



December 5, 2023

The Office of Management and Budget
725 17th Street, NW
Washington, DC 20503

Response to AI Memo Request for Comment

Submitted Electronically

Dear Cindy Martinez:

The American Psychological Association Services, Inc. (APA Services) respectfully submits the following response to the request for comment on the governance, innovation, and risk management for agency use of artificial intelligence (AI). APA Services is the companion organization of the American Psychological Association, which is the nation's largest scientific and professional nonprofit organization representing the discipline and profession of psychology, as well as over 146,000 members and affiliates who are clinicians, researchers, educators, consultants, and students in psychological science.

Helping ensure that technology shapes the future for the better requires understanding the psychology of human-technology interaction. Just as new technological tools are emerging, the psychological science on human interaction with new products and services is also developing. A burgeoning area of psychological science involves the development, use, and impact of artificial intelligence.

Psychologists who study human intelligence are unlocking ways to enhance artificial intelligence and to ensure its ethical and appropriate use. Applied psychologists are studying how important characteristics of AI technologies impact identity and psychological wellbeing. Clinicians are navigating an increasingly complex environment where AI tools can assist in evaluation and diagnosis. Meanwhile, research points out that the deployment of AI is not a one-size-fits-all solution and is perceived differently by different groups of people. For example, even when factors such as age, gender, education, technology skills, and country-level GDP per capita are held constant, those living in societies with greater inequality are more likely to perceive AI as a threat to their role in the workplace.¹

APA stands in support of the current draft guidance being put forward to heads of agencies related to the governance and use of AI. We believe it is essential for there to be an all-

¹ Shoss, M. K., & Ciarlante, K. (2022). Are Robots/AI Viewed as More of a Workforce Threat in Unequal Societies? Evidence From the Eurobarometer Survey. *Technology, Mind, and Behavior*, 3(2: Summer 2022). <https://doi.org/10.1037/tmb0000078>

of-government approach to governing new artificial intelligence tools and their development. This requires adequate coordination and cross-coordination across government and the involvement of psychologists, well-versed in the research necessary to make the rollout and governance of these new technologies successful. We appreciate a focus on the responsible development of new AI tools and encourage more research funding and access to be main components of this effort. Similarly, the risk-mitigation focus of this proposal is essential to safeguard against the most harmful impacts of these new technologies. We believe this memo represents an important step in the right direction and appreciate the Administration's focus on this issue.

Below are additional considerations the Office of Management and Budget and others in federal agencies should consider when seeking to understand, develop, and regulate new AI-based technologies.

AI Standards

Policymakers and employers alike should be aware of the risks and benefits inherent to developing and using AI in the workplace. To maximize the efficiency and effectiveness of AI, and minimize damages to labor markets, workplaces, and employees, APA believes that artificial intelligence should:

- **Be Human-Centric** – People must remain a central part of technologies involving artificial intelligence, including in the development, deployment, and evaluation stages.
- **Elevate Ethics & Diversity** – Diversity must be present in the individuals, backgrounds, perspectives, and thought processes of those responsible for the creation and deployment of new technologies. Companies must endeavor to embed diversity and reduce unconscious and conscious discrimination in both the development of AI tools and their deployment.
- **Be Well-Researched** – Psychological science focused on the development and deployment of artificial intelligence is a relatively new area of study and more research is needed to ensure effective development of these technologies and to understand the impact of these technologies on human behavior.
- **Be Effective & Risk Adverse**– New uses for AI technologies are discovered every day and deployment of these tools is often done without proper consideration of ethics and efficacy. AI must be properly tested, and the use of these technologies must be limited to applications where it is safe and effective in achieving the intended purpose.
- **Prioritize Privacy** – The data collection fueled by the increased deployment of AI tools is leading to increased risks for individual data privacy. Companies must redouble their efforts to protect data and ensure privacy around data collected by AI technologies. This should include copyright and trademark protection.
- **Maintain Copyright & Intellectual Property Rights** – Current AI tools and programs exist in violation of copyright and intellectual property standards required of other

similarly situated technologies, more must be done to ensure compliance with current standards and new standards must be created where necessary.

