

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

**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.



October 1, 2021

[REDACTED]  
[REDACTED]  
[REDACTED]

Submitted electronically [REDACTED]

**RE: 86 FR 39081: Request for Information on an Implementation Plan for a National Artificial Intelligence Research Resource**

Dear Ms. Wigen:

The American Psychological Association (APA) is grateful for the opportunity to respond to the White House Office of Science and Technology Policy (OSTP) and the National Science Foundation's (NSF) [Request for Information \(RFI\): Implementation Plan for a National Artificial Intelligence Research Resource \(86 FR 39081\)](#). We thank OSTP and NSF for seeking wide input to inform federal activities attempting to democratize advanced cyberinfrastructure in the US.

APA is the leading scientific and professional organization representing psychology in the US, with more than 122,000 researchers, educators, clinicians, consultants, and students as its members and affiliates. APA's mission is to promote the advancement, communication, and application of psychological science and knowledge to benefit society and improve lives. As a broad field, psychology can contribute to the development of artificial intelligence (AI), enhance its positive impacts for individuals and society, and reduce unintended negative consequences. Our association is also committed to increasing access to scientific and educational tools for underrepresented research institutions and investigators, and to the ethical and responsible development of AI systems. We applaud the National AI Research Resource (NAIRR) Task Force's efforts to develop representative datasets, appropriate educational tools, and support mechanisms to improve equitable engagement with AI resources in the US.

The following are responses to the questions prompted by the implementation roadmap:

**Question 2: Which capabilities and services provided through the NAIRR should be prioritized?**

The Task Force has proposed several potential capabilities and services necessary to maintain a shared computing infrastructure and facilitate equitable access to resources for researchers across the country. Among these capabilities, it is essential to prioritize the appropriate choice and use of metadata – this will enhance the understanding, organization, and use of curated

datasets under NAIRR.<sup>1</sup> However, sufficiently appropriate metadata is only one necessary ingredient to meet the work plan for NAIRR as it is enumerated. For example, Section E references “high-quality government data sets” where “high-quality” implies that there is appropriate measurement, in place of mainstream measures that may be superficially appealing. The [AERA/APA/NCME Standards for Educational and Psychological Testing and Society for Industrial and Organizational Psychology \(SIOP\) Principles for the Validation and Use of Personnel Selection Procedures](#) are robust examples of valuable contributions from psychological science, including decades of psychometric and statistical research relevant to the successful execution of NAIRR’s workplan and priorities.

Per Sections D-G, human factors (HF) psychologists<sup>2</sup> are critical to the incorporation of usability, and human-systems integration into AI and data-sharing technologies.<sup>3</sup> HF researchers are necessary to ensure the AI data-sharing systems<sup>4</sup> and cyberinfrastructures developed will be useful, accessible, privacy-assuring, and fair to all parties involved. Experts who have developed, revised, implemented, and maintained taxonomies, including psychologists (and the rich ontologies supporting them) must be included to construct dataset platforms that are strong, scalable, useful, and proven capable of preserving privacy.<sup>5</sup>

**Question 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?**

Psychological science is vital to the ethical and responsible development of AI. Some AI systems have been trained on large data sets of human attributes or demographics that integrate biases related to gender, race, and other characteristics. These systems then replicate the biases in their interactions with humans, with implications for equity and fairness. Psychologists’ research on the various forms of resulting bias and the detrimental impacts are being used to develop data sets that are less biased and AI systems that can detect and compensate for biases in data.

---

<sup>1</sup> Blask, K., Gerhards, L., & Jalynskij, M. (2020, February 1). *Metadata in Psychology 2.0: What researchers really need – Study description of the data referring to the online survey conducted in the BMBF-funded project PsyCuraDat*. PsychArchives. <http://dx.doi.org/10.23668/psycharchives.2757>

<sup>2</sup> Brée D.S. (1988). *Artificial Intelligence and Cognitive Psychology: A New Look at Human Factors*. In: van der Veer G.C., Mulder G. (eds) *Human-Computer Interaction*. Springer. [https://doi.org/10.1007/978-3-642-73402-1\\_17](https://doi.org/10.1007/978-3-642-73402-1_17)

<sup>3</sup> Tenopir, C., Rice, N.M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., Sandusky, R.J., (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLOS ONE*, 15(3). <https://doi.org/10.1371/journal.pone.0229003>

<sup>4</sup> Tsuji, S., Bergmann, C., & Cristia, A. (2014). Community-Augmented Meta-Analyses: Toward Cumulative Data Assessment. *Perspectives on Psychological Science*, 9(6), 661–665. <https://doi.org/10.1177/1745691614552498>

