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.
Attn: Wendy Wigen, NCO  
Office of Science and Technology Policy and the National Science Foundation  
2415 Eisenhower Avenue  
Alexandria, VA 22314  

Submitted electronically via email:

Dear Ms. Wigen and NAIRR Task Force Members:

On behalf of Oracle America, Inc., thank you for the opportunity to provide input on FR Doc. 2021-15660, “Request for Information on an Implementation Plan for a National Artificial Intelligence Research Resource” (RFI). Oracle is a global leader in data management solutions, information technologies and cloud computing, serving 430,000 companies across 175 countries, and supporting hundreds of research projects and millions of students in more than 100 countries through its outreach programs. From its beginning, Oracle has been committed to supporting and advancing research, investing $56 billion in research and development over the past decade. Oracle appreciates the importance of developing a shared research infrastructure to provide Artificial Intelligence (AI) researchers and students across scientific disciplines with access to computational resources, high-quality data, educational tools, and user support, and fully supports this effort. We believe if the NAIRR initiative succeeds, it could be for AI what DARPA’s sponsorship of Ethernet and Internet research was for the creation of the modern information age.

In the following pages, we will address the specific points of the RFI in more detail. At a high level, we believe the NAIRR Taskforce and the resulting resource should be grounded in five primary principles to achieve the outcomes set forth in the RFI and ensure the NAIRR resource is able to evolve as technology evolves:

1. A common, flexible, open software framework that supports easy exchange of and collaboration around data, AI model, and compute resources (“run anywhere”);
2. Open standards for data and AI models;
3. Open standards for fairness, privacy and security in AI;
4. Open, sustainable architecture that avoids dominance by any single vendor or entity; and
5. Support for a diverse range of students, educators and researchers.

**Question 1** - Oracle offers the following feedback on goals E-I:

**Regarding Goal E** - The dissemination and use of government data sets and security and access controls are fundamentally connected issues related to accessing, sharing and managing data in a shared infrastructure model. We believe NAIRR can and should play a critical role in disseminating data, providing infrastructure to facilitate use of that data, and protecting the rights and privacy of individuals that could be violated by the availability of data collected by the government and by corporate and academic entities.
Although the Federal government currently publishes vast quantities of data via organizations like NOAA and NASA, and resources like www.data.gov, inconsistent data management practices and structures, data set size, compute costs, network bandwidth limitations, and inconsistent use licenses can make this data difficult and costly to understand, utilize, and combine with other data sets. Data sets that contain personally identifiable information (PII) present additional challenges: while HIPPA protects health data to some degree, similar controls are not available to protect the privacy of individuals in other large data collections, such as faces and license plates captured by street view mapping processes or video doorbell monitors. When researchers combine public and/or anonymized data sets with data sets containing PII, as for medical clinical trials, further use and privacy complications ensue.

We recommend a “data-as-a-service” (DaaS) approach from the outset that sets forth clear policies and practices related to access, security, metadata, search, retention, and use, with specified processes for updating these policies and practices over time as technologies and data evolve. While to date, DaaS mostly has been used to monetize data sets for business reasons, we propose something new – Data as a Service for Research – or DaaSR – implemented to enable wider use of more data for AI modeling and research, consistently and securely, and with built in safeguards to manage data ownership, data provenance, data security, and data governance.

A DaaSR approach to NAIRR could provide the following benefits for data set owners and users:

- Avoidance of vendor lock-in
- Standard metadata for data content, security, search and lifecycle
- Ability to support data in a wide range of formats
- Simplified search across data sets to enable more researchers to find and use the data they need
- Consistent data management and practices to enable combination of data sets and compliance with retention and reproducibility requirements
- Separation of the analysis/presentation layer from the actual data itself (this is the most open approach)
- Ease of administration, collaboration and audit

To ensure that NAIRR meets its goals of being generally usable and widely accessible, in establishing a DaaSR approach, we recommend that NAIRR:

- engage with standards organizations to define open standards, a data taxonomy, metadata, and best practices for data to ensure interoperability, using existing standards, modalities, and taxonomies where possible,
- define and apply consistent and appropriately differentiated data licenses across all government-collected data,
- establish privacy and security standards and security policy profiles to govern data access and use in ways that balance individual privacy protections with advancing research interests,
• ensure data is encrypted at rest as well as in transit to prevent any unintentional breaches, while including the encryption method and routines in the metadata, and
• establish clear and consistent rules for data retention and maintenance, and preservation and maintenance of the workflows and tools that enable use of that data.

