

美国心理学会代表委员会

关于“人工智能与心理领域”的政策声明

编译：刘明子，卫垆圻

美国心理学会（American Psychological Association, APA）成立于1892年，是美国心理学学术组织。APA的使命是促进心理科学和知识的进步、交流和应用，以造福社会，改善人们的生活。APA代表委员会（APA Council of Representatives）是APA的主要决策机构。2024年8月该机构发布了一份关于“人工智能与心理学领域”（Artificial Intelligence and the Field of Psychology）的政策声明。以下为该政策声明的原文编译稿。

本政策声明旨在充分认识到人工智能技术，包括机器人和聊天机器人技术、数据收集和分析、机器学习算法和自动化决策技术等对当前和未来可能产生的变革性影响，并寻求明确心理科学和知识在指导不同形式AI的应用和分析方面发挥关键作用，以及AI对心理学研究、培训、实践和应用的现实和潜在影响。APA申明心理学在应对AI对社会、健康和福祉挑战和机遇上，以及AI的伦理和隐私方面会发挥重要作用，并致力于识别、减轻和消除AI的有害影响，同时确保社会中的每个个体都能从AI技术的进步中最大程度获益。

心理科学对AI领域的影响

AI技术正越来越多地融入人们的日常生活(Matheny et al., 2020; Bankins et al., 2023; Odekerken-Schröder et al., 2020)。鉴于心理学在人类认知、行为和情感方面有着深入的科学理解，心理学应该在塑造和研究安全高效的AI系统方面发挥至关重要的作用，包括人们如何使用并与这些系统互动，研究系统对社会的影响和贡献。心理学亟需在改善人机互动上给出指导性意见，促进公平、道德和公正地使用AI(Landers & Behrend, 2023)。

AI对心理科学领域的影响

AI将影响当前和未来的心理学及其子学科的教育、培训、实践和研究(Götz

et al., 2023; Fan et al., 2023)。基于上述 AI 对心理学领域的重要影响，APA 决议要继续研究、分析、批判和促进 AI 技术在心理学领域的应用。

APA 机构及心理学学科在 AI 领域的独特作用可以分为以下三个相互关联的领域，主要包括以下三个方面：

塑造 AI 的社会影响力(Shaping AI's Societal Impact)

AI 对医疗保健、教育、刑事法律体系、交通出行、工作等社会系统都有突出贡献。在此背景下，心理科学所提供的宝贵见解与知识，可以帮助理解 AI 正在如何——以及将会如何塑造社会模式与趋势，这包括虚假信息和偏见的产生、延缓和减轻(Garimella & Chauchard, 2024)；有效数字素养工具的创建、实施与改进(Tiernan et al., 2023)；以及 AI 对儿童、成人和整个社会的行为模式和福祉的影响(Odekerken Schröder et al., 2020; Tang et al., 2023)。

通过 AI 技术促进人类健康与福祉(Using AI to Promote Health and Wellbeing)

通过与心理学专家合作，AI 技术可以更好地促进人口健康、提高医疗保健可及性、减少医疗的不公平性、增加早期干预、调整疾病预防策略、降低疾病风险并提高整个社会的人类福祉(Matheny et al., 2020; Ophir et al., 2022; Pierson et al., 2021)。只有将心理学纳入到 AI 领域的研究、开发、部署和分析，才能发挥上述 AI 技术的潜力。

处理和维持 AI 相关的道德和隐私问题(Addressing and Upholding Ethics and Privacy Related to AI)

人工智能的效率、准确性和解决问题的能力必须在正确或错误使用这类技术的道德和隐私问题的背景下加以考虑。长期以来，心理学研究一直站在识别和减弱道德侵犯问题的最前沿，这类问题也在心理科学的长期帮助下得到解决。总体而言，在 AI 技术背景下制定的新政策应当纳入道德和隐私问题相关的心理学专业知识(Tay et al., 2022)。

