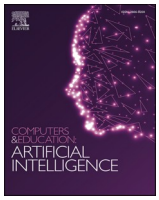
Computers and Education: Artificial Intelligence 4 (2023) 100130

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence



Empowering learners for the age of artificial intelligence

A B S T R A C T

Although studied for decades by the research community, artificial intelligence (AI) in education has recently sparked much public debate with the wide-spread popularity of systems such as ChatGPT and DALL-E. Existing literature offers a wealth of research on design, deploying and evaluating AI-driven systems in education. However, the challenges related to the growing influence of AI in the society, calls for revisiting research foundations of AI in education in order to inform decision-making in policy and guide future research. This special issue of *Computers & Education: Artificial Intelligence* brings 11 papers that explore the theme of empowering learners for the age of AI. The contributions of these papers can be broadly grouped into seven main themes: *intersection between AI and humans* that looks at the space of coordination; *assessment* that explores challenges and opportunities afforded by the use of AI in educational assessment; *explainability* in *AI* as a critical need for humans in education to understand and trust AI; *design for learning* that offers principles for designing AI-driven systems and educational opportunities; *conceptual AI and learning* exploring the need for the development of new theories of learning and their connections with existing theoretical foundations in education; *accurate predictions* and their role in future education; and *applications* of AI in classrooms and educational systems. The findings of these studies highlight pressing research and policies challenges and opportunities that arise with the broad penetration of AI in education. They also emphasize the need for future research that addresses issues of *ethics, bias and farness* in the use of AI in education; challenges associated with *data sources and ownership* as the key fuel and enabler of present-day AI generation; *AI literacies and competencies* of stakeholders who use and are impacted by AI in education; identification of *effective learning and teaching practices* with the use of AI; and policy development to increase *responsiveness of education systems* to rapid changes driven by AI.

Since AI's founding, researchers, theorists, and pundits have overstated how it would impact humanity. While periodic moments of dramatic achievement have occurred, such as DeepBlue and AlphaGo, most progress with AI has been behind the scenes and in labs. Public declarations of AI's coming dominance have continually failed to materialize, or when they do, to materialize much later than forecast. By late 2022, after a year of rapid developments with generative AI (technologies such as text to image, large language models that produce reasonably coherent text with only short prompts, and text to video tools) it appears that the early glimpses of how AI might intersect with humans is starting to emerge. Acts of creation are no longer the domain of human endeavor as AI produces novel text and images through GPT-3 and Stable Diffusion. Whether the *appearance* of creativity is an indication of *actual* creativity remains the domain of philosophers and cognitive scientists as they explore theory of mind, states of knowing, and even consciousness. For the general public, these advances of AI raise tantalizing prospects of future human-AI intersections in a range of creative and knowledge work, but also worrying concerns about bias, ethics, fairness, and accuracy. As critically are questions regarding the longer term role for humans to interact with AI systems (Johnson et al., 2013).

For the education sector specifically, consequential questions have emerged regarding innovations such as essay writing technologies. Notably - how should learners be assessed when a paper can be created of plausible writing quality in a matter of seconds with only a few strategic prompts. A new concern arises: how do researchers and educators define and articulate the *space of negotiation* between what AI is and does in learning and educational settings and how that relates to human learning and cognition.

This special issue is a substantive engagement around how learning will be impacted and the types of challenges the entire education sector faces. It is important to note that while AI has now burst out of university and corporate labs, even dedicated researchers have been taken aback by the apparent sophistication of large language models like ChatGPT. In late November 2022, prominent AI researchers shared their surprise at both the quality of ChatGPT's outputs and its periodic comically nonsensical responses. Education is concerned with preparing people to navigate complex futures. Going forward, this will involve active engagement and collaboration with AI. Significant challenges exist in laying core foundational theories of human and machine learning intersections, including machine-machine intelligence (Rahawan, 2019), and how to apply those frameworks into education systems.