<sup>5</sup> Hesse, B. W. (2018). Can psychology walk the walk of open science? *American Psychologist*, 73(2), 126–137. <https://doi.org/10.1037/amp0000197>

Given evidence that AI can reproduce discrimination and bias against individuals and groups, it is imperative to leverage psychological science and examine people's expectations about and reactions to the fairness and potential discrimination of AI versus human agents. An emerging line of research suggests that people expect AI to be less biased than humans in some cases, and are less outraged when they learn of bias from an AI versus human actors.<sup>6</sup> Algorithms appear less discriminatory than humans, perhaps incorrectly engendering trust and comfort from human users. Given the massive and increasing influence of AI on people's lives, it is critical to better appreciate how people understand and react to such influence, especially when the AI is perceived to be biased or unfair.

There are some fundamental research opportunities the AI research community must investigate. AI Ethics and Psychology is an evolving discipline essential to the study of how AI learns from society and humans<sup>7</sup> and how AI makes consequential decisions in critical settings.<sup>8</sup> Studies have demonstrated that AI automatically learns implicit biases from language corpora and accordingly perceives the world in a biased manner. These implicit biases that have been documented in social psychology for decades include racial, gender, sexuality, ability, and age attitudes.<sup>9</sup><sup>10</sup> Moreover, these findings provide insights about how language might be impacting the social cognition of both AI and humans.

There are, additionally, ethical implications for what AI learns, how AI learns, and AI's subsequent decision-making. For example, developing transparency enhancing algorithms for measuring and simulating AI bias and equity would make it possible to analyze the ethical implications of AI in a variety of domains including natural language and computer vision.<sup>11</sup> Alternatively, these AI methods could examine and analyze current and historical social and human cognition.<sup>12</sup> This research program would allow for understanding how AI is co-evolving

---

<sup>6</sup> Jago, A. S., & Laurin, K. (2021). Assumptions About Algorithms' Capacity for Discrimination. *Personality and Social Psychology Bulletin*. <https://doi.org/10.1177/01461672211016187>

<sup>7</sup> Caliskan, A., Bryson, J.J., & Narayanan, A., (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183-186. [10.1126/science.aal4230](https://doi.org/10.1126/science.aal4230)

<sup>8</sup> Pandey, A., & Caliskan, A., (2021). *Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy's Price Discrimination Algorithms*. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. 822-833.

<sup>9</sup> Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27. <https://doi.org/10.1037/0033-295X.102.1.4>

<sup>10</sup> Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>

<sup>11</sup> Steed, R., & Caliskan, A. (2021). A set of distinct facial traits learned by machines is not predictive of appearance bias in the wild. *AI Ethics* 1, 249–260. <https://doi.org/10.1007/s43681-020-00035-y>

<sup>12</sup> Caliskan, A., & Lewis, M. (2020, July 16). Social biases in word embeddings and their relation to human cognition. <https://doi.org/10.31234/osf.io/d84kg>

with humanity, as AI is shaping society and impacting individuals' lives in an accelerated manner and at an unprecedented scale. While beyond the scope of this Request for Information, there remains no comprehensive regulation for auditing how AI impacts equity and fairness in democratic societies.<sup>13</sup> Consequently, these promising research areas of computer and information science contribute to data-driven policy making and law while having implications for psychology, political science, sociology, linguistics, and philosophy.

**Question 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?***

NAIRR's ability to democratize access to AI R&D may rely upon the ethical and responsible development of AI systems. Ensuring ethical and responsible AI R&D will require experts who understand not only the meaning of equity, fairness, and bias, but also the nature of AI technologies and algorithms<sup>14</sup> – and importantly, the samples and settings to which they apply.<sup>15</sup> Computer scientists and engineers must enlist the social, behavioral, and psychological sciences to ensure the development less harmful AI systems that have the potential to spark positive social change.

APA looks forward to a close collaboration as this plan is further developed and implemented. Please contact with any questions, I can be reached at [REDACTED]

Sincerely,

[REDACTED]  
Mitch Prinstein, PhD  
Chief Science Officer

---

<sup>13</sup> Caliskan, A. (2021, May 10). *Detecting and mitigating bias in natural language processing*. Brookings Institution. <https://www.brookings.edu/research/detecting-and-mitigating-bias-in-natural-language-processing/>

<sup>14</sup> McLennan, S., Fiske, A., Celi, L.A. *et al.* 2020). An embedded ethics approach for AI development. *Nature Machine Intelligence* 2, 488–490. <https://doi.org/10.1038/s42256-020-0214-1>

<sup>15</sup> Sloane, M., & Moss, E. (2019). AI's social sciences deficit. *Nature Machine Intelligence* 1, 330–331. <https://doi.org/10.1038/s42256-019-0084-6>