APA 力求在最大程度上确保美国心理学会的工作和上述三个领域的工作依据以下 7 个原则进行：

(1) 为了确保 AI 技术和系统的伦理发展、应用及评估始终以人为本，需以研究人类行为和经验、以人为本的心理学科为底层逻辑，将伦理和人权置于中

心地位。

(2) 通过心理学研究促进公平并识别和减少偏见，从而促进 AI 的安全性和有效性，最大限度地减少未知的、未预期的和不公平的影响。此外，需将多元化作为人工智能技术、算法和数据的开发、传播、监测和改进的核心。

(3) 促进心理科学的创新、讨论和传播，从而进一步理解和评估人工智能技术的发展和应用，以确保它们可以造福儿童、成人，甚至是整个社会。

(4) 确保政策制定者了解心理科学在影响人工智能工具和系统方面的专业知识和关键作用，并在监管、立法和司法的努力中融入该类专业知识，从而探索管理和规范人工智能领域的方法。

(5) 通过信息素养培训向公众传授符合其发展水平的相关知识，使他们能够使用心理学原理，能够最好地区分人工智能和现实，理解人工智能的含义，并尽可能地从人工智能中获益。

(6) 为寻求在心理科学的创新、教学和实践中使用 AI 技术的人提供指导和反馈。

(7) 为寻求开发人工智能技术的人提供指导和反馈，以确保他们以一种基于心理科学并坚持道德的方式进行开发。

参考文献

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2023). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, 45(2), 159-182. <https://doi.org/10.1002/job.2735>

Fan, J., Sun, T., Liu, J., Zhao, T., Zhang, B., Chen, Z., Glorioso, M., & Hack, E. (2023). How well can an AI chatbot infer personality? Examining psychometric properties of machine-inferred personality scores. *Journal of Applied Psychology*, 108(8), 1277-1299. <https://doi.org/10.1037/apl0001082>

Garimella, K., & Chauchard, S. (2024). How prevalent is AI misinformation? What our studies in India show so far. *Nature*, 630(8015), 32-34.

<https://doi.org/10.1038/d41586-024-01588-2>

- Götz, F. M., Maertens, R., Loomba, S., & van der Linden, S. (2023). Let the algorithm speak: How to use neural networks for automatic item generation in psychological scale development. *Psychological Methods*. *Advance online publication*. <https://doi.org/10.1037/met0000540>
- 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>
- 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://doi.org/10.1001/jama.2019.21579>
- Odekerken-Schröder, G., Mele, C., Russo-Spena, T., Mahr, D., & Ruggiero, A. (2020). Mitigating loneliness with companion robots in the COVID-19 pandemic and beyond: an integrative framework and research agenda. *Journal of Service Management*, *31*(6), 1149-1162. <https://doi.org/10.1108/JOSM-05-2020-0148>
- Ophir, Y., Tikochinski, R., Brunstein Klomek, A., & Reichart, R. (2022). The hitchhiker's guide to computational linguistics in suicide prevention. *Clinical Psychological Science*, *10*(2), 212-235. <https://doi.org/10.1177/21677026211022013>
- Pierson, E., Cutler, D.M., Leskovec, J. et al. (2021). An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nature Medicine*, *27*, 136–140. <https://doi.org/10.1038/s41591-020-01192-7>
- Tang, P. M., Koopman, J., Mai, K. M., De Cremer, D., Zhang, J. H., Reynders, P., Ng, C. T. S., & Chen, I-H. (2023). No person is an island: Unpacking the work and after-work consequences of interacting with artificial intelligence. *Journal of*

Applied Psychology. Advance online publication.
<https://doi.org/10.1037/apl0001103>

Tay, L., Woo, S. E., Hickman, L., Booth, B. M., & D'Mello, S. (2022). A conceptual framework for investigating and mitigating machine-learning measurement bias (MLMB) in psychological assessment. *Advances in Methods and Practices in Psychological Science*, 5(1), 25152459211061337.
<https://doi.org/10.1177/25152459211061337>

Tiernan, P., Costello, E., Donlon, E., Parysz, M., & Scriney, M. (2023). Information and Media Literacy in the Age of AI: Options for the Future. *Education Sciences*, 13(9), 906. <https://doi.org/10.3390/educsci13090906>

原文标题:

Artificial Intelligence and the Field of Psychology

原文链接:

<https://www.apa.org/about/policy/artificial-intelligence-psychology>