**Regarding Goal F -** Over time, more data will be made available, in differing formats (text, graph, video, audio, etc.), and more privacy concerns and protections will be introduced. In short: the challenges with publishing, using and protecting data are only going to grow. The existence of a consolidated research data resource like the NAIRR will amplify the already material risks that bad actors may try to access and use data for nefarious purposes, and the expanded application of AI contributes to a growing risk that individual privacy and autonomy may be compromised, even accidentally, by collation of apparently innocuous, anonymized, individual data points.

To ensure the successful implementation, security and sustainability of the NAIRR, these challenges must be addressed strategically. Using a DaaSR model, access to data sets is enabled according to graduated security policies that reflect the risks associated with each data set, while data can be located anywhere and searched from anywhere. In such a model, data is “fluid” but controlled, enabling data owners to control cost, access to and retention of their data sets within an agreed set of guidelines and policies defined by NAIRR. These guidelines and policies should include a defined data catalog structure and consistent metadata to tag and define data sets. This overarching policy and practice infrastructure can and should evolve over time, will improve data searches and access to useful and allowed data, and enables data owners to easily expand or restrict data access as needed.

To protect and preserve both data and infrastructure, we recommend that the NAIRR develop guidelines that express the best modern security practices and access control principles, grounded in open standards and continually updated to reflect both improving technical standards and emerging risks in privacy and cyber-security. We suggest that the NAIRR adopt a risk-based layered approach to security and access control, where minimal risk data and resources are made highly accessible and widely available (e.g., to secondary school students) with minimal complexity for the user. In other words, the NAIRR should develop and adopt procedures to protect data, intellectual property and computing resources that are proportionate to the value and risk associated with those resources. A deliberate and integrated DaaSR approach would enable strong, consistent, and flexible data and security policies, practices, and implementations that can evolve as technologies and AI evolve.

A comprehensive security design strategy provides security assurance through identity management – the process of authenticating and authorizing security principals. Identity management services should be used to authenticate and grant permission to users, partners, customers, applications, services, and other entities. A successful design strategy will also endeavor to classify, protect, and monitor sensitive data assets using access control,
encryption, and logging in the NAIRR data repository system, and will place controls on data at rest and in transit.

The evolution of standards for APIs, WebServices, metadata and so on make it possible to reduce dependencies on features of the specific platform that the data resides on. As data from different sources and different countries are contributed, each of those data sets may or may not come with different use and security restrictions, such as the European Union GDPR requirements. Because it will be impossible to identify all possible restrictions and requirements from the outset, establishing core requirements up front and building an adaptable system that can evolve with requirements over time is critical.

In addition, the NAIRR should play a leadership role in setting security and privacy standards in both academia and industry and provide appropriate technology and educational material to help researchers meet these standards. The NAIRR also should develop and deliver other educational materials regarding the use of NAIRR infrastructure and data sets, working with academic institutions and educational organizations to develop curricular materials that ensure students and researchers at all levels down to secondary schools can understand, access, use and analyze data.

Regarding Goal G - We believe the NAIRR can and should play a critical role in evaluating and defining research data practices and AI use and implementation guidelines to ensure equity and preservation of individual privacy and civil rights and civil liberties. Just because we can do something with technology does not mean we should do it; appropriate guidelines and guardrails must be in place to avert a “tragedy of the commons” when it comes to data and AI.

