government, and academia on AI. Given the rapid development and adoption of AI technologies in the commercial space, the need for consistent dialogue between the government and the private sector to inform research priorities, from both technical and social impact perspectives, cannot be understated. We recommend a regular cadence of dialogues such as quarterly discussions between the public and private sector. We encourage the NAIRR Task Force to continue to identify ways to incorporate regular

---

5 https://www.incits.org/contentAsset/raw-data/7a702e19-e075-400c-92cb-88ad83fda0d1/reportFile/fb9b572d-49f2-4b4b-8074-a84f3edad942.pdf
6 https://www.incits.org/contentAsset/raw-data/688802a6-4aeb-4333-9782-0c8b855ba040/reportFile/bd433a99-972d-4af0-8c2f-cc712a82516d.pdf
7 https://www.nist.gov/artificial-intelligence
8 https://foldingathome.org/?lng=en-US
9 https://www.cloudbank.org/
private sector participation to ensure AI is advanced in a fashion that is broadly beneficial to all Americans.

Many emerging AI technologies are designed to perform a specific task, such as assisting human employees or making tasks easier. Our ability to adapt to rapid technological change is critical and we must continue to be prepared to address the implications of AI on the existing and future workforce. By leveraging public-private partnerships – especially between industry partners, academic institutions, and governments – we can expedite AI R&D, democratize access, prioritize diversity and inclusion, and prepare our workforce for the jobs of the future. We recommend that government tap into the commercial space and when appropriate, form more public-private partnerships to maximize the potential of AI. For example, the Japanese government set up its country’s biggest research center through the government-backed Riken Institute, which involved 20 companies and research entities, with the goal of developing applicable AI in the medical and financial fields within 10 years. Another example of a successful public-private arrangement is the COVID-19 High Performance Computing Consortium, which brings together public, private, and academic participants to facilitate access to high-performance computing power to conduct COVID-19 research. Similar consortia may be helpful to entertain as follow-ons or complements to the initial FFRDC-governed resource.

As the NAIRR matures, we encourage the NAIRR Task Force to look to the Information and Communications Technology Supply Chain Risk Management (ICT SCRM) Task Force as a potential exemplar. The ICT SCRM Task Force has been a highly successful public-private mechanism, which has developed products and tools to address some of the most pressing supply chain security challenges in the United States. As such, we encourage the Task Force to look to the structure of the ICT SCRM Task Force as a guiding example of how a public-private partnership on AI may be similarly arranged, with private sector and government co-chairs leading the group and considering what working groups might be necessary to address challenges that may emerge as the NAIRR grows.

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?

Facilitating Access to Datasets. As we mentioned in response to question 1E, we believe the success of many promising uses of AI will depend to a large extent on the availability of training data. However, the lack of a comprehensive data-sharing strategy in the United States may make it more difficult for the NAIRR to democratize access to AI R&D. Therefore, we recommend that the USG, with assistance from the NAIRR, develop a balanced framework for the responsible use of data. By leveraging large and diverse datasets and increased computing power and ingenuity, AI developers and other stakeholders are empowered to innovate across industries to find solutions that will meet the needs of individuals and society in unprecedented ways. AI driven medical diagnostics can alert doctors to early warning signs to treat patients more capably. Increasingly intelligent systems are capable of monitoring large volumes of

11 https://covid19-hpc-consortium.org/