This special issue has been in development for over 18 months and reflects leading thinking about AI and learning from scholars around the world. While it is a substantial contribution, it is part of what will become a crescendoing conversation regarding what it means to learn with AI. This conversation is now happening in labs and public spaces around the world and including the broad representation from all regions, sectors, and societies is vital to ensuring the future of education remains focused on the human student. Our theme for this issue is *Empowering Learners for the Age of Artificial Intelligence* - a focus on the need for AI in education to improve the experience of learners and the impact of their education.

Our special issue includes eleven articles and introduces surprising synthesis around a range of themes. Each of these papers contributed to

https://doi.org/10.1016/j.caeai.2023.100130

Available online 3 February 2023

a number of themes, so our summary below is intended to identify those areas where the contribution was to a primary theme. These themes are best seen as a network, rather than silos. We identified seven themes: intersection between human and machine, assessment, explainability and ethics, design, conceptual, predictions, and application.

The first and most prominent theme centres on what we describe as "the space of coordination" - the **intersection between AI and humans**. What is being coordinated in this space? Currently, almost everything related to learning is under examination. Molenar (2022) focuses on tasks that can be "offloaded from humans to AI" (p. 1) and then "onloaded from AI to humans" (p.1), focusing specifically on a hybrid systems to help develop self regulated learning skills in young learners. She offers a hybrid model of regulation, emphasizing four regulation degrees: AI, co-regulation, shared-regulation, and self-regulation with an intent to support SRL developing through interaction with AI. Markauskaite et al. (2022) explore this middle space by engaging a number of experts in the learning analytics and AI fields to articulate some direction of an "interconnected, fast changing world" where "agentic machines" (p. 1) are the norm. Their Table 1 captures the range of perspectives regarding what capabilities are needed, how we develop those capabilities in humans, and how we study the development of those capabilities. Siemens et al. (2022) suggest that intelligence is too broad a term to provide utility when assessing which specific tasks or activities should be conducted by human or machine. Instead, they argue that cognition and discrete cognitive activities should be the foundation for evaluating and assessing how people learn, sensemake, and make decisions when agents with cognitive capacity are present.

Assessment is another prominent theme, not only in this special issue, but a growing interest in large language models (Floridi & Chiriatti, 2020) and their ability to assist students in auto-generating essays. At the time of writing this editorial, there is wide ranging conversation in newspapers, news programs, and conferences regarding the utility of essays in colleges given the performance of ChatGPT. Swiecki et al. (2022) draw a contrast between standard assessment paradigm - where "a predefined set of items … is used to infer claims about learner's proficiency in one or more traits" (p.1) - and the emerging AI for assessment where tasks can be generated, peers can be recommended to grade work, and grading itself can be done automatically. They suggest that a number of inefficiencies in the existing model can be addressed in this new approach, including a move towards continuous, more authentic, and adaptive assessment. However, a number of challenges remain in this assessment model, including "sidelining professional expertise", "black-boxing" accountability by placing decisions in the hands of programmers, restricting the role that pedagogy plays in assessment, limiting accountability and the scope of learning, and surveillance pedagogy. Despite these substantial challenges, the authors remain optimistic that AI can at least partially address the limitations of the existing assessment model.

The third major theme of this issue relates to **explainability of AI**. Explainability refers to the ability of humans to understand and trust AI. This need will increase in importance, especially as the intersections between human and AI require greater collaboration and coordination (see Carvalho et al., 2022; Markauskaite et al., 2022; Molenar, 2022; Siemens et al., 2022). Khosravi et al. (2022) introduce the XAI-ED, detailing how existing explainable AI approaches can be applied to the education sector. XAI-ED incorporates insight from numerous fields, including cognitive and learning sciences, AI, human computer interaction, and learning analytics in order to present a "means for educational tool developers and researchers" (p. 4) with the means to support the "creation of trustworthy, AI-augmented, sociotechnical systems" (p. 18). They then present four case studies that detail how XAI-ED is reflected in different AIED systems: learner sourced, open-ended learning environments, writing analytics, and team-based learning. They conclude by advocating for broader discussion around explainable AI in education, increased research and adoption of these models, and involvement of multiple stakeholders in developing tools and systems.

