

March 4, 2022

Office of Science and Technology Policy
Submitted via email to AI-RFI@nitrd.gov

Re: RFI Response on the Update of the National Artificial Intelligence Research and Development Strategic Plan (FR Doc. 2022-02161)

Dear Dr. Alondra Nelson,

Data & Society Research Institute is pleased to submit a response to the Request for Information (RFI) published by the Office of Science and Technology Policy (OSTP) on updating the National Artificial Intelligence Research and Development Strategic Plan (“Strategic Plan”).

Our organization is an independent, nonprofit research institute studying the social implications of data-centric technologies and automation. We are working to produce empirical research that challenges the power asymmetries created and amplified by technology in society. We are pleased to see this commitment to updating the Strategic Plan. As we have shared in recent comments to OSTP¹, it is essential that we develop AI policy and governance mechanisms that are responsive to the prevalence of AI systems that enable discriminatory practices and that expose marginalized communities to harm.

Throughout our comment, we reiterate the importance of long-term research funding that challenges, rather than consolidates, corporate control over the AI research field. In order to meaningfully reckon with mounting evidence of the harmful impacts of large-scale AI systems, the Strategic Plan must introduce additional research directions for programs that facilitate meaningful democratic control of AI and related technologies. We also encourage reframing that centers equity and anti-discrimination as first principles and critical components of successful AI research.

¹"Democratize AI? How the proposed National AI Research Resource falls short." *AI Now and Data & Society*. (Oct. 5, 2021).

<https://medium.com/@AINowInstitute/democratize-ai-how-the-proposed-national-ai-research-resource-falls-short-96ae5f67ccfa>

1. Background and context: Understanding the “AI boom”

To understand the complex environment surrounding AI research, it’s helpful to recall that the current turn to AI is primarily a product of significantly concentrated corporate resources—namely vast computation, massive data, and the capital required to attract and retain scarce AI talent.² Placing human beings as active agents in the creation of AI tools and frameworks, as well as stakeholders with a vested interest in algorithmic outcomes is especially important before we can truly assess and understand any technological advances achieved thus far. **The so-called “advances” in AI that have been celebrated since the early 2010s were not due to breakthroughs in AI research and innovation. They were predicated on newly available access to powerful computation and to massive amounts of web data.** Then, as now, these are resources that a handful of powerful tech companies have in large supply thanks to ad tech-driven surveillance business models, that few others can avail themselves of without going through these companies first.³

The past decade’s Big Tech-led turn to AI profoundly shaped academic computer science disciplines as well. It served to redirect computer science research toward AI-related questions and approaches favored by these companies. In particular, the influx of money and attention produced a turn toward resource-intensive research and development. Work that could avail itself of expensive and scarce industry computing and data was heralded as “cutting edge.” This created an uneasy and conflicted environment for university AI research, in which the dependence on large tech company funding, infrastructure, and data was recognized by practitioners, but not often openly acknowledged.

Policy proposals that disrupt the dynamic of concentrated corporate power and the effects this has on academia are important and necessary. However, this also presents us with a set of thorny questions, at the center of which is: how do we validate the importance of federally funded R&D while also reducing the power and control of the handful of companies currently dominating AI research, and how do we ensure that determinations about whether—if at all—AI is developed and deployed are subject to more democratic deliberation. Directing attention and resources to

²While the field of Artificial Intelligence is ostensibly oriented around making machines intelligent, in practice, most AI systems rely on big data - the collection and processing of massive datasets, identifying patterns and probabilities within them and codifying them into a predictive mathematical model. See Broussard, Meredith. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge: MIT Press.

³Whittaker, Meredith. “The Steep Cost of Capture,” ACM Interactions, Vol. XXVIII.6 Nov-Dec 2021. Forthcoming.

these questions is essential for OSTP's efforts to ensure that data-driven technologies reflect and respect our democratic values.⁴

2. Large-scale government investment in shared computing and data infrastructure will entrench corporate control over the AI field, contrary to the Biden Administration's bold stance against the power of large tech companies in society.

The National AI Initiative Act calls for regular updates to the Strategic Plan, which may include activities that overlap with National AI Research Resource (NAIRR) aims, such as “[providing or facilitating] the necessary computing, networking, and data facilities for artificial intelligence research and development.”⁵ If these resources were to take the form of shared, national cloud infrastructure, as envisioned through the NAIRR proposal, we foresee the inability to successfully implement those plans without further entrenchment of corporate influence and control over the AI research field.⁶ What is being proposed through the NAIRR is an extension of industry-dependent resources, not the construction of resources that would challenge or reduce the centralized power of the large tech players.

