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

RFI Response To: OSTP and NSF "Request for Information (RFI) on an Implementation Plan for a National Artificial Intelligence Research Resource" issued 23 July 2021 [1]

Date: Thursday, 30 September 2021
From[1]:   Representatives from the National Oceanic and Atmospheric Administration (NOAA) Artificial Intelligence Executive Committee (NAIEC) and the Center for Artificial Intelligence (NCAI)

Contact: Dr. Rob Redmon, NOAA/NESDIS NCAI Lead

Background
The contributors to this RFI response are representatives from the NOAA Artificial Intelligence Executive Committee (NAIEC) and the NOAA Center for Artificial Intelligence (NCAI) that is authorized via the National AI Initiative Act of 2020 (DIVISION E, SEC. 5001) [2]. Through becoming law on January 1, 2021, the initiative provides a coordinated authorization for the entire Federal government to accelerate Artificial Intelligence (AI) research and application for the Nation's economic prosperity and national security.

NOAA has developed several strategies to harness the potential of emerging technologies, including the NOAA AI strategy [3] and is unlocking the full potential of Cloud, Data, and AI/Machine Learning (ML) capabilities in support of its mission.

Response
(Italicized Times New Roman text in this section is taken verbatim from the RFI notice.)

1. "What options should the Task Force consider for any of roadmap elements A through I, and why?"

Response: All elements listed are important to the development of a high quality roadmap. We have provided suggestions for consideration to items:
- *"A. Goals and metrics for success,*
- *D. Access to curated data sets and educational tools,*
- *E. Dissemination and use of high-quality government data sets,*
- *G. Civil rights and civil liberties requirements,*
- *I. Agency roles and responsibilities."*

A. "Goals for establishment and sustainment of a National Artificial Intelligence Research Resource and metrics for success;"

Response: While intended for NOAA's mission areas (e.g. ocean observations and exploration, weather, climate, etc), the goals from NOAA's AI strategy could be mapped and leveraged broadly. Below are a few ideas for the NAIRR to consider against their use cases.

---

[1]Any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

- Goal 1: Establish an efficient organizational structure and processes to advance AI
  - Conduct gap analysis of agency, academic, and industry needs and readiness to support and leverage a NAIRR (e.g., with data sets and services, educational materials, and open source development tools),
  - Leverage existing OSTP to United States Government (USG) agency Working Groups to build cross agency coalitions to share data, expertise, and resources,
  - Encourage agency staff exchanges to the NAIRR to capture broad perspectives including from NOAA,
  - Success metric: Breadth of USG, Academia, and Industry engagement
- Goal 2: Advance AI research and innovation, and research to applications
  - Leverage existing agency Federal funding mechanisms to develop competitive grants, X-prizes, Cooperative Research and Development Agreements (CRADAs), Hackathons, Other Transactional Authorities (OTAs), etc.,
  - Collaborate with Industry "AI for Good" activities,
  - Create AI-ready data, Trustworthy, Explainable, Equitable, and computationally efficient AI/ML measures and standards to infuse as requirements for Federal funding opportunities,
  - Develop interoperable Federal data lakes—Cloud optimized repositories with structured AI-ready data and baseline open source development support—of information products with free upload/download (aka egress),
  - Success metrics:
    - (a) Diversity of stakeholders involved including opportunities for students and historically underrepresented groups;
    - (b) Diversity of novel research activities supported and their potential mappings to economically valuable and/or decision-making applications;
    - (c) Economic value across sectors from applications transitioned to an Operational, Decision Support Service or Commercial setting;
    - (d) Federal efficiencies realized from implementing improved workflows
- Goal 3: Strengthen and expand AI partnerships
  - Develop USG agency level partnerships
  - Create a USG, International, Academia, Industry Advisory Working Group,
  - Foster and enhance partnerships and synergies with enabling technologies and activities (e.g., scientific data stewardship, commercial Cloud and AI pipelines),
  - Encourage agency, academia and industry staff exchanges,
  - Develop guidance for the treatment of intellectual property rights (IPR) resulting from AI partnerships, with an aim towards ensuring open approaches to AI development,
- Goal 4: Promote AI proficiency
  - Encourage the emergence of domain and domain agnostic Communities of Practice,
  - Encourage agencies, academia, and industry to collaborate on curated libraries of "Learning Journey's",

---

[1]Any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

○ Support proficiency in AI/ML via co-productive and competitive hackathons, certificate and higher educational programs, leveraging partnerships in Goal 3.

