Federal Register Notice 86 FR 46278, https://www.federalregister.gov/documents/2021/08/18/2021-17737/request-for-information-rfi-on-an-implementation-plan-for-a-national-artificial-intelligence, October 1, 2021.

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

ANTHROP\C

September 30, 2021

Submitted by email

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

Anthropic welcomes the opportunity to provide feedback to the Office of Science and Technology Policy (OSTP) and the National Science Foundation (NSF) in response to a Request for Information (RFI) on the development of the National Artificial Intelligence Research Resource (NAIRR). Anthropic is an AI safety and research company working to build reliable, interpretable, and steerable artificial intelligence (AI) systems. We're an organization with backgrounds in research, engineering, and policy, and we approach AI development from a cross-disciplinary perspective.

We believe progress in AI benefits from a broad range of stakeholders understanding the technology and its effects, and participating actively in its development. In addition to democratizing AI resources for the research and academic community, the NAIRR provides an opportunity for the U.S. government to measure and monitor progress in AI research and development (R&D), by virtue of creating a shared infrastructure that many stakeholders can use. With detailed information about the sorts of computationally-intensive projects researchers want to do, the U.S government will be better able to identify future research directions and funding needs[1].

We appreciate the opportunity to share our perspective on how the Task Force may wish to implement the technical and organizational design of the NAIRR. At a high level, we propose the NAIRR:

- Serve as a centralized marketplace for access to compute resources, made available by U.S. cloud providers;
- Prototype a system to access easily usable compute facilitated by supercomputers from the Department of Energy;

[1] Whittlestone, J., & Clark, J. (2021). Why and How Governments Should Monitor AI Development. a*rXiv*. arxiv.org/pdf/2108.12427.pdf

- Provide funding for a variety of AI research projects through a two-track proposal review process: one for Principal Investigator(s) (PIs) from U.S. academic institutions, and another reserved for a handful of large-scale, computationally-intensive research projects;
- Support the diversification of AI research talent in the U.S. by allocating a portion of funds to PIs that have not previously received NAIRR funding for work in AI;
- Encourage a significant portion of research conducted in the NAIRR be contributed to open source platforms in the form of code and research published as preprints;
- Develop a set of success criteria for the NAIRR, in order to generate information about how useful it is, and provide a means to iterate on program and technical design choices as needed.

Summary

We must ensure that the academic community has sufficient access to the infrastructure and services required for research into AI, much of which has become computationally-intensive. To achieve this, we recommend that the Task Force consider a phased approach to the development of the NAIRR, prioritizing the most critical and readily-available components of a shared compute resource first, in order to enable researchers to be able to replicate experiments done at the frontier of resource-intensive AI technology development. As these foundational components are made available, the NAIRR can be enhanced in subsequent iterations with the addition of important, though less time-sensitive, services and capabilities.

If successful, the NAIRR and its accompanying programs will democratize access to AI R&D by making infrastructure technology available to a broader range of academic and research institutions.

Compute: While strengthening the AI ecosystem will require investments in a number of important areas (e.g. education and training, dataset availability, etc.), arguably none are more strategic or immediately pressing as access to compute resources. OpenAI notes that the amount of compute used to train some large AI systems has doubled every 3.4 months since 2012, compared to the previous two-year doubling period for AI systems[2]. This indicates that access to large-scale computational resources has been critical in the creation of various AI breakthroughs — a point also made by Richard Sutton, an AI professor and pioneer of reinforcement learning, in his essay 'The Bitter Lesson'[3].

In this essay, Sutton argues that "the biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin." He concludes that "one thing that should be learned from the bitter lesson is the great power of general purpose methods, of methods that continue to scale with increased computation even as the available computation becomes very great." This suggests that many AI

[2] Amodei, D., & Hernandez, D. (2018, May 16). *AI and Compute*. OpenAI. https://openai.com/blog/ai-and-compute/
[3] Sutton, R. (2019, March 13). *The Bitter Lesson.* http://www.incompleteideas.net/IncIdeas/BitterLesson.html

breakthroughs not only *benefit* from access to large-scale computational infrastructure, but the discovery of these breakthroughs may *require* access to increasingly large amounts of compute to be found. As this trend continues, the infrastructure required to develop these systems will become increasingly cost-prohibitive for all but a handful of entities, most of which are private corporations.