There is no scenario in the short or mid-term future where large scale computational resources adequate to the task of expanding access to bigger-is-better AI research resources could be created and maintained by institutions meaningfully separate from the large tech platform companies. These companies provide more than raw computing power: the computational environments they own and license provide the tools and research environments that define how AI research gets done. Most policy discussion thus far, as has happened within recent media policy history, “involves technical terms that enable easy obfuscation.”⁷ There is no plausible path forward in which such a resource would not be dependent on existing tech industry platforms, tools, and resources.

⁴Lander, Eric, & Nelson, Alondra. "Americans Need a Bill of Rights for an AI-Powered World." *Wired*. (Oct. 8, 2021). <https://www.wired.com/story/opinion-bill-of-rights-artificial-intelligence/>

⁵Science and Technology Policy Office. Request for Information to the Update of the National Artificial Intelligence Research and Development Strategic Plan. *Federal Register*. (Feb. 2, 2022). <https://www.federalregister.gov/documents/2022/02/02/2022-02161/request-for-information-to-the-update-of-the-national-artificial-intelligence-research-and>

⁶Science and Technology Policy Office. Request for Information (RFI) on an Implementation Plan for a National Artificial Intelligence Research Resource. *Federal Register*. (July 23, 2021). <https://www.federalregister.gov/documents/2021/07/23/2021-15660/request-for-information-rfi-on-an-implementation-plan-for-a-national-artificial-intelligence>

⁷Crawford, Susan. (2013). *Captive Audience: The Telecom Industry and Monopoly Power in the New Gilded Age*, Yale University Press.

This is tacitly acknowledged in repeated calls to constitute the NAIRR via public-private partnerships.⁸ This arrives at a time when the concentrated power and influence these companies exert is increasingly under scrutiny, including by the Biden Administration itself, which has taken a clear stance that a small number of dominant platforms are using their power to extract monopoly profits.⁹ An effort that aims to “democratize” AI research by investing money in companies that dominate their market, will only further entrench these firms’ power and reach. This will make it harder to check the power of these companies through regulation and public pressure, and will potentially stifle new innovation, as a small amount of proprietary tech and resources will inherently create limits or boundaries on new research directions.

As we raised in our NAIRR comment, working to democratize access to cyberinfrastructure and to fuel AI research could instead take the form of investments that help us meaningfully ensure democratic control and deliberation over AI. In order to do this, the Strategic Plan could call for additional funding for under-resourced research domains, taking leadership from those most harmed by inequitable uses of AI, and investing in meaningful public control over these powerful technologies.¹⁰ This funding could be used to 1) create and fund scholarships for underrepresented students and sociotechnical research disciplines, 2) invest in fellowships that place sociotechnical scholars in federal agencies, and 3) create and maintain public engagement mechanisms that allow communities most harmed by these systems to have a say in AI R&D.

Public-private partnerships also form a central component of the Strategic Plan. While public-private partnerships can take many forms and are relied on in many parts of government, it’s important to view them cautiously in the context of federal AI R&D strategy, given the fact that many companies who may be eager to partner with the government on AI research are the same companies consolidating research power and resources in this field.

In order to better understand the tradeoffs of public-private partnerships in the AI research field, we advise OSTP to convene public workshops with civil society, academia and industry to develop a nuanced approach to public-private partnership that balances

⁸Stanford Institute for Human-Centered Artificial Intelligence. "National Research Cloud Call To Action" <https://hai.stanford.edu/national-research-cloud-joint-letter>; The National Security Commission on Artificial Intelligence. "NSCAI Submits First Quarter Recommendations to Congress." (Apr. 1, 2020). <https://www.nscai.gov/2020/04/01/nscai-submits-first-quarter-recommendations-to-congress-2/>

⁹Executive Office of the President. Executive Order on Promoting Competition in the American Economy. (Jul. 9, 2021). <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/07/09/executive-order-on-promoting-competiti-on-in-the-american-economy/>

¹⁰The National Science Foundation. "NSF partnerships expand National AI Research Institutes to 40 states." (Jul. 29, 2021). https://www.nsf.gov/news/news_summ.jsp?cntn_id=303176

the benefits of cross-sector collaboration with the risks of further consolidating power and resources within industry.