Accordingly, in addition to developing policies and practices for the work the NAIRR sponsors, we recommend that the Task Force investigate the broader social implications of AI research and technology beyond that work. We expect that the Task Force’s work will encompass ensuring equitable access to compute and data resources for researchers and educators at a wide range of institutions while also addressing the privacy and civil liberties implications of the use of data and the research enabled by that data. The Task Force should consider issues such as “rights to be forgotten” and challenges raised by differences in international law concerning data and AI systems and access to data and systems by researchers within the country of origin and researchers outside the country of origin.

Specifically, we recommend that the NAIRR should (1) form a taskforce that includes academic and government researchers, computer scientists, industry representatives, non-profit organizations, legal and ethics scholars who also have an understanding of modern computing technologies to identify the privacy, security, civil rights and civil liberties risks associated with the aggregation of massive amounts of data and the application and advancement of AI; (2) lead and engage public discussion of the risks created by data and AI; (3) form specialized taskforces comprised of academic and government researchers, industry representatives and computer scientists to implement the recommendations of the first taskforce, including developing technical and organizational standards for privacy and
security in AI systems, potentially including standards akin to Human Subjects protocols for appropriately high-risk situations, and (4) develop and publish educational tools and guidelines for broad public understanding of the benefits of and risks inherent to data and advancing AI.

The work to evaluate and maintain appropriate guidelines and policies will be an ongoing effort because changes in technology and the legal frameworks associated with data and AI will require regular appraisal of data sets and the AI systems that exploit them. While we do not believe the NAIRR or other branches of the government should be in the business of defining “representative data sets,” we do believe the NAIRR has a critical role to play in ensuring that data sets are appropriately and transparently curated and identified by metadata to enable researchers to evaluate the appropriateness of data sets for their specific projects and goals. Similarly, the NAIRR is well situated to devise standards for the implementation of AI to ensure responsibility, transparency, and that AI innovation is serving the public good and consistent with American goals, values and laws.

Regarding Goal H - given the importance of the NAIRR to maintaining the long-term global leadership of the United States in research, the long-term sustainability of the NAIRR as a public resource is essential. The increasing centrality of computational methods in research across disciplines and the potential solution to compute access that NAIRR represents will naturally attract public and private partnerships and funding. However, private entities may represent competing interests, and partnerships necessarily introduce complexity. Given that the NAIRR is intended as a public research resource, the foundational, sustaining funding should come from public sources. Private-public partnerships (PPPs) and private investment, both in cash and in kind, can provide solid supplemental funding aligned to specific and project-driven interests.

Because compute can be expensive to provide, and computationally-intensive research interests must be balanced with environmental costs of delivering compute, we encourage the Taskforce to consider tiered and subscription-based models for NAIRR use, along with rewarding researcher contributions to open resources (e.g., algorithms and research data sets) and providing technical assistance to optimize projects to use the NAIRR facilities most effectively. For example, we have found that researchers and start-ups need “white glove” infrastructure support to train large “Foundation” models on the Oracle Cloud.

In terms of sustainability, a public private partnership funding model, with the government providing the initial funding for the establishment and governance of the NAIRR, and the private and educational sectors contributing to the capital and operational costs, would ensure the long term viability of the NAIRR. Funding organizations such as the NSF and NIH could preference the use of NAIRR assets, creating a continual funding stream.

Regarding Goal I – the NAIRR should consider developing a heterogeneous infrastructure to support AI education and research from primary school levels through to advanced, university-based academic, government and commercial researchers. This involves provisioning diverse compute resources and data resources appropriate to the communities
that the NAIRR supports, as well as the infrastructure required to meet cybersecurity and privacy standards that the NAIRR requires. Rather than develop and maintain its own infrastructure, we recommend that NAIRR utilize government, academic and commercial cloud computing and data resources from multiple vendors. This will enable the NAIRR to offer the latest computing resources to users. We also recommend that the NAIRR develop a common abstraction layer that enables users to develop AI systems in the same way across all vendors (“write once, run anywhere”). The NAIRR should adopt open standards for both data and compute resources to enable frictionless mobility from one vendor to another. As outlined above, this likely will require the NAIRR to take a leadership role in developing or extending such standards.