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;"

Response: The following capabilities should be considered:
1. Extend the ai.gov web accessible portal with features including AI-enabled Application Programming Interfaces (APIs) for researchers to search for and discover curated AI-ready datasets, educational materials, workshops and other engagement opportunities,
2. Develop and manage a Cloud resource pool with AI ready development sandboxes that users can apply for time to use,
3. Develop interoperable data lakes of curated Federal datasets, with baseline open-source development tools leveraging open repositories (e.g. Git). These curated data should be AI-ready including machine actionable documentation, readily working software tools, and adhere to Findable Accessible, Interoperable, Reusable (FAIR) standards,
4. Develop platform to include interactive notebooks (e.g. Jupyter or similar) as well as an interactive forum for exchanging community ideas and resolving technical challenges (e.g. Slack or similar),
5. Develop a framework to assess the AI-readiness of government produced data sets.

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

Response: The NAIRR should facilitate the development of Trustworthy AI standards and metrics to identify and minimize bias, and infuse these standards as required elements in funding opportunities. The NAIRR should include a mechanism for capturing biases discovered in operational products, which could be used to spur additional evaluation and development.

I. "Parameters for the establishment and sustainment of the National Artificial Intelligence Research Resource, including agency roles and responsibilities."

Response: OSTP should consider developing new or leveraging existing OSTP to USG agency Working Groups with representation from each agency; agencies contribute to national AI-ready data standard, and Trustworthy AI metrics.

In addition, agencies will need to devote significant effort to evaluating and improving the AI-readiness of their open data assets, ideally coordinated through agency Chief Data Officers. Establishment of a National Research Resource must acknowledge and support this crucial data

[1]Any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

curation, and standards metrics development work.

2. "Which capabilities and services (see, for example, item D above) provided through the NAIRR should be prioritized?"

Response: We recommend consideration of all capabilities suggested in item D above with particular emphasis on the creation of AI-ready development sandboxes connected to Federal data lakes of curated AI-ready data and baseline open-source development support.

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?"

Response: Trustworthy and Explainable AI (XA) best practices should be developed, training promoted and assessments performed to identify and minimize bias throughout the Research to Operation/Application/Commercialization pipeline. In particular the NAIRR should support the development of diverse teams assuring focus on AI issues.

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

Response: The NAIRR could leverage and extend its prior efforts to develop AI readiness matrices and could capitalize and support existing agency level strategies and developing implementations such as NOAA's AI, Cloud and Data strategies, and new Center for AI (noaa.gov/ai), as well as the recently launched National Science Foundation's National AI Institutes Program [4].

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

Response: Participating agencies could develop cross-agency pollinating activities: (a) announcements of opportunity, (b) CRADAs, (c) OTAs (e.g. NOAA's OTA with Google), (d) competitive X-prizes, (e) co-productive and competitive hackathons, (f) and training events. One exemplar is the NASA Center of Excellence for Collaborative Innovation (CoECI), who collaborates with innovators across NASA and the Federal Government to generate ideas and solve important problems by working with global communities via the NASA Tournament Lab. NOAA has several examples of direct experience in working with CoECI and others to conduct international competitive X-prizes (e.g. https://www.fishnet.ai/), including the recently concluded Model the Magnetic Field (MagNet) (https://ngdc.noaa.gov/geomag/mag-net-challenge.html) challenge which received thousands of potential model solutions that converged towards the best achievable performance and resulted in several prize awards to the international data science community.

[1]Any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

Another possible opportunity for public-private partnerships is in data preparation and dissemination (e.g. NOAA's Big Data Program). Private data brokers have enormous capability to clean, harmonize, and distribute data. Perhaps there is an opportunity for those brokers to achieve a public relations win by showing that these capabilities can also be used for good. For example, imagine a private data broker agreeing to develop a data infrastructure to bring together climate data from various Federal agencies, clean and harmonize the data, transform it into cloud-optimized formats, establish open APIs, and release it to the public for free as an AI-ready climate database, alongside archival of the most critical of these improvements in a national archive alongside the source data.

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?"*

Response: One significant challenge along the path to democratized access to AI R&D resources may include data transfer costs from data provider repositories to the user's NAIRR processing environment. A mitigation step for this issue could be for AI R&D resources to include fully functional AI-ready sandboxes including AI-and Analysis-ready datasets and AI/ML processing tools & pipelines, as well as encouraging development using open source libraries and platforms.

Additionally, restrictive intellectual property rights (IPR) are a concern, and one mitigation step is to develop guidance for the treatment of intellectual property rights (IPR) resulting from AI partnerships, with an aim towards ensuring open approaches to AI development

Another significant challenge is today's lack of interoperability between Cloud provider architectures and data formats, inclusive of USG data products, and a mitigation step might be to include this need in the development of an AI-ready standard, along with encouraging the use of open source libraries and platforms.

References:
[1] Request for Information (RFI) on an Implementation Plan for a National Artificial Intelligence Research Resource, issued 23 July 2021, with due date extended to 1 October 2021.
[2] National AI Initiative Act of 2020 (DIVISION E, SEC. 5001) became law on January 1, 2021
[3] NOAA Artificial Intelligence Strategy (signed), Version 2020-09-17.
[4] New NSF AI Research Institutes to push forward the frontiers of artificial intelligence, 26 August 2020.

[1]Any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.