We also encourage additional historical analyses of analogous large-scale R&D investments from the federal government, in order to identify strategies for facilitating cross-sector collaboration, securing ethical and effective public-private partnership, and ensuring investment in technological development is rooted in maximizing public interest.¹¹

3. The Strategic Plan must begin to reckon with mounting evidence of the harmful impacts of large-scale AI systems, including discriminatory consequences for marginalized groups and long-term climate impact of this scale of computing.

We encourage additions to Strategy 3 that engage more deeply with the body of critical research, press coverage, investigative reporting, and public discussion that has generated significant evidence of AI's harms¹², and raised fundamental research questions¹³ about the ability of AI systems to operate safely and transparently in sensitive social domains.¹⁴

While proponents of rapidly expanding the use of AI often point to the potential for this technology to stimulate economic growth, many people—particularly marginalized communities—are already subject to the worst excesses, mistakes, and harms perpetuated by the oppressive and extractive use of powerful algorithmic technology.¹⁵ These harms are not

¹¹ For example a historical analysis of U.S. media policy and radio technology. See Victor Pickard, (2022). "The Great Reckoning: Lessons from the 1940s media policy battles". *Knight Columbia*. <https://knightcolumbia.org/content/the-great-reckoning>, and See Pickard, Victor. (2021). A New Social Contract for Platforms: Historical Lessons for the Digital Age, in *Dealing with Digital Dominance: Joining up the Policy Solutions* (Damian Tambini & Martin Moore eds).

¹² Sisson, Patrick. Housing discrimination goes high tech: How algorithms, ad targeting, and other new technologies threaten fair housing laws." *Curbed*. (Dec. 17, 2019). <https://archive.curbed.com/2019/12/17/21026311/mortgage-apartment-housing-algorithm-discrimination>; Kirchner, Lauren. "Access Denied: Faulty Automated Background Checks Freeze Out Renters." *The Markup*. (May 28, 2020). <https://themarkup.org/locked-out/2020/05/28/access-denied-faulty-automated-background-checks-freeze-out-renters>; Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 143(1), 30-56, <https://doi.org/10.1016/j.jfineco.2021.05.047>.

¹³ Birhane, A. (2021). Algorithmic injustice: a relational ethics approach. *Patterns*, 2(2), 100205.

¹⁴ Alkhatib, A. (2021, May). To live in their utopia: Why algorithmic systems create absurd outcomes. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-9).

¹⁵ Noble, Safiya Umoja. (2018). *Algorithms of oppression*. New York University Press. <https://nyupress.org/9781479837243/algorithms-of-oppression/>; Benjamin, Ruha. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*, Polity Press. <https://www.wiley.com/en-us/Race+After+Technology:+Abolitionist+Tools+for+the+New+Jim+Code-p-9781509526437#:~:text=Presenting%20the%20concept%20of%20the,ultimately%20doing%20quite%20the%20opposite>

abstract—these are tangible, often irreversible and irreparable harms that perpetuate inequality.¹⁶ **Any effort that focuses on accelerating economic growth while these harms continue relatively unchecked, raises critical questions about whether certain communities are considered expendable and what kinds of harms are allowable in our society in the name of economic growth.**

Furthermore, many subcategories of AI, especially those that emphasize the research and development of large-scale data and computing (often called Large Language Models or LLMs), are uniquely prone to perpetuating social harms and entrenching biases.¹⁷ These large-scale models, which are trained on troves of internet data from sources exhibit persistent discriminatory outputs.¹⁸ Large-scale AI models are built on mass surveillance, which disproportionately impacts marginalized communities,¹⁹ without implementing meaningful mechanisms for accountability or consent of the public.²⁰ **Their carbon cost is also substantial: the amount of processing required to train AI models is both financially and environmentally resource-intensive, and these costs are only likely to expand given industry standards that tie performance metrics to the size of the dataset used to train the model.²¹ As a whole, the tech industry is responsible for a global carbon footprint comparable to the aviation industry, and data centers make up 45% of this footprint.²²**

The Strategic Plan can begin to acknowledge these harms by being explicit about the wide range of people, organizations, and disciplines that also contribute to AI R&D. Major contributions to

¹⁶ Hill, Kashmir. Wrongfully Accused by an Algorithm. *The New York Times*. (Aug. 3, 2020). <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>.