In digital education, thoughtful and intentional design is critical to providing learners with the structure and support required during the learning process. **Design for learning** figures prominently in our special issue as well. Kay et al. (2022) promote design of learning data as part of the overall design process. Open Learner Models (OLMs) represent what learners know and the skills they possess. While an OLM can be seen from a range of perspectives, including the teacher, the implicit model in a system, the model of a set of learners, Kay et al. focus on the individual model of learners. As such, OLMs capture individual student progress and can serve as a progress indicator. Their approach to OLM involves adopting core concepts from AI in education (such as scrutability and model development) and applying this to learning design process, supported by data design. Carvahlo et al. (2022) take a broader view by focusing on learning in an AI world. They detail design implications, notably around defining the "*problem space* for education design in a world of AI" (p. 3) and advancing pedagogies for unknown futures. They recommend co-design approaches with humans in design in general, but when applied to AI settings, they conclude by articulating the need for teaming up with AI in various learning settings. This raises an unaddressed need in education today: design models for human-AI interactions have not yet been developed that provide guidance on how AI supports, augments, and directs learning.

Conceptual AI and Learning is the fifth theme of this issue. Do existing theories of learning (and related areas such as sensemaking, decision making, and self regulation) require new theoretical constructs? Can constructivism suitably capture the learning and design needs of AI-human learning interactions? All articles confronted this at a basic level. Yazdanian et al. question the implications of AI on systems, Poquet et al. (2021) on the experiences of various learning transitions, Siemens et al. (2022) on the relationship learner cognitive have with AI, Khosravi et al. (2022) on explainabilty, Molenar (2022) on self-regulation, and Kay et al. (2022) on learner models. Taken as a whole, these papers reveal the need for foundational discussions about learning theory and conceptualizations of learning actions and behaviours in AI-human settings.

A sixth theme relates to the holy grail of humanity: the ability to see into the future and make **accurate predictions**. Universities have started grappling with concerns about curriculum relevance, especially in technical fields where new skills and job categories quickly arise. Yazdanian et al. (2022) target this area of prominence given the complexity and pace of change of the modern workforce. They argue that they "are able to predict future emerging skills with good precision" (p. 1) and by doing so, "AI can help enable educational institutions to keep up with rapid changes in the labour market" (p. 9). This analysis can provide universities with a quick response to new trends, but the ability of universities to move at a pace fast enough to answer this call remains uncertain at best. Carvalho et al. (2022) address the design challenges of these complex issues, but it is important to emphasize that universities are competing with large technology providers, such as LinkedIn, who have granular insights into geographic regions, a topic that Yazdanian et al. raise for future work.

A final theme is on the **application** and use of AI. AI is not a future technology. It is currently in use in classrooms and courses around the world. Howard et al. (2022) explore educational data journeys. A data journey is presented as surfacing "how data was produced and used across different sites of practice" (Howard et al., 2022, p. 2). This journey has direct implications for AI and particularly, in the elements that underpin policy, teacher work and activities, literacies, and general educational data work. Data is used to achieve something in educational settings. In schools, this often involves reporting, tracking student progress, pursuing various state-level targets and personalized education. The authors conclude by arguing for future educational data journey research to meet the increased presence of emerging AI technologies. Poquet et al. (2022) offer a practical application where AI addresses the transition spaces that arise as learners move through life. Luckin et al. (2022) introduce *AI readiness*, noting that adoption requires

literacies and skills and an institutional focus on understanding the differences between human and AI.

1. Concerns regarding the future

We also note a series of four concerns that will shape how education systems adopt AI. Advances in automation and computation alone are not sufficient to ensure broad scale adoption. Ethics, data ownership, AI literacies, and systemic resistance to change are all areas that require greater scrutiny.

Ethics, bias, and fairness remain central to discussions of AI's growing influence. The data that is used to train AI is generally data that has bias embedded within it or within the process of creating algorithms. This bias then produces results that can be problematic over even harmful to certain populations. While responses such as explainable AI are possible avenues to improving the fairness of algorithms, important insights may be gleaned by adopting the methods of cognitive scientists who are used to working with the "black box" of the human brain (see Taylor & Taylor, 2021).