Developing an educational framework for AI is critical to both advancing AI innovation and ensuring responsible and ethical AI and data curation and use. The NAIRR should develop educational materials for primary and secondary students, undergraduate and graduate level students and researchers, and professional governmental, commercial and academic researchers. Importantly, these materials must be cross-disciplinary and include both the technical aspects of AI and the ethical, social and organizational issues involved in the development and deployment of AI. These materials can include a set of “best practices” studies demonstrating how AI models can be developed using NAIRR in an ethical, socially-aware fashion. The educational organizations of industry partners, for example, Oracle Academy and Oracle University, as well as other private sector educational resources, and organizations like the Computer Science Teachers Association (CSTA), CS4All, AI4All, AAAI, ACM and IEEE can and should be important partners in developing educational resources. Of note, the ACM Education Council has been working for a number of years on curricular guidelines at the intersection of computing, data science, and machine learning (a form of AI).

Regarding questions 2 and 3, we believe the NAIRR presents an opportunity both to effect a significant increase in the quantity and quality of scientific research and to expand the number, diversity, and perspectives of researchers engaging in the active pursuit of new knowledge and discovery. Therefore, we propose a tripartite organization of the NAIRR to focus specifically on and support the following three distinct activities, to ensure continued US leadership in AI research and technologies, and address equity and diversity concerns in education about and access to AI technologies and related opportunities across all of society:

1. Provide computing and data resources to support academic research into novel AI technologies, novel applications of AI to other academic disciplines, and the effects and implications of AI and related activities, e.g., data collection, on society. This will involve providing very substantial resources at the scale required to train “Foundation” models from scratch, as well as the more modest resources required to apply such models to a diverse range of applications, including basic science, healthcare, finance, etc. Support for this basic and applied AI research is essential for maintaining international competitiveness across a wide range of industries.

2. Provide computing and data resources to support undergraduate educational activities at universities and 2- and 4-year colleges, integrated with appropriate educational materials
dealing with AI and related technologies, e.g., cybersecurity, privacy, etc., and the social and ethical implications of these technologies. Modern AI models require specialized hardware (e.g., modern GPU-enabled computers) not available at many undergraduate institutions, especially non-R1 universities and 2- and 4-year colleges. It is critical that we make education in AI technologies available to all students at all educational institutions, for equity and diversity reasons, as well as economic ones. US industry has a deep and growing need for graduates with expertise in all areas of AI; to meet this need and remain competitive internationally, we must support AI education at the undergraduate level.

3. Provide educational and teaching resources at the primary and secondary school levels, integrated into math and science curricula, for teaching and learning about AI and related technologies, including interactive learning platforms. There is a fundamental lack of understanding and trust in AI and related technologies throughout society today, and without a broader social understanding of AI technologies, we run the risk that large sections of society will not be able or willing to take full advantage of the benefits AI can offer. Appropriate primary and secondary school educational materials, integrated into science and/or math curricula together with related topics such as statistics, could help students and teachers understand AI technologies and their risks and benefits. An interactive learning platform, perhaps based on a web notebook, together with appropriate teacher professional development courses and teaching materials, will need to be a primary focus of this activity.

**Regarding question 4,** there are existing organizations and resources that can serve as initial building blocks for both the research and educational missions of the NAIRR.

The considerate compute resources required for state-of-the-art AI research are substantial. For example, training one “Foundation” model carries cloud computing costs of about $100K. Federal research funders, like the NSF and the NIH, provide grants in the same order of magnitude as these costs, so similar grant proposals and funding mechanisms could be used to support academic AI research. While there are multiple models for providing computing and other resources to academic researchers, we believe NSF’s CloudBank is the closest extant model to what we recommend for the NAIRR. Cloud computing infrastructure of the kind required to build advanced AI systems already exists within US industry, and this kind of infrastructure (rather than traditional supercomputers) is ideally suited to developing the next generation of AI systems. Access and authentication to the NAIRR resource for researchers, regardless of funding source, would be managed by the NAIRR.