¹⁷ Bender, Emily M., Gebru, Timnit, McMillan-Major, Angelina, & Shmitchell, Shmargaret. (2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). <https://dl.acm.org/doi/10.1145/3442188.3445922>

¹⁸ Abid, Abubakar, Farooqi, Maheen, and Zou, James. (2021). "Persistent anti-muslim bias in large language models." *In Proceedings of the 2021 AAI/ACM Conference on AI, Ethics, and Society*, pp. 298-306. <https://arxiv.org/pdf/2101.05783v1.pdf>

¹⁹ Georgetown Law Center on Privacy and Technology. The Color of Surveillance." Conference. (Nov. 7, 2019). <https://www.law.georgetown.edu/privacy-technology-center/events/color-of-surveillance-2019/>

²⁰ Bender, Emily M., Gebru, Timnit, McMillan-Major, Angelina, & Shmitchell, Shmargaret. (2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). <https://dl.acm.org/doi/10.1145/3442188.3445922>

²¹ Schwartz, Roy, Dodge, Jesse, Smith, Noah A., and Etzioni, Oren. (2020). "Green ai." *Communications of the ACM* 63, No. 12 54-63. <https://doi.org/10.1145/3381831>.

²² Dobbe, Roel, and Whittaker, Meredith. "AI and Climate Change: How they're connected, and what we can do about it." *AI Now Institute*. (Oct. 17, 2019). <https://medium.com/@AINowInstitute/ai-and-climate-change-how-theyre-connected-and-what-we-can-do-about-it-6aa8d0f5b32c>

this field are coming from scholars in disciplines like law,²³ anthropology,²⁴ and history,²⁵ and also from community organizations²⁶ gathering qualitative evidence and elevating the lived experience of the people who are surveilled, assessed, and otherwise subject to AI's determinations and predictions. These disciplines and approaches are most capable of analyzing and addressing the structural issues with AI and related technologies and should be included in a broad and flexible definition of AI R&D.

In a joint letter addressed to OSTP in July 2021, a coalition of civil rights and technology organizations called on the department to center civil rights concerns in AI and technology policy, emphasizing the need to fully incorporate the Biden Administration's Executive Order on Racial Equity into its AI policy priorities. This letter drew on years of evidence about the harms that AI is already causing, including perpetuating housing, financial services, and hiring discrimination.²⁷ This letter also calls on OSTP to "ensure that federal investment in research and development of AI technologies includes significant and immediate research on anti-discrimination measures and ways that AI systems can be used to advance equity, as well as investment in strategies to increase equity, diversity and inclusion in the tech industry."

While the Strategic Plan calls for R&D to develop AI architectures that incorporate societal concerns, we encourage reframing that centers equity and anti-discrimination as first principles and critical components of successful AI research. This is essential for facilitating rigorous, cross-disciplinary work, and avoiding the belated incorporation or retrofitting of ethics onto AI research or systems that were completed without the benefit or perspective of these disciplines.

²³Richardson, Rashida, Schultz, Jason M. & Southerland, Vincent M. (2019). "Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems." *AI Now Institute*. <https://ainowinstitute.org/litigatingalgorithms-2019-us.html>

²⁴Elish, Madeline Clare, and Watkins, Elizabeth Anne. "Repairing Innovation: A Study of Integrating AI in Clinical Care." *Data & Society Research Institute*. (Sept. 30, 2020). <https://datasociety.net/library/repairing-innovation/>; Brayne, Sarah. (2020). *Predict and Surveil: Data, Discretion, and the Future of Policing*. Oxford University Press. <https://global.oup.com/academic/product/predict-and-surveil-9780190684099?cc=us&lang=en&>

²⁵Bouk, Dan. "House Arrest: How An Automated Algorithm Constrained Congress for a Century." *Data & Society Research Institute*. (Apr. 14, 2021). <https://datasociety.net/library/house-arrest/>

²⁶For example, the Movement Alliance Project <https://movementalliance.org/about/>, Data for Black Lives <https://d4bl.org/>, Detroit Community Technology Project <https://detroitcommunitytech.org/>, Fight for the Future <https://www.fightforthefuture.org/>

²⁷Akselrod, Olga. "How Artificial Intelligence Can Deepen Racial and Economic Inequities." *American Civil Liberties Union*. (Jul. 13, 2021). <https://www.aclu.org/news/privacy-technology/how-artificial-intelligence-can-deepen-racial-and-economic-inequities>

4. The Strategic Plan should reflect the strategic importance of impact assessments as an accountability and governance mechanism for AI R&D.