- **Be Transparently Deployed** – The prevalence of AI tools deployed in the creation of and distribution of content is often unknown by the end user, fueling misunderstandings and misconceptions about the technologies; companies should take steps to increase transparency around the presence of these tools to ensure individuals are aware of the role AI might play in impacting their experience in real world and online environments.
- **Further Digital Literacy** – It is essential that new avenues of education be paired with the deployment of new technologies. Users should be informed about the way new AI technologies work, how they might impact their lives, and how best to interact with these platforms to minimize negative impacts.

As agencies are considering areas within the federal government to centralize AI leadership, APA believes the Office of Science and Technology Policy is well situated to create cross-cutting government policies related to AI. We also believe that there is a role for the National Science Foundation, the National Institutes of Health, specifically the National Institute of Mental Health, and the National Institute on Minority Health and Health Disparities. It is imperative that the government take concrete steps now to form governing bodies and research efforts to better understand the impact of AI.

Data privacy training should be embedded into processes for the development, deployment, and evaluation of AI tools. AI technologies collect and store large amounts of data, including personally identifiable data, and increasingly health-care related data. Federal and state privacy standards must be reexamined to adequately ensure that privacy protections are adequately robust in the face of increasingly complex technologies. Adequate steps must be taken to ensure that the data collected by these tools is meeting privacy laws and copyright policies.

As with a requirement for data privacy, so should there be requirements for transparency around AI tools. Access for researchers should not be limited to only those that work within the company developing the tools and mechanisms must be built into new technologies that easily allow for researchers to access and analyze anonymized data.

Supporting Medical Innovation

While we need to continue to support and expand the mental and behavioral health care workforce, artificial intelligence and innovative technologies have the potential to positively impact the health care system and the future of service delivery. For example, AI technologies may improve patient care by enhancing diagnostic precision, allowing for more individualized treatment, and improving engagement. AI technologies may assist psychologists and other health professionals in their clinical work by automating administrative tasks and supporting clinical decision making. Additionally, AI could foster the early detection of behavioral health concerns and enable the scaling of interventions to reach a much broader segment of the population than

currently can access care. For example, digital therapeutics can be used to deliver evidence-based interventions on a scale that vastly exceeds the capacity of the current clinician workforce. These developments, if realized, represent meaningful improvements in the lives of individuals, significant benefits to overall population health, and possible decreases in health care costs.

Technology aimed at identifying and treating social isolation, mental and physical health issues, and providing new tools from transportation to caregiving show promise in improving health.² Some of these tools provide a means for helping predict health risks or recommend personalized treatment plans, and others interface directly with patients in the form of therapeutic chatbots.³

Yet, the use of AI technology within mental health care raises several ethical considerations. First, AI can be susceptible to bias; if the datasets used to train AI technologies are incomplete or reflect systemic biases, then the resulting AI technologies also will reflect those biases. Ultimately, this could exacerbate existing disparities and inequities. Secondly, given the sensitivity of mental health data, there needs to be considerable attention to data safeguarding to prevent public harm. Users of apps, chatbots, and other AI-based technology must be comprehensively informed about the privacy implications, including what personal data is being collected, how it is stored, who has access to it, and what happens to their data after they stop using the tool. Thirdly, greater transparency is needed regarding the development and technology underlying AI-based tools (while protecting copyright and intellectual property), as well as rigorous testing of the AI-based tools to foster public trust, to establish efficacy, and to minimize unintended consequences.

Ultimately, finding a balance between continued health care innovation and public safety is imperative. Psychologists should be at the forefront of AI development and employment to ensure that human-technology interactions are effective and ethical.

