2024

On Human-Centered AI in Education

Sai Gattupalli
University of Massachusetts Amherst

Robert W. Maloy
University of Massachusetts Amherst

On Human-Centered AI in Education

Sai Gattupalli & Robert Maloy
Department of Teacher Education and Curriculum Studies
College of Education
University of Massachusetts Amherst
sgattupalli, rwm@umass.edu

Abstract

As AI advances, human-centered principles are key to harnessing its benefits ethically. We explore scaling human-centered AI to enrich education. Thoughtfully implemented, AI could enable personalized, equitable learning and amplify teachers' strengths, and also facilitate more intuitive human-AI collaboration. However, benefits require mitigating risks around privacy, bias, transparency, and social-emotional impacts. Multidisciplinary teams should research embedding ethics into systems. Policymakers need to develop guardrails for privacy, fairness and accountability. Schools should pilot applications cautiously and demand explainable AI. Diverse voices must guide tool development to enhance autonomy and inclusion. With care, human-centered AI may propel an educational renaissance that uplifts learning and the human spirit. But we must build this future on moral foundations serving all students.

Introduction

Artificial Intelligence (AI) is reshaping our lives through mobile phones, smart TVs, advanced cars, and many more technologies, including its latest development: GenAI or Generative Artificial Intelligence tools. This advanced AI, demonstrated by its ability to pass numerous standardized tests (Terwiesch, 2023; Newton & Xiromeriti, 2023; Jung et al., 2023; Bommineni et al., 2023; Biever, 2023; Schubert, 2023), is revolutionizing benchmarks in job qualifications. Its emergence poses pivotal questions about AI's role in education and society. In the field of human-centered AI (HCAI), these developments take on added significance. HCAI is not merely about technological progress but about enhancing human capabilities and aligning AI with ethical and societal values (Cheng et al., 2021), fundamentally redefining how we approach learning and personal development in an increasingly AI-integrated world. Recent breakthroughs in machine learning techniques have rapidly expanded AI's capabilities, from personalized learning to intelligent assistants (Xu et al., 2021; Aggarwal, 2018; Deng, 2018). Yet, as AI increasingly impacts human lives, concerns arise about job displacement, privacy erosion, and ethical alignment (Sinha et al., 2023).

HCAI addresses these challenges by prioritizing human values, context, and agency (Xu, Dainoff, et al., 2022; Renz & Vladova, 2021). It champions AI as a tool to augment human potential, driven by guiding principles like transparency, accountability, and fairness (Shneiderman, 2022; Auernhammer, 2020). HCAI in education particularly emphasizes personalized and equitable learning, supporting teachers, and fostering human-machine collaboration, while mitigating risks to privacy and social development (Yang et al., 2021; Cheng et al., 2021).

By embedding human values from the outset, HCAI promises not only a technological leap in education but a reaffirmation of its humanistic goals. In this paper, we explore how HCAI can revolutionize education, ensuring AI serves as a catalyst for learning, accessibility, and joy, grounded in an ethical foundation for the common good.

Why Education Needs Human-Centered AI

Education aims to nurture individuals and prepare them for a society that values human-centered approaches. It's essential to ponder whether AI can support educational goals without compromising its fundamental values. Human-Centered AI (HCAI) in education emerges as a crucial element in this context (Renz & Vladova, 2021).

Traditional AI systems, often focused on efficiency and predictive power, can inadvertently perpetuate biases and overlook the diverse needs of educational environments (Fountain, 2022; Baker & Hawn, 2021). HCAI, in contrast, is dedicated to aligning AI with the diverse goals, needs, and preferences of educational stakeholders, including students, teachers, parents, and policymakers (Shneiderman, 2022). This approach emphasizes involving these stakeholders in every stage of AI's development, from design to evaluation, ensuring a balance between human control and AI automation. Crucially, it underscores the importance of creating safe, reliable, and trustworthy AI systems in educational settings (Domfeh et al., 2022).