To ensure researchers can fully participate in AI R&D, the NAIRR must help close the gap in infrastructure access between academic and industrial research labs. For example, at Anthropic there are multiple GPUs per researcher, whereas when we talk to academic labs we tend to find the inverse: that there are sometimes multiple researchers per GPU. The Task Force may wish to conduct an analysis of compute requirements for large-scale AI projects in both academia and industry to quantify this gap and determine the amount of funding necessary to support AI R&D through the NAIRR. On the technical side, a prototype NAIRR could serve as a centralized marketplace to acquire compute resources, leveraging the best infrastructure from U.S. cloud providers and existing public sector resources such as supercomputers managed by the Department of Energy.

Research Funding: With funding made available through a two-track proposal review process, successful grantees could use funds in the form of marketplace credits to acquire the infrastructure services that best meet the needs of their particular research efforts. One track could look to existing review processes such as National Science Foundation grants[4] or Broad Agency Announcements (BAAs) from the Defense Advanced Research Projects Agency (DARPA)[5], and solicit proposals from PIs in a method that is rigorous, transparent, and competitive. A parallel track would be used to fund a handful of large-scale, computationally-intensive projects and would be subject to a more robust review process. To encourage an equitable distribution of funding, we recommend reserving a portion of funds each year for PIs that have not previously received NAIRR grants for AI research.

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

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

Prior to the development of the NAIRR, the Task Force should establish a set of metrics to evaluate at predefined intervals during the lifecycle of the NAIRR to assess the success of the initiative. Some preliminary metrics the NAIRR may consider evaluating include: the scale of AI projects academia is able to develop as a result of resources unlocked by the NAIRR, the usability of NAIRR resources, the number of PIs receiving a research grant for the first time, the

[4] National Science Foundation. *Merit Review.* https://www.nsf.gov/bfa/dias/policy/merit_review/
[5] Defense Advanced Research Projects Agency. *Office-wide Broad Agency Announcements*. https://www.darpa.mil/work-with-us/office-wide-broad-agency-announcements

geographic diversity of PIs receiving funding, the number of publications enabled by the NAIRR, and the number of open source contributions to NAIRR-enabling infrastructure.

One metric of success would be to determine whether or not the NAIRR can help enable academic researchers to build models that match the resource scale (amount of compute used in training, size of datasets trained on, and so on) of those built by industry — this is important for developing AI research capacity in academia that is able to replicate (and therefore, critique and analyze) the systems being developed in industry.

Adoption and use of the NAIRR will depend heavily on its usability, which should be regularly measured as an additional indicator of success. The Task Force may implement various usability metrics indicating how long it takes researchers to get up and running with an experiment after receiving funding. Usage patterns and researcher surveys might also reflect the usability of the NAIRR, leading to technical or program design changes in later implementations. We elaborate further on how Anthropic enhances usability for our own infrastructure in Section 2, below.

The NAIRR presents a unique opportunity to help diversify the range of actors contributing to AI R&D. Providing funding and technical tools to researchers builds on a long partnership between the federal government and academia in discovering the scientific breakthroughs that power the U.S. economy. It is critically important that the NAIRR provide opportunities to newer entrants in the AI research community in order to both diversify the current talent pool and create a broader foundation for the U.S. research community.

To do this, we recommend the NAIRR reserve a portion of each year's funding (e.g. 10% of total grant dollars) for PIs that have not previously received NAIRR grants for AI research. This means that, beginning in its second year of operation, the NAIRR would start setting aside 10% of total grant dollars for researchers who have previously not applied or had their NAIRR grant denied. Using data from the first year, the NAIRR could prioritize outreach efforts to PIs and academic institutions that did not submit a grant proposal but would be eligible for these reserved funds, and may wish to prioritize PIs from traditionally underrepresented demographics, in line with broader federal initiatives around fairness and equity.

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

In initial implementations of the NAIRR, we recommend the direction of research be led by a "bottom up" approach. By this we mean that PIs would submit proposals in areas of AI R&D that are aligned with their goals, capabilities, and expertise, rather than a government directed focus on a handful of research areas. This may encourage AI proposals from a more diverse range of study (e.g. linguistics, biology, security) and potential to develop currently unforeseeable scientific breakthroughs.

