



Towards Human-Centred Artificial Intelligence for Education

Stephen Bezzina¹, Achilina Hasna Chariuc²

University of Malta, Malta¹
Università eCampus, Italy²

Abstract

Artificial Intelligence (AI) is revolutionising education by promising new opportunities for adaptivity, personalisation, accessibility, and data-driven decision-making. However, beyond the hype, the design, development and application of AI for education must be embedded in ways that enhance rather than impair or compromise human capabilities. To this extent, Human-Centred AI (HCAI) is a branch of Artificial Intelligence which emphasises the central role of the human element in the AI revolution. Consequently, this approach accentuates the need to ensure that AI is transparent, trustworthy, and explainable and that it respects privacy and human rights. This is particularly relevant to education, where the interaction between students, teachers, parents, administrators and all stakeholders remains primarily human in nature. To this extent, this qualitative study aims to explore the perspectives of AI specialists and educators on HCAI for education. A focus group was held with two AI scientists and two educators, and thematic analysis was used to identify and interpret patterns of meaning within the data. The main themes that emerged from this focus group reiterate important current debates in the world of education, from personalised adaptive learning and automated grading to trust, explainability, accessibility and inclusivity. The findings suggest that AI already plays and will keep on playing a major role in underpinning the future of education. Consequently, it must be rooted in a human-centric framework prioritising human well-being, capability and ethics, underpinned by pedagogy-driven AI design, educator training and policy development. Possible further research in the area involves examining the impact of human-centred AI in real-world educational settings to understand better the benefits, challenges and limitations that arise during implementation.

Keywords: artificial intelligence, human-centred, education, trust, explainability, ethics

1. Introduction

Artificial intelligence (AI) is increasingly present in many aspects of our daily lives, often helping to simplify them. Algorithmic processing already determines resource allocation, marketing strategies, the prevention of natural events and human behaviour [1]. Furthermore, AI is transforming the educational landscape by offering unprecedented possibilities for adaptability, customisation, improved accessibility, and decisions based on data analysis [2]. In fact, the impact of digital technologies, content and processes can be seen across the educational ecosystem (including schools, higher education and also informal and non-formal learning), affecting all aspects of education (such as curricular reform, teaching and learning practices, assessment, initial and continuing teacher professional development) and encompassing all stakeholders (teachers, learners, school leaders, parents and policymakers).

2. Human-Centred Artificial Intelligence

Beyond the hype, the ubiquity of technology, in particular the ever-growing application of AI in education, must be embedded in ways that enhance rather than impair or compromise human capabilities. In this regard, human nature and beings must be placed at the centre of human-machine interactions. Human-Centred AI (HCAI) is a branch of AI which emphasises the central role of the human element in the AI revolution [3]. Consequently, this approach accentuates the need to ensure that AI respects human beings and their rights. This is particularly relevant to education, where the interaction between all stakeholders remains primarily human in nature. HCAI aims to revolutionise teaching and learning practices while accelerating progress towards the UN's Sustainable Development Goal 4 (SDG 4) of equitable and inclusive education for all [4].

In this regard, explainability, trustworthiness, and privacy are the three main tenets underpinning HCAI [5]. These factors are crucial in building AI systems that can be trusted by the users, as they are often vulnerable to attacks, biased against underrepresented groups, and lack adequate user privacy protection. Explainability, as it pertains to HCAI, is about fostering a clear and transparent understanding of how an AI system functions. This is especially important in the educational setting.



Teachers, students, and other educational stakeholders need to comprehend the reasoning behind AI-driven recommendations or decisions. Consequently, this level of transparency builds trust in the system and allows for its better integration into the educational environment. To this extent, trustworthiness, in the context of HCAI, is the creation of fair and justifiable systems based on reliable and ethical decisions. In an educational environment, it is of paramount importance that AI systems provide fair and unbiased suggestions and/or decisions. Moreover, as the third tenet of HCAI, privacy posits that AI systems must handle personal data responsibly, ensuring that sensitive information about students' learning patterns, performance, and personal characteristics is protected and not misused. The use of such data should be transparent, and users should have control over how their data is used. The intersection of these tenets is also of significant importance, as the pursuit of one could interfere with the achievement of another. For instance, the need for data privacy might conflict with the desire to explain the system output in detail. Therefore, building HCAI systems requires a balanced approach, taking into account the joint optimisation and trade-offs between multiple aspects of explainability, trustworthiness, and privacy.

3. Research Question

The aim of this paper is to broaden the conceptual understanding of HCAI in and for education by examining the perspectives of two AI specialists and two educators on the application and potential of HCAI within the educational sphere. To address this objective, the following research question has been formulated: *How can HCAI contribute to the enhancement of educational practices?*

Accordingly, this paper endeavours to explore the potential and limitations of HCAI in education through a qualitative discussion with experts in the field, and to provide practical insights into the design, development, and implementation of HCAI in educational settings. By analysing the views and opinions of the experts, the paper aims to elucidate the potential of HCAI to transform educational practices.

4. Methodology

In order to meet the research objective, an online focus group lasting ninety minutes was organised with professionals in AI and education. Two AI specialists; a male with 3 years of experience (AIS-1) and a female with 6 years of experience (AIS-2) and two educators; a male with 7 years of experience (E-1) and a female with 2 years of experience (E-2) participated in the focus group. The session, led by the authors, began with a brief introduction to the study's purpose before inviting the participants to share their insights and experiences. The discussion explored not only the specific elements of HCAI but also the participants' underlying perceptions and potential benefits posed by its integration into educational practices.