HCAI also focuses on fostering human-AI collaboration and co-creation, especially in learning and teaching scenarios (such as Alfredo et al., 2023). This aspect is key to developing AI literacy and equipping students for the future (Ng et al., 2021; Casal-Otero et al., 2023). By integrating HCAI, educational institutions can ensure AI tools respect and enhance human roles, thereby augmenting human development and agency.

Moving forward, the responsible scaling of HCAI in education demands a participatory approach. This means extensive collaboration among students, parents, educators, technologists, and community members in shaping AI to support human development (Schoenherr et al., 2023; Capel & Brereton, 2023). It also necessitates stringent ethical guardrails around privacy, bias, and transparency. AI tools in education should be developed with a nuanced understanding of the social and psychological realities of learning, ensuring they are driven by the voices of teachers and students to enhance autonomy and self-actualization (Luckin & Cukurova, 2019; Schiff, 2020).

By adopting a human-centered approach, AI has the potential to transform education (Yang et al., 2021). It can offer personalized, equitable instruction, alleviate the burden on teachers, and enable more intuitive human-computer interactions. The successful integration of HCAI in education hinges on a collective commitment to ethical principles and shared values, ensuring AI supports and enriches the human experience in education.

Key Opportunities for Scaling Human-Centered AI in Education
The potential of HCAI in education is transformative, aligning AI systems with human values, needs, and goals (Renz & Vladova, 2021; Yang et al., 2021; Shneiderman, 2022; Ozmen Garibay et al., 2023). For researchers, teachers, students, and policymakers, the focus should be on high-potential applications that can revolutionize learning experiences when implemented with care and collaboration.

- Personalized Learning: HCAI can develop adaptive learning systems that customize education for each student, analyzing data from exercises and assessments to provide immediate, relevant feedback (Yang et al., 2021; Mørch & Andersen, 2023). These systems, grounded in learning science principles and promote a growth mindset, ensure they support rather than limit students' learning.
- Intelligent Tutoring: AI tutors adapting to various learning styles can offer expert guidance anytime (such as MathSpring by Arroyo et al., 2014; Lin et al., 2023). As virtual mentors, they can support goal-setting and reflection, but their design needs to consider the socio-emotional aspects of learning to avoid creating frustration. However, poorly designed systems might lack nuance or frustrate learners. Developing AI tutors that model human wisdom requires extensive research into the social dimensions of learning.
- Teacher Empowerment: HCAI can alleviate teacher workloads by automating administrative tasks and offering insightful curriculum or student support suggestions

(Andersen et al., 2022). These tools should respect teacher autonomy and privacy, avoiding excessive monitoring. Other opportunities lie in improving school management to empower teachers. AI could assist with administrative tasks, freeing up time and reducing burnout (Johnson, 2023). Teachers might get AI-generated insights about how to refine their curriculum or support individual students. The key will be designing tools to provide helpful analysis without excessive surveillance.

- Learning at scale: One major opportunity is using AI for personalized learning at scale (Tan, 2023). Adaptive learning systems can tailor instruction, remediation, and enrichment to each student's strengths and needs. Analyzing patterns in exercises and assessments, AI tutors can provide timely feedback and suggest resources to fill knowledge gaps. Researchers must develop these AI tools to align with learning science (Van Mechelen et al., 2023), foster growth mindsets, and avoid pigeon-holing students.

Realizing these opportunities will require extensive participatory research. Multi-disciplinary teams must investigate AI techniques that model the social and emotional dimensions of learning. Schools should pilot applications in constrained settings before scaling. And communities must have a voice in shaping how AI is applied locally.

With care and foresight, HCAI could profoundly enrich education. Students might learn faster and more joyfully while developing life skills for an AI-influenced world. Teachers could focus more attention on the human side of instruction. And new possibilities for human-AI collaboration could enhance knowledge in ways beyond imagination today. The future starts now—we must begin building it with both wisdom and optimism.