In addition to the proposal review committees, a central oversight body should be established to handle day-to-day management of the NAIRR and monitor the overall direction of AI research in the U.S. Because the NAIRR will facilitate many new AI research projects, it will be valuable to track the projects conducted in the NAIRR and to closely analyze their outcomes, especially projects that lead to subsequent beneficial social goods or which generate meaningful economic activity. This will help the government develop information to guide future funding and research efforts[6].

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

Consider a Phased, Iterative Approach to the Development of the NAIRR

The Task Force may wish to consider a phased and iterative approach to the build-out and implementation of the NAIRR. In order to keep pace with the rapid acceleration of AI research and development, the Task Force should prioritize the most critical and readily-available components of a national research cloud, and make them highly usable, for the first iteration. Once the first phase of the NAIRR has been rolled out, subsequent phases can focus on the development of complementary services and capabilities for the resource.

We recommend the Task Force leverage existing infrastructure options to get off the ground quickly, meanwhile providing the research community with highly advanced technology. To train large-scale AI models, researchers must have access to a sufficient number of GPUs, CPU clusters to generate large datasets, moderately large amounts of storage to store datasets, good networking bandwidth, software to schedule jobs on the infrastructure (e.g. Kubernetes), and tooling to diagnose both hardware and software bugs. Instead of expending valuable resources developing a new 'government cloud' to provide these services, the NAIRR should utilize a combination of public cloud service providers and supercomputers managed by the Department of Energy National Labs (further described in Section 4, below).

High Demand for Compute

To determine technical specifications for the NAIRR, the Task Force may wish to benchmark against those from the national computing platforms of other countries. Australia's high-performance computing (HPC) system in its National Computational Infrastructure contains a total of 155,000 CPU cores and 640 GPUs[7], while Canada's advanced research computing (ARC) platform provided approximately 233,000 CPU cores and 2,610 GPUs in 2021[8]. To give a sense of demand for these services, this past year Canada's ARC platform saw its highest number

[6] Whittlestone, J., & Clark, J. (2021). Why and How Governments Should Monitor AI Development. a*rXiv*. arxiv.org/pdf/2108.12427.pdf
[7] NCI Australia. *HPC Systems.* https://nci.org.au/our-systems/hpc-systems
[8] Compute Canada. (2021). *2021 Resource Allocations Competition Results*. https://www.computecanada.ca/research-portal/accessing-resources/resource-allocation-competitions/rac-2021-results/

of applications for technical resources in its nine year history, representing a 10% year-over-year increase in proposals from 2020. The ARC, however, was only able to award 40% of the total compute requested and 22% of the total GPUs requested, indicating a growing need from researchers that outstrips current supply[9].

Closing the compute gap between academia and industry will require ambitious technical (and financial) investment. Funding aside, a truly competitive NAIRR in the U.S. might provide 100,000 A100-equivalent GPUs and 1 million CPUs. These figures, as significant as they are, would represent an ideal research resource and one likely built well into the future. To help put these figures (and associated significant cost) in perspective, it may be useful to benchmark against leading U.S.-based digital infrastructure providers which collectively spent nearly $97 billion on capital expenditures in 2020[10].

The initial version of the NAIRR should be large enough to launch with an ability to allocate non-trivial computational and data processing resources to a handful of projects. Specifically, we imagine in its first year, the NAIRR should allocate something on the order of 30% of its computational capacity to a small number (10 or fewer) of computationally-intensive projects. Such an approach would allow researchers to 'stress test' the capacity of the NAIRR to facilitate large-scale experimentation, and to identify bugs in initial implementation, which could be fixed before adding further capabilities to the NAIRR. It would also differentiate the NAIRR projects from other forms of scientific research enabled by existing federal funding programs — the NAIRR would have a concrete incentive to do a set number of computationally significant experiments per year.

Prioritize Usability of Infrastructure and Research Interfaces

Access to sufficient infrastructure alone will not guarantee the success of the NAIRR; usability will also play a key role. For AI research, cloud computing systems and HPC systems aren't always easy to use 'out of the box'. Researchers may have to invest in building tooling to make them usable for specific scientific experiments. At Anthropic, we've generally found widely-available commercial clouds are easier to build software tooling on top of than HPC infrastructure. However, our investments have been non-trivial — we have multiple full-time engineers whose main job is making our infrastructure cluster stable and usable for our researchers. A successful NAIRR would need to do the same.