The compute resources required for teaching purposes in higher education are much less, and could be managed by a block grant to each educational institution. The allocated resources would be managed by the institution, and student accounts and access to the NAIRR would be governed at the institutional level, in accordance with policies and processes established by the NAIRR.

At the primary and secondary education levels, the required resources need to be planned and allocated as an integrated part of the regular curriculum, made available via the federal Department of Education to the relevant state-level educational authorities, who then would manage allocation and access in accordance with policies and processes established by NAIRR.
Regarding question 5, public-private partnerships (PPP) will be essential to the long-term efficacy of the NAIRR. AI-related computational research workloads have become sufficiently large and complex that no single organization – government or commercial – will be able to adequately and efficiently service all researchers. For example, through Oracle for Research we receive requests for AI-related research projects requiring up to 1M high performance compute hours per year. Servicing these workloads is important for the United States to remain at the leading edge of AI research and innovation, and PPPs are a viable mechanism for achieving this. PPPs can engage in open standards development and implementation, support efficient networking infrastructure development and availability, and ensure the efficient and timely availability of computational resources for research and education.

In a PPP, multiple partners can combine their resources, infrastructure, and responsibilities, allowing the NAIRR to allocate resources more efficiently and thereby use some of these resources to provide compute to researchers from across the academic spectrum. Furthermore, by combining resources, infrastructure, and skills from members of the public, private, and educational sector, PPPs benefit from an economy of scale, leveraging multiple technologies.

At the same time, new models for PPPs must be developed; these models must take into consideration commercial, academic and public interests and strive for simplicity and efficiency in contracting and funding models. While the aforementioned NSF-funded CloudBank project may provide a starting point, it also exposes myriad challenges with creating a simple user interface and pricing model while accommodating different university requirements and commercial considerations and various relationships between various government organizations and private entities. Other existing PPP models to consider are those that have been designed to handle, protect, and share medical or personal data. Additionally, changes to government funding practices are needed to ensure that grant funding and related university overhead fees do not combine to create perverse incentives for researchers to elect less efficient computational infrastructure options.

Regarding question 6, as a new initiative, the NAIRR will be competing with existing public and private AI projects. The NAIRR will need to differentiate itself by providing large amounts of high quality compute and data resources to researchers at all levels of the academic spectrum.

The NAIRR must also position itself to address the dire shortage of compute resources among the educational institutions that train developers in AI skills. By focusing on this need, the NAIRR will help address the significant lack of opportunities to work with high performance computing among diverse communities traditionally underserved by the educational system.

The NAIRR must also broaden its services beyond the traditional AI research community. New kinds of AI technology (such as Deep Learning “Foundation” models) are significantly easier to use than earlier AI, and the NAIRR should make these user-friendly models available to non-computer scientists in fields throughout the natural sciences, social sciences, and the humanities.

Democratizing access to AI research, development and education requires the NAIRR to address both technical and organizational challenges. The NAIRR will need to ensure equitable access and control at the data level, the network/infrastructure level, and the application layer. It will also
need to overcome limitations related to identity and access management (IAM), privileged access management (PAM), and performance and bandwidth, as well as the ability to support the required diversity of access restrictions based on data and resource policies.

To balance the need to ensure wide and equal access with the need to protect privacy and data, Oracle recommends that the NAIRR adopt a “zero trust” approach to security and access control. Zero trust is an IT security approach that keeps sensitive data safe while staying compliant with privacy regulations. As the use of cloud services rapidly expands, there are new risks associated with compromised or stolen credentials of a privileged administrator or application and increased potential for data theft, and for cyber criminals to conduct cyber fraud, because effective security controls are often an afterthought. Zero trust makes it possible for organizations to regulate access to systems, networks, and data without giving up control. Therefore, the number of organizations that are moving to a zero trust security model (meaning trusting nothing) is growing, so that they can safeguard data with security controls that restrict access to the data according to a specific policy.