Challenges and Risks
In the field of HCAI in education, profound challenges require vigilant attention. Ethical development of AI necessitates addressing privacy concerns, algorithmic biases, and ensuring transparency. For instance, rigorous audits of algorithms, as demonstrated by the Amazon hiring algorithm's bias case (IBM Data & AI Team, 2023), highlight the need for ethical oversight, particularly in sensitive areas like admissions. This principle extends to education, where educators, administrators, and policymakers must be aware of similar risks. Balancing data collection for personalized learning with privacy protections calls for clear policies around consent and data use. Additionally, the challenge of biases in algorithms demands careful testing and contextual understanding. The importance of transparency, such as the debate between open vs closed source software, is highlighted by the complexity of AI models which can obscure decision-making processes, affecting human agency and accountability. UNESCO's recent

guidance on generative AI in education underlines further issues like digital poverty, regulatory challenges, unauthorized content use, unexplainable models, and the risks of "deepfakes", pointing towards the urgent need for comprehensive teacher training (Fengchun & Wayne, 2023). Teachers are crucial in leading AI literacy efforts, equipping them to navigate AI's complexities in the classroom. Addressing these challenges with a commitment to transparency and participation, guided by human values, is essential. Responsibly managed, AI has the potential to significantly enrich learning, but it must be built upon moral foundations that affirm our shared humanity.

Conclusion

As AI continues its rapid advancement, the field of education stands at a crossroads. We must thoughtfully shape the role of these emerging technologies to align with humanistic educational values and goals. HCAI offers a guiding paradigm—developing AI tools not merely for efficiency, but expressly to enhance human capabilities and learning experiences. This approach opens up transformative possibilities in education, from personalized and equitable instruction at scale to more intuitive human-AI collaboration.

However, realizing this potential requires vigilant efforts. Extensive research is needed so that AI systems model the social and emotional dimensions of learning. Pilot applications should be implemented cautiously to identify pitfalls before broader rollout (such as ChatGPT and Claude AI). Policymakers must institute guardrails regarding privacy, fairness and accountability. And diverse voices, especially teachers and students, should drive tool development to focus squarely on enhancing autonomy and inclusion for all.

Fundamentally, AI is not destiny but a choice. We must collectively build the future of AI in education on ethical foundations serving the common good. If developed thoughtfully through a human-centered lens, AI could propel an educational renaissance—amplifying the human spirit through more enlightened, empowering learning for all. But we should never lose sight of the supreme aim—an educational system that uplifts our shared humanity. The promise and perils of AI force us to re-examine and reaffirm our moral commitments. With wisdom and courage, we can shape emerging technologies to enrich education ethically for generations to come.

References

Aggarwal, C. C. (2018). Training deep neural networks. In *Neural Networks and Deep Learning* (pp. 105–167). Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-94463-0_3

Akbarighatar, P., Pappas, I., & Vassilakopoulou, P. (2023). A sociotechnical perspective for responsible AI maturity models: Findings from a mixed-method literature review. *International Journal of Information Management Data Insights*, *3*(2), 100193. https://doi.org/10.1016/j.jjimei.2023.100193

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2023, December 20). *Human-Centred learning analytics and AI in education: A systematic literature review*. arXiv.Org. https://arxiv.org/abs/2312.12751

Andersen, Renate, Gjølstad, Eli, Mørch, & Anders. (2022). Integrating human-centered artificial intelligence in programming practices to reduce teachers' workload. *CEUR Workshop Proceedings*, *3136*, 30–35.

Anderson, J., & Rainie, L. (2018, December 10). *Artificial intelligence and the future of humans*. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2018/12/10/artificial-intelligence-and-the-future-of-humans/

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A Multimedia Adaptive Tutoring System for Mathematics that Addresses Cognition, Metacognition and Affect. *International Journal of Artificial Intelligence in Education*, *24*(4), 387–426. https://doi.org/10.1007/s40593-014-0023-y

Auernhammer, J. (2020, September 10). Human-centered AI: The role of Human-centered Design Research in the development of AI. *DRS2020: Synergy*. http://dx.doi.org/10.21606/drs.2020.282

Baker, R. S., & Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, *32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