As a result of these tooling investments, researchers at Anthropic are able to run experiments that are both relatively easy to launch (as in it doesn't take an individual researcher much effort to

[9] Compute Canada. (2021). *2021 Resource Allocations Competition Results*. https://www.computecanada.ca/research-portal/accessing-resources/resource-allocation-competitions/rac-2021-results/

[10] Note that this CapEx figure includes spend on land, corporate offices, etc., in addition to cloud and data center infrastructure. Fitzgerald, C. (2021, February 5). *Follow the CAPEX: Cloud Table Stakes 2020 Retrospective.* Platformonomics. https://www.platformonomics.com/2021/02/follow-the-capex-cloud-table-stakes-2020-retrospective/#*

boot up a large-scale multi-machine job), as well as efficient (as in the tooling helps increase the efficiency with which we utilize our cloud computing resources). These improvements ultimately help increase our experimental throughput. Researchers participating in the NAIRR may similarly need to build custom tooling to improve infrastructure usability and experiment efficiency. Those that do so should be encouraged to open source those solutions, helping other participants spend more time on research and less time on engineering needs.

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?

The NAIRR provides a unique opportunity to promote responsible research practices, and to gather valuable data on AI safety characteristics from aggregated research findings. These findings may be used to develop best practices in AI R&D, as well as to inform future research areas.

The NAIRR can reinforce responsible AI development by encouraging participating researchers to assess and document the broader societal impacts of their work. The Task Force may look to existing models such as the Conference on Neural Information Processing Systems (NeurIPS), an annual machine learning and computational neuroscience conference, which recently asked researchers to include a 'broader impact' section in submitted research papers[11]. While the NeurIPS statements ask researchers to reflect on broader impacts at the conclusion of their research, Stanford University has developed an Ethics and Society Review board that requires researchers to consider potential negative impacts of their work as a prerequisite to receiving funding[12]. Implementing similar requirements for NAIRR-enabled research may encourage participants to thoughtfully consider not only the benefits of their work, but potential harms and societal consequences, as well.

In collaboration with the research community, the government might also use the NAIRR as an opportunity to develop and iterate on assessment mechanisms to monitor the robustness, fairness, or bias of AI systems developed via NAIRR infrastructure. While evaluations to assess bias in some AI systems already exist, collaboration in a controlled research environment is likely to reveal areas of policy interest that do not have well-established benchmarks, and to yield assessment techniques that the whole community can benefit from. The government could use findings from this research conducted in the NAIRR to prioritize investment in specific areas of measurement, which has been shown to unlock follow-on work in academia and other parts of the economy[13].

---

[11] Neural Information Processing Systems Conference. (2020, February 19). *Getting Started with NeurIPS 2020*. Medium. https://neuripsconf.medium.com/getting-started-with-neurips-2020-e350f9b39c28

[12] Bernstein, M. S., Levi, M., Magnus, D., Rajala, B., Satz, D., & Waeiss, C. (2021). ESR: Ethics and Society Review of Artificial Intelligence Research. *arXiv*. https://arxiv.org/abs/2106.11521

[13] Whittlestone, J., & Clark, J. (2021). Why and How Governments Should Monitor AI Development. a*rXiv*. arxiv.org/pdf/2108.12427.pdf

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

Industry Infrastructure: Cloud Providers

Rather than a government-led effort to build net new infrastructure solely for the purposes of the NAIRR, the entity ultimately responsible for its development should leverage the advanced capabilities of industry available today — specifically, the infrastructure that has been developed relating to rentable compute and data processing systems. Utilizing the infrastructure of industry cloud providers will accelerate implementation of the NAIRR, and the NAIRR can seek to use multiple infrastructures in parallel, which will encourage competition among industry vendors during the infrastructure bidding process.

The NAIRR management entity should establish an open procurement process to solicit bids from a range of providers supplying infrastructure services and establish a panel of technical experts to review submitted proposals. The procurement terms should be clear on required functionality and desired outcomes, but otherwise remain open to a range of providers. Widening the field of available infrastructure providers also enables greater choice for researchers, allowing them to pick the provider (or providers) that most effectively meets their needs. After implementing selected providers in the first iteration of the NAIRR, the Task Force may consider a recurring procedure to solicit new bids on a regular basis (e.g. every 3-5 years) informed by the trends and feedback of researchers participating in the NAIRR.