Medical Ethics and Protecting Patients

APA wholeheartedly agrees that new AI tools should centralize transparency and effectiveness in their development and deployment. Unexpected or unacknowledged collection of data being employed in the training of AI models raises serious privacy and transparency

² Matheny, M. E., Whicher, D., & Israni, S. T. (2020). Artificial intelligence in health care: a report from the National Academy of Medicine. *Jama*, 323(6), 509-510. <https://jamanetwork.com/journals/jama/article-abstract/2757958>; Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H. C., ... & Jeste, D. V. (2021). Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 6(9), 856-864. <https://doi.org/10.1016/j.bpsc.2021.02.001>; Dhelim, S., Chen, L., Ning, H., & Nugent, C. (2023). Artificial intelligence for suicide assessment using audiovisual cues: A review. *Artificial Intelligence Review*, 56(6), 5591-5618. <https://doi.org/10.1007/s10462-022-10290-6>

³ Fiske, A., Henningsen, P., & Buyx, A. (2019). Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *Journal of Medical Internet Research*, 21(5). <https://www.jmir.org/2019/5/e13216/>

concerns. The prevalence of AI tools is often unknown by the end user, fueling misunderstandings and misconceptions about the technologies; companies should take steps to increase transparency around the presence of these tools to ensure individuals are aware of the role AI might play in impacting their experience in real world and online environments. The use and exploitation of data will have unknown consequences on the perception and experience of AI tools.⁴ AI, like humans, is varied; yet people often see AI as homogeneous. This may result in AI failures being treated as global rather than specific, affecting perception and trust in AI.⁵ Trust between humans is earned. Humans, AI, and the context in which AI is employed determine trust in technology, and the limits of this trust must be closely considered.⁶

Transparency requirements should be paired with public education efforts centered around technological literacy. Users should be informed about the way new AI technologies work and use user-provided data, how they might impact their lives, and how best to interact with these platforms to minimize negative impacts.

Deployment of AI tools is often done without proper consideration of ethics and efficacy. AI must be properly tested prior to widespread deployment and the use of these technologies must be limited to applications where it is safe and effective, minimizing unknown, unintended, and inequitable impacts. These new AI technologies can be greatly improved by psychological science.⁷ User experience researchers and designers with strong backgrounds in psychological science can play an important role in ensuring new tools are more human, more effective, and more ethical.

The data collection being fueled by the increased deployment of AI tools is leading to increased risks for individual data privacy. Companies must be required to redouble their efforts to protect data and ensure privacy around data collected by AI technologies, especially in the case of sensitive information possessed by health care providers and when copyrighted material is present. Companies must also be more transparent related to user data practices and data used for AI training.

Practical Uses for AI in the Workplace

APA acknowledges that new AI technologies can have profound positive impacts on the workplace and worker productivity when developed, tested, and deployed in a responsible way. Already, psychologists are not only working on understanding perceptions of AI in the

⁴ Sethumadhavan, A. (2019). Trust in Artificial Intelligence. *Ergonomics in Design*, 27(2), 34–34. <https://doi.org/10.1177/1064804618818592>

⁵ Longoni, C., Cian, L., & Kyung, E. J. (2023). Algorithmic Transference: People Overgeneralize Failures of AI in the Government. *Journal of Marketing Research*, 60(1), 170–188. <https://doi.org/10.1177/00222437221110139>

⁶ Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2023). Trust in Artificial Intelligence: Meta-Analytic Findings. *Human Factors*, 65(2), 337–359. <https://doi.org/10.1177/00187208211013988>

⁷ Gopnik, A. (2017). Making AI More Human. *Scientific American*, June 2017. https://www.alisongopnik.com/Papers_Alison/scientificamerican%20Gopnik0617-60.pdf

workplace, but also are involved in developing AI for more effective use in the workplace. For example, psychologists recently used job descriptions, in tandem with natural-language processing algorithms, to estimate the knowledge and skills necessary for existing and even future jobs.⁸