Biever, C. (2023, July 25). *ChatGPT broke the Turing test — the race is on for new ways to assess AI*. Nature. https://www.nature.com/articles/d41586-023-02361-7

Bommineni, V. L., Bhagwagar, S., Balcarcel, D., Davatzikos, C., & Boyer, D. (2023). *Performance of chatgpt on the MCAT: The road to personalized and equitable premedical learning*. Cold Spring Harbor Laboratory. http://dx.doi.org/10.1101/2023.03.05.23286533

Capel, T., & Brereton, M. (2023, April 19). What is Human-Centered about Human-Centered AI? A Map of the Research Landscape. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. http://dx.doi.org/10.1145/3544548.3580959

Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, *10*(1). https://doi.org/10.1186/s40594-023-00418-7

Chen, Y., Clayton, E. W., Novak, L. L., Anders, S., & Malin, B. (2023). Human-Centered design to address biases in artificial intelligence. *Journal of Medical Internet Research*, *25*, e43251. https://doi.org/10.2196/43251

Cheng, L., Varshney, K. R., & Liu, H. (2021). Socially responsible AI algorithms: Issues, purposes, and challenges. *Journal of Artificial Intelligence Research*, *71*, 1137–1181. https://doi.org/10.1613/jair.1.12814

Data, I., & Team, A. (2023, October 16). Shedding light on AI bias with real world examples. *Security Intelligence*. https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/

Deng, L. (2018). Artificial intelligence in the rising wave of deep learning: The historical path and future outlook [perspectives]. *IEEE Signal Processing Magazine*, *35*(1), 180–177. https://doi.org/10.1109/msp.2017.2762725

Domfeh, E. A., Weyori, B., Appiahene, P., Mensah, J., Awarayi, N., & Afrifa, S. (2022). *Human-Centered Artificial Intelligence, a review*. Authorea, Inc. http://dx.doi.org/10.22541/au.166013641.15972664/v1

Fazelpour, S., & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, *16*(8). https://doi.org/10.1111/phc3.12760

Fengchun, M., & Wayne, H. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Fountain, J. E. (2022). The moon, the ghetto and artificial intelligence: Reducing systemic racism in computational algorithms. *Government Information Quarterly*, *39*(2), 101645. https://doi.org/10.1016/j.giq.2021.101645

Geiger, R. S., Tandon, U., Gakhokidze, A., Song, L., & Irani, L. (2023). Rethinking Artificial Intelligence: Algorithmic bias and ethical issue. *International Journal of Communication*, *18*(0).

Johnson, K. (2023, September 15). Teachers are going all in on generative AI. *WIRED*. https://www.wired.com/story/teachers-are-going-all-in-on-generative-ai/

Jung, L. B., Gudera, J. A., Wiegand, T. L. T., Allmendinger, S., Dimitriadis, K., & Koerte, I. K. (2023). ChatGPT passes German state examination in medicine with picture questions omitted. *Deutsches Ärzteblatt International*. https://doi.org/10.3238/arztebl.m2023.0113

Kunj Bihari, M., & Tyagi, V. (2022). Blockchain technology: Past, present, and future. *Indian Science Cruiser*, *36*(2), 52. https://doi.org/10.24906/isc/2022/v36/i2/212547

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

Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring

systems toward sustainable education: A systematic review. *Smart Learning Environments*, *10*(1). https://doi.org/10.1186/s40561-023-00260-y

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences‑driven approach. *British Journal of Educational Technology*, *50*(6), 2824–2838. https://doi.org/10.1111/bjet.12861

Margetis, G., Ntoa, S., Antona, M., & Stephanidis, C. (2021). HUMAN‑CENTERED DESIGN OF ARTIFICIAL INTELLIGENCE. *HANDBOOK OF HUMAN FACTORS AND ERGONOMICS*, 1085–1106. https://doi.org/10.1002/9781119636113.ch42

Mørch, A. I., & Andersen, R. (2023). Morch-Andersen-CoPDA2023-Formatted-May17. *CEUR Workshop Proceedings (CEUR-WS.Org)*.