Public Sector Infrastructure: High Performance Computing (HPC) Assets

Along with using a system based on commercial cloud technology, the NAIRR might also utilize existing and future public compute assets — specifically, supercomputers operated by the Department of Energy. The Task Force should explore repurposing these HPC assets for training AI workloads via the NAIRR. This will generate valuable information about the cost tradeoff between industry cloud infrastructure and government-owned HPC infrastructure, while also creating evidence about the costs required to make such infrastructure usable by researchers (a key success criteria of the NAIRR).

It is unknown to us whether existing HPC systems exhibit the same usability as commercial cloud computing resources; we suspect that, for frontier AI research, they don't. By evaluating how easy researchers find using HPC platforms compared to cloud platforms, the NAIRR can generate information about where to make future investments. Concretely, if it became obvious that researchers were having difficulty launching large-scale experiments on HPC infrastructure, and these experiments took longer to launch than those on commercial clouds, then that could lead to the NAIRR prioritizing investments in increasing the usability of HPC assets.

Leverage Existing Research Review Processes, While Building New Ones

As with many of the elements of the NAIRR, the research selection process represents an opportunity to build on existing programs while iterating on new parallel structures. In determining how to review and allocate funding for research proposals, the Task Force may consider a dual-track approach where a significant portion of funds (e.g. 70%) are open to PIs from any U.S. academic institution, while the remaining funds (e.g. 30%) are reserved for a small number of computationally-intensive projects. The 10% allocation for new PIs could be used for either general or computationally-intensive research proposals.

The general project track might leverage existing processes such as the NSF grant review process or DARPA's Broad Agency Announcements, with review panels composed of AI experts representative of the academic community at-large. This type of thorough and transparent review process, combined with a reserved 10% allocation for new PIs, may help distribute funding to institutions in areas that don't typically receive large-scale grants. Brookings suggests that nearly 87 metropolitan areas could be potential centers for AI R&D, on top of the 21 metropolitan areas with prominent academic institutions already receiving significant federal funding for AI projects[14].

As mentioned, the NAIRR could also encourage researchers to think about bold, large-scale efforts and support them by funding a handful of computationally-intensive projects. Given the scale of these projects, PIs, academic institutions, or multiple academic institutions in collaboration would be eligible to apply. The review process for selecting these projects will likely involve an additional layer of deliberation relative to the general project grants, and could also incorporate priorities from the broader scientific community. For instance, the NAIRR could work with the NSF to identify a small number of computationally-intensive project areas, such as those that relate to AI and climate change, and then request project proposals in these areas.

In some cases, existing review processes may be too slow for promising research proposals from well-known institutions that are already leaders in AI R&D (e.g. the National Artificial Intelligence Research Institutes[15]). In order to accelerate their research and access to critical resources, the Task Force may consider developing a form of NAIRR access that involves minimal review for PIs who need only a modest amount of compute resources. This could further increase the impact of the NAIRR as a core piece of research infrastructure, though we believe the greatest value of the NAIRR will come from unlocking large-scale, computationally-intensive experiments.

[14] Muro, M., & Liu, S. (2021, September). *The geography of AI: Which Cities Will Drive the Artificial Intelligence Revolution?* Brookings. https://www.brookings.edu/wp-content/uploads/2021/08/AI-report_Full.pdf
[15] National Science Foundation. *National Artificial Intelligence (AI) Research Institutes*. https://beta.nsf.gov/funding/opportunities/national-artificial-intelligence-research-institutes

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?

Encourage Open Source Research Contributions

One failure mode of the NAIRR would be a scenario in which the U.S. government invests significant funding to democratize access to AI R&D, but the research made possible by these efforts is kept within a small cohort of actors or the discovering institution itself. To overcome this potential limitation, a significant portion of NAIRR projects should yield open source insights, either by way of software written to facilitate large-compute experimentation or published results in the form of research preprints. While we ultimately defer to the NAIRR Task Force to determine the optimal percentage of open source contributions, we recommend it be significant enough such that all participating researchers have ample work product to further assess and build upon.

Conclusion

We applaud the work of the Office of Science and Technology Policy, National Science Foundation, and the Task Force to develop a shared research ecosystem. Increasing academic access to the components of today's advanced AI technology will build on a long and successful collaboration between academia and the public sector in creating transformative technologies and advancements across the U.S. economy. We urge the Task Force to consider program design choices that leverage existing advanced infrastructure, distribute funding equitably and efficiently, and ensure the use of responsible research practices. We appreciate the opportunity to submit this response and would be delighted to answer any questions that may arise.