In a recent survey fielded by APA and The Harris Poll, workers were asked about their perceptions of AI in the workplace.⁹ 51% of respondents who worry about AI said their work has a negative impact on their mental health, compared with 29% of workers not worried about AI making some or all of their job duties obsolete in the future. The survey shows 46% of workers worried about AI making part or all of their job duties obsolete. We see that worry about AI is more disproportionately felt among workers with less education, high school graduates worry about AI at a rate 10 percentage points higher than those that completed a four-year degree. Similar disparities were noted in workers of color and younger workers.¹⁰ This upheaval will impact the job satisfaction of workers as well as their security. 46% of individuals worried about AI intend to look for a new job in the next year, compared with 25% of those who did not report being worried. These disparities are likely to grow as the prevalence of AI technologies impact more parts of the workforce. It is essential to understand and create resources to serve the workers most vulnerable to worry and displacement by AI.

It is important to note that psychologists are not only working on understanding perceptions of AI in the workplace, but also are involved in developing AI for more effective use in the workplace. For example, psychologists recently used job descriptions, in tandem with natural-language processing algorithms, to estimate the knowledge and skills necessary for existing and even future jobs.¹¹ This work underscores the importance of including psychological scientists in the development, regulatory, and evaluation of these new technologies.

AI and Working Conditions

Three AI challenges facing policymakers are working conditions, discrimination, and job displacement. There have been recent reports that highlight the poor working conditions of those individuals responsible for training and inputting data into new AI tools.¹² Often for low wages and centralized to workers in the global south, the work necessary to create tools like ChatGPT comes at a high human cost. We know that workers in these conditions are likely to have a lower

⁸ Putka, D.J., Oswald, F.L., Landers, R.N. *et al.* (2023). Evaluating a Natural Language Processing Approach to Estimating KSA and Interest Job Analysis Ratings. *Journal of Business and Psychology*, 38, 385–410. <https://doi.org/10.1007/s10869-022-09824-0>

⁹ American Psychological Association. (2023). 2023 *Work in America* survey: Artificial intelligence, monitoring technology, and psychological well-being. <https://www.apa.org/pubs/reports/work-in-america/2023-work-america-ai-monitoring>

¹⁰ *Id.*

¹¹ Putka, D.J., Oswald, F.L., Landers, R.N. *et al.* (2023). Evaluating a Natural Language Processing Approach to Estimating KSA and Interest Job Analysis Ratings. *Journal of Business and Psychology*, 38, 385–410. <https://doi.org/10.1007/s10869-022-09824-0>

¹² Perrigo, B. (2023, January 18). Exclusive: The \$2 Per Hour Workers Who Made ChatGPT Safer. Time. <https://time.com/6247678/openai-chatgpt-kenya-workers/>

level of life satisfaction and higher prevalence of depression and anxiety.¹³ These workers are also exposed to high level of graphic images, which is an activity associated with a number of other mental health symptoms.¹⁴ Because of lack of transparency requirements currently being imposed on companies developing new AI technologies, there is not more information about these workers and their condition. The federal government can and should do more to prevent these working conditions and potential exploitation that can lead to mental health harm.

The European Union, in their recently proposed EU AI Act, have taken an important step to recognize the potential for unfairness and discrimination that new AI tools present.¹⁵ Stating that “‘diversity, non-discrimination and fairness’ means that AI systems shall be developed and used in a way that includes diverse actors and promotes equal access, gender equality and cultural diversity, while avoiding discriminatory impacts and unfair biases that are prohibited by Union or national law,” this regulation seeks to minimize the negative impacts we know are possible with AI. Adoption of a similar proposal at the federal level in the U.S. would be an important step forward in avoiding discrimination and furthering fairness. A similar framework has been proposed by psychological scientists that can also be considered.¹⁶

We are beginning to understand how AI technologies might displace individual workers or entire sections of our economy. Researchers “argue that the conditions for AI to either enhance or threaten workers’ sense of identity derived from their work depends on how the technology is functionally deployed (by complementing tasks, replacing tasks, and/or generating new tasks) and how it affects the social fabric of work.”¹⁷ Agencies have the ability to set resources aside to prepare for these negative impacts and create an environment that helps to minimize them.