About Oracle
Our responses to the specific points of the RFI are informed by ongoing Oracle projects, some of which are described here to provide further context.

*Oracle’s Implementation of AI across the Enterprise*
Oracle has been using AI to solve hard industry problems for many years, infusing AI into some of the most widely used enterprise applications in the world. Oracle now makes its AI capabilities and platforms, including general-purpose deep learning “Foundation” models trained on the Oracle Cloud, available to developers and data scientists to use directly in their own applications. With this, Oracle’s AI Platform enables users to build truly intelligent systems rather than simply incorporating general purpose AI models, by embedding AI models in SaaS applications and by using the domain expertise, data and knowledge from user applications. Oracle’s AI Platform includes the following:

- **Data Science Service** – a managed cloud service that helps data science teams rapidly build and deploy machine learning models across every step of the data science lifecycle from data labeling to orchestrated pipelines for data science workflows
- **AI Services** – includes perceptual AI services for areas like computer vision, speech and natural language processing, and Decision Services for business scenarios like anomaly detection, forecasting and recommendations and built on top of Oracle’s Data Science Service to make it easy for developers to apply AI to their solutions without data science expertise
- **Business AI Services** – higher level, compound AI solutions to solve for specific business scenarios, such as the Oracle Digital Assistant conversational platforms or Document AI

*Oracle Research and Development in AI*
Machine learning (ML) is an important subset of AI, and the work of the [Machine Learning Research Group](#) (MLRG) at Oracle spans a wide range of topics in ML and natural language processing (NLP), including aspects of ML devoted to fairness and privacy and how they apply to the business problems Oracle’s product groups focus on; NLP systems for model-based sentiment
analysis, named entity recognition, entity linking, co-reference resolution and product attribute extraction; and standard image recognition models and trying to understand how those models might help in training image recognition models for cervical cancer detection. In addition, the MLRG is responsible for one of the most scalable topic modeling algorithms that can train a model on billions of documents in hours on a cluster of computers, and is currently working on approaches to building large-scale, multilingual contextual embedding “Foundation” models that are resistant to many types of errors.

Also developed by Oracle Labs, KeyBridge uses advanced machine learning techniques to detect complex, unknown, or zero-day activities of intruders in terabyte-scale, unstructured, or semi-structured data streams. Oracle is using KeyBridge to monitor activities against our most valuable assets in the cloud infrastructure, e.g., data exfiltration, reconnaissance, and brute-force identity attacks.

**Oracle’s Engagement with the Research and Startup Communities**

**Oracle for Research** works collaboratively with researchers worldwide to accelerate results and inform the development of tools and services for research. Researchers working with Oracle for Research have used AI and ML to accelerate drug discovery, support virtual reality molecular dynamics simulations, make advances in 3D pathology tools, and more. **Oracle for Startups** enables new companies to launch and grow on Oracle Cloud in a virtuous cycle of innovation. Oracle for Startups has worked with AI startups like DeepZen, GridMarkets, and Aleph Alpha. The Eleuther research consortium and start-ups such as Aleph Alpha are using the Oracle Cloud to build open-source versions of GPT-3 style “Foundation” models, choosing to build their open-source “Foundation” models on the Oracle Cloud because it meets their extreme infrastructure needs and provides the best price-performance trade-offs.

Oracle appreciates the opportunity to provide input on this important topic. We know how to collaboratively develop and share infrastructure that creates opportunities for advancement and global leadership and are eager to work with you and others to create an accessible infrastructure that enhances and accelerates the position of the United States as a world leader in AI, across industries and research disciplines. We look forward to working with the Task Force, OSTP, NSF and other agencies and your collaborators to develop a national AI infrastructure that provides broad access, education and expanding opportunities for research and innovation in AI and beyond.