This study adopts a qualitative approach, employing thematic analysis to identify and interpret patterns of meaning within the data gathered from the focus group [6]. An audio recording of the session was transcribed and subsequently analysed using NVivo™ to uncover key themes. Each segment was reviewed individually, generating codes that were then analysed. The final step, involved refining the codes and grouping them into distinct thematic categories, which are discussed in the subsequent section.

5. Results

The thematic analysis unearthed four primary themes associated with the potential advantages of incorporating HCAI into education practices. These themes include personalised adaptive teaching and learning, automated grading, fostering understanding and trust, and promoting accessibility and inclusivity.

The first theme, personalised adaptive learning, highlights how HCAI can facilitate a tailored learning experience. AIS-1 stated that HCAI's capacity to adapt educational content and pace based on individual student's progress and needs enhances personalised learning. Further, E-1 argued that the provision of dynamic, real-time feedback tailored to each student's needs supports their learning journey and aids in surmounting specific learning hurdles. AIS-2 noted the vital role of continuous assessment in HCAI, emphasising that it helps to gradually track students' performance and learning patterns incessantly. As E-2 believes, such continuous assessment enables the adaptation of teaching strategies to optimise learning outcomes, further bolstering the effectiveness of personalised adaptive learning.

The second theme, automated grading, underscores the benefits of HCAI in teaching, learning and assessment tasks. AIS-1 stated that HCAI can streamline grading, thereby reducing the burden on educators while ensuring consistent and unbiased evaluations. Additionally, E-1 argued that automated grading facilitates immediate feedback for students, encouraging continuous learning and timely enhancements in understanding. AIS-1 also noted the role of HCAI in producing data-driven insights through automated grading. Such insights, he suggested, can shed light on students' learning trajectories, thereby providing useful information for educators to refine their teaching methods. As E-2



believes, these insights can result in better educational outcomes, demonstrating the transformative potential of automated grading through HCAI.

The third theme delved into understanding and trust in HCAI systems. AIS-2 stated that trust in these systems necessitates transparency, allowing users to comprehend the decision-making processes employed by AI within the educational domain. Moreover, E-2 argued that trust is cultivated when HCAI systems carry out tasks reliably, accurately, and effectively, thus improving educational results. E-2 noted the ethical aspects of trustworthy HCAI, which include adherence to ethical guidelines, respect for privacy, and unbiased, secure, and fair decision-making. E-1 believes that these considerations are fundamental to the establishment and maintenance of trust in HCAI systems in education, underscoring the importance of these principles in the successful integration of AI in educational settings.

The fourth theme highlighted the role of HCAI in fostering accessibility and inclusivity in learning environments. AIS-2 stated that HCAI has the potential to provide assistive technologies that support learners of varied abilities, thereby promoting inclusivity in education. In agreement, AIS-1 argued that HCAI can also facilitate learning beyond traditional physical classrooms, thereby providing educational opportunities to students irrespective of their geographical locations. In addition, AIS-2 noted that HCAI can address language barriers by providing translation services and language support, thereby making learning materials accessible to non-native speakers or learners with different language proficiencies. Both E-1 and E-2 believed that these features of HCAI not only enhance the inclusivity of education but also contribute significantly to its accessibility for a diverse range of learners.

6. Conclusion

The design, development, and implementation of AI tools in and for education necessitate that these are in turn grounded in pedagogical underpinnings. The findings of this study suggest that AI already plays and will keep on playing a major role in underpinning the future of education. Consequently, it must be rooted in a human-centric framework prioritising human well-being, capability and ethics, underpinned by pedagogy-driven AI design, educator training and policy development. Such an endeavour requires a collaborative, interdisciplinary approach, with a specific focus on engaging educators in the design process. This ensures that these AI solutions are tailored to meet teachers' and students' needs and enhance the pedagogical process. Consequently, the successful integration of HCAI into educational settings demands robust training for educators. Such professional development equips teachers with the requisite skills and knowledge to effectively harness the potential of HCAI principles, thereby facilitating enhanced pedagogical outcomes. As such, it is essential to establish clear-cut policies and regulatory frameworks to guide the use of HCAI in educational contexts. These should encompass regulations around data collection, utilisation, privacy, accountability for AI-based decisions, and guidelines to prevent bias. Moreover, policies should outline ethical considerations and ensure equitable access to AI tools, taking into account the need for the requisite technological infrastructure and resources to support the successful implementation of HCAI. Thus, with careful planning and strategic implementation, HCAI has the potential to transform the educational landscape, promoting inclusive, effective, and engaging human-centred teaching and learning experiences.

References

- [1] Strong, A. I. (2016). Applications of artificial intelligence & associated technologies. Science [ETEBMS-2016], 5(6).
- [2] Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. IEEE Access, 8, 75264-75278.
- [3] Lepri, B., Oliver, N., & Pentland, A. (2021). Ethical machines: The Human-Centric use of Artificial Intelligence. IScience, 24(3), 102249.
- [4] UNESCO. (2023). Artificial Intelligence in Education. UNESCO. Available at: <https://www.unesco.org/en/digital-education/artificial-intelligence> (Accessed: June 8, 2023)
- [5] Butt, M. A., Qayyum, A., Ali, H., Al-Fuqaha, A., & Qadir, J. (2023). Towards secure private and trustworthy human-centric embedded machine learning: An emotion-aware facial recognition case study. Computers & Security, 125,103058.
- [6] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.