AI and Workplace Bias

¹³ Howell, R. T., & Howell, C. J. (2008). The relation of economic status to subjective well-being in developing countries: A meta-analysis. *Psychological Bulletin*, 134(4), 536–560. <https://doi.org/10.1037/0033-2909.134.4.536>; Ridley, M., Rao, G., Schilbach, F., & Patel, V. (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science*, 370(6522), eaay0214. DOI: 10.1126/science.aay0214

¹⁴ Holman, E. A., Garfin, D. R., Lubens, P., & Silver, R. C. (2020). Media Exposure to Collective Trauma, Mental Health, and Functioning: Does It Matter What You See? *Clinical Psychological Science*, 8(1), 111–124. <https://doi.org/10.1177/2167702619858300>

¹⁵ DRAFT Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts. (n.d.). https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf

¹⁶ Landers, R. N., & Behrend, T. S. (2023). Auditing the AI auditors: A framework for evaluating fairness and bias in high stakes AI predictive models. *American Psychologist*, 78(1), 36–49. <https://doi.org/10.1037/amp0000972>

¹⁷ Selenko, E., Bankins, S., Shoss, M., Warburton, J., & Restubog, S. L. D. (2022). Artificial intelligence and the future of work: A functional-identity perspective. *Current Directions in Psychological Science*, 31(3), 272–279. <https://doi.org/10.1177/09637214221091823>

The early evidence shows that decisions about AI and how it is implemented reflect the world view and values of the human beings who design them and set policy for how it is used. 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.

As with other areas of technology, there remains an imperative for diversity, ethics, and inclusion to be a central component of any oversight model. Requiring diversity training, especially in teams developing AI tech can be an important first step to ensuring teams developing these new products are well versed in the components of DEI, which in turn can help to reduce the possibility of a biased product.¹⁸

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.¹⁹ Alternatively, these AI methods could examine and analyze current and historical social biases and human cognition.²⁰ Research programs focused on this area would allow for understanding how AI is co-evolving with humanity, as AI is shaping society and impacting individuals' lives in an accelerated manner and at an unprecedented scale.

Given evidence that AI can reproduce discrimination and bias against individuals and groups, it is imperative the government 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 while being less outraged when they learn of bias from an AI versus human actors.²¹ Algorithms appear less discriminatory than humans, perhaps incorrectly engendering trust and comfort from human users.

Without the inclusion of psychological science, we risk harming already disadvantaged populations and creating systems that perpetuate harmful stereotypes and bias. AI systems are often trained using large data sets of human attributes or demographics that have the potential to integrate biases related to gender identity, race, and other characteristics. These systems then

¹⁸ Cowgill, Bo and Dell'Acqua, Fabrizio and Deng, Sam and Hsu, Daniel and Verma, Nakul and Chaintreau, Augustin, Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics (June 1, 2020). In Proceedings of the 21st ACM Conference on Economics and Computation (pp. 679-681), Columbia Business School Research Paper Forthcoming, Available at SSRN: <https://ssrn.com/abstract=3615404> or <http://dx.doi.org/10.2139/ssrn.3615404>

¹⁹ 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>.

²⁰ 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>.

²¹ Jago, A. S., & Laurin, K. (2021). Assumptions About Algorithms' Capacity for Discrimination. *Personality and Social Psychology Bulletin*. <https://doi.org/10.1177/01461672211016187>.



spread the biases in their interactions with humans or other technology-informed systems, 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. Findings from this research should be incorporated into future deployments of artificial intelligence tools, especially when being funded or used by the federal government.

Thank you again for the opportunity to weigh in on these questions. To the extent that our psychologists can serve as a resource to your office, please feel free to reach out to Corbin Evans (Senior Director Congressional & Federal Relations, Science Advocacy) at CEvans@apa.org.

Sincerely,

A handwritten signature in black ink, appearing to read 'Katherine B. McGuire', with a long, sweeping horizontal line extending to the right.

Katherine B. McGuire
Chief Advocacy Officer
American Psychological Association Services, Inc.