The Strategic Plan doesn't yet require the measurement of impacts of AI systems supported through these strategies. The harmful impacts of AI systems must be better understood in order to effectively and ethically commit public resources to those efforts.

Multiple jurisdictions are exploring algorithmic impact assessments as a means for regulating algorithmic systems and protecting the public interest. Data & Society's research on impact assessments maps the challenges of constructing algorithmic impact assessments by analyzing their use in other domains, including finance, environment, human rights, and privacy.²⁸ To ensure the effective development of algorithmic impact assessments as a governance mechanism, this report presents a framework for evaluating impact assessment processes. This framework deepens our understanding of the mutual shaping of accountability and practices of measuring harms through impacts.

Our research shows that impact assessments are a reliable way to ensure that companies study, explain, and report on how their proposals will affect society. Not only do impact assessments illuminate the harms of specific activities or products, they also establish a much needed standard for the disclosure of where and why algorithms are being deployed, who was consulted, and what steps were taken to mitigate or prevent risks.

A recent legislative proposal seeks to mandate impact assessments in instances where automated decision systems are used to make critical decisions about our lives, demonstrating the viability of impact assessments as an accountability mechanism.²⁹ However, more research is needed to understand how best to develop and use impact assessment methodologies. Additional research is also needed to understand how best to assess the impacts of basic research, along the lines of

²⁸Moss, Emanuel, Watkins, Elizabeth Anne, Sing, Ranjit, Elish, Madeline Clare, & Metcalf, Jacob. "Assembling Accountability: Algorithmic Impact Assessment for the Public Interest." *Data & Society Research Institute*. (June, 29, 2021).

<https://datasociety.net/library/assembling-accountability-algorithmic-impact-assessment-for-the-public-interest/>.

²⁹ See "Wyden, Booker and Clarke Introduce Algorithmic Accountability Act of 2022 To Require New Transparency And Accountability For Automated Decision Systems." (Feb. 3, 2022).

<https://www.wyden.senate.gov/news/press-releases/wyden-booker-and-clarke-introduce-algorithmic-accountability-act-of-2022-to-require-new-transparency-and-accountability-for-automated-decision-systems#:~:text=Y.%2C%20to%20introduced%20the%20Algorithmic.every%20aspect%20of%20Americans%20lives.>

the technology impact work once conducted by the Office of Technology Assessment.³⁰ Without this research, the status quo in which companies conduct their own impact assessments (if at all) and don't publish the results, remains in place. This makes it difficult for researchers, policymakers, and the general public to understand and contest the impacts of AI systems.

We encourage OSTP to incorporate impact assessments into the Strategic Plan in two ways:

- **Impact assessments should be incorporated into Strategy 1 as a mechanism for determining which long-term investments in AI research maximize public interest while minimizing or preventing harm.** Our ability to control AI systems depends largely on what we know about how they work, where they are deployed, and in which ways they affect the general public. Impact assessments can make this information visible and can be useful for encouraging researchers to consider and mitigate negative downstream impacts of their research in advance of deployment. This is particularly important to do in advance of being granted access to public funding and public datasets.
- **Impact assessments should also be incorporated into Strategy 3 and Strategy 6 as areas where additional research is needed.** These assessments can expose preventable harms, encourage consultation with affected communities, and standardize the information we have available for further research about which AI systems are used in which contexts and for which purposes. The development of methodological standards for these assessments is especially critical for ensuring impact assessments are done in the public interest, and to prevent the proliferation of assessments that manipulate or obscure harmful impacts of AI systems.

Thank you for your openness to feedback on revisions to the Strategic Plan. We encourage further public engagement on areas the Strategic Plan can support that facilitate meaningful democratic control of AI and related technologies. We look forward to supporting OSTP in this effort.

Sincerely,

Brittany Smith, Policy Director
Melinda Sebastian PhD, ACLS Leading Edge Fellow

³⁰Congressional Research Service Report. "The Office of Technology Assessment: History, Authorities, Issues, and Options." (Apr. 29, 2020). <https://sgp.fas.org/crs/misc/R46327.pdf>