Newton, P. M., & Xiromeriti, M. (2023). *ChatGPT performance on MCQ exams in higher education. A pragmatic scoping review*. Center for Open Science. http://dx.doi.org/10.35542/osf.io/sytu3

Ng, D. T. K., Leung, J. K. L., Chu, K. W. S., & Qiao, M. S. (2021). AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology*, *58*(1), 504–509. https://doi.org/10.1002/pra2.487

Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S. M., Garibay, I., Grieman, K., Havens, J. C., Jirotka, M., Kacorri, H., Karwowski, W., Kider, J., Konstan, J., Koon, S., Lopez-Gonzalez, M., Maifeld-Carucci, I., … Xu, W. (2023). Six human-centered artificial intelligence grand challenges. *International Journal of Human–Computer Interaction*, *39*(3), 391–437. https://doi.org/10.1080/10447318.2022.2153320

Renz, A., & Vladova, G. (2021). Reinvigorating the discourse on human-centered artificial intelligence in educational technologies. *Technology Innovation Management Review*, *11*(5), 5–16. https://doi.org/10.22215/timreview/1438

Schiff, D. (2020). Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI & SOCIETY*, *36*(1), 331–348. https://doi.org/10.1007/s00146-020-01033-8

Schoenherr, J. R., Abbas, R., Michael, K., Rivas, P., & Anderson, T. D. (2023). Designing AI using a human-centered approach: Explainability and accuracy toward trustworthiness. *IEEE Transactions on Technology and Society*, *4*(1), 9–23. https://doi.org/10.1109/tts.2023.3257627

Schubert, M. C. (2023). Performance of large language models on a neurology board–style examination. *JAMA Network Open*, *6*(12), e2346721–e2346721. https://doi.org/10.1001/jamanetworkopen.2023.46721

Shneiderman, B. (2020). Human-Centered artificial intelligence: Reliable, safe & trustworthy.

International Journal of Human–Computer Interaction, *36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Shneiderman, B. (2022). Defining reliable, safe, and trustworthy systems. In *Human-Centered AI* (pp. 53–56). Oxford University PressOxford. http://dx.doi.org/10.1093/oso/9780192845290.003.0007

Sinha, A. L., Muley, H., Ghosh, J., & Sarode, P. (2023, June 1). AI based Desktop Voice Assistant for Visually Impared Persons. *2023 8th International Conference on Communication and Electronics Systems (ICCES)*. http://dx.doi.org/10.1109/icces57224.2023.10192894

Tan, S. (2023). Harnessing artificial intelligence for innovation in education. In *Learning Intelligence: Innovative and Digital Transformative Learning Strategies* (pp. 335–363). Springer Nature Singapore. http://dx.doi.org/10.1007/978-981-19-9201-8_8

Terwiesch, C. (2023). Would chat GPT3 get a Wharton MBA. *A Prediction Based on Its Performance in the Operations Management Course. Wharton: Mack Institute for Innovation Management/University of Pennsylvania/School Wharton.*

Van Mechelen, M., Smith, R. C., Schaper, M.-M., Tamashiro, M., Bilstrup, K.-E., Lunding, M., Graves Petersen, M., & Sejer Iversen, O. (2023). Emerging technologies in K–12 education: A future HCI research agenda. *ACM Transactions on Computer-Human Interaction*, *30*(3), 1–40. https://doi.org/10.1145/3569897

Xu, W. (2019). Toward human-centered AI. *Interactions*, *26*(4), 42–46. https://doi.org/10.1145/3328485

Xu, W., Dainoff, M. J., Ge, L., & Gao, Z. (2022). Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI. *International Journal of Human–Computer Interaction*, *39*(3), 494–518. https://doi.org/10.1080/10447318.2022.2041900

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.-W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., … Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, *2*(4), 100179. https://doi.org/10.1016/j.xinn.2021.100179

Yang, S. J. H., Ogata, H., Matsui, T., & Chen, N.-S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence*, *2*, 100008. https://doi.org/10.1016/j.caeai.2021.100008