The topic of **data sources and ownership** was not prominent in the articles in this special issue, but as both datasets and models increase in size and the related computation needs increase dramatically beyond the scope of what an academic lab can afford, the presence of big technology companies becomes increasingly important. Many of the datasets used today to train models are open. However, data sets comparable to ImageNet or LAION to build learning models do not yet exist. This is partially due to the challenges of multi-faceted needs for creating even a rudimentary model of learner behaviour or knowledge, in contrast with ImageNet, where variables are constrained. The organizations that have large enough datasets to begin creating learner models are often private or for-profit (Microsoft, Google, Instructure). For educators, addressing a range of concerns around AI in education - including bias, ethics, and fairness issues - will require generating large open datasets for developing, training and validating models.

With the sudden public interest in AI, driven by the success of large language models, the need for **literacies and competencies** of all faculty, teachers, and learners in higher education becomes apparent (Long & Magerko, 2020; Ng et al., 2021). What should the general public know about AI? How do general competencies in AI differ from the reskilling of society in computer science literacies of the last several decades? Here vital questions emerge: what should academics (both students and faculty) know about AI and what should the general public know? How will entire sectors of society be reskilled and whose responsibility it is to initiate and support that reskilling? Should AI literacy be a state and national initiative? Or should the public and private education marketplaces be left to address this need? Regardless of how those decisions are made, having a functional understanding of what AI is, what it does, and possible implications on individuals and societies seem like a fundamental and basic need.

Trustworthiness and reliability of AI technologies will remain an open challenge. Many AI technologies that received much prominence recently are grounded in the concept of generative AI that is built on transformer-based architectures for training large language models. Such AI technologies generate exceptionally convincing human-like textual responses across a range of different genres. However, the architectural designs of such technologies do not have notions that can guarantee factual truthfulness and reason over causal and temporal relationships (Marcus & Davis, 2019). They can also produce eloquent responses written in an authoritative style on nonsensical topics (e.g., financial implications of pension plans and aged care for immune cells). Arguably, to address these issues requires a fundamental paradigm shift from contemporary AI technologies (DARPA, 2022; Marcus & Booch, 2023). While we work with the present (generative) incarnation of AI technologies, we see a need to identify ***effective learning and teaching practices*** that will harness the weaknesses of generative AI technologies as opportunities for promoting higher-order learning (e.g., analyze and scrutinize outputs produced by ChatGPT). In this process, we can not expect teachers and educators alone to fix the problem of AI in their classroom, but we need to have an extensive involvement of researchers, technology developers, and policy makers.

In late 2022, as generative AI technologies, including ChatGPT, gained increased media attention, the response by universities and academics is cause for concern. The impact of universities as institutions is measured in centuries and millennia in how humanity's knowledge is discovered and shared. As a result, universities are not measured by their rapid responses to potential trends. While this aspect of higher education systems is to be lauded in ensuring that small, but highly hyped, trends do not overwhelm the lofty long term goals of universities in supporting society and democracies while raising the quality of life for all people, this **slowness of universities and school systems to change** and respond at a systemic level to dramatic and possibly existential trends, and in the process, to conceive new contributions to a society where AI is prominent, is worrying. For university leadership, vital discussions emerge. First, do the early indicators of successful AI approaches in classrooms scale for large numbers of learners and increasingly diverse learner populations? Secondly, how fast should systems respond to AI and how aggressively should systems remake themselves in response to AI?

This special issue represents front line research on dramatic advances over the past several decades. Many AI researchers have seen trends rise and seen the field enter a number of "AI winters" where research funding and progress slow down. The constellation of trends, including computing advancements, scope and quality of data, and advances in algorithms, suggests that the short term future is one of continued advancement. As AI proceeds, however, educators and society in general face a new reality: what will we teach and how will we teach when artificial agents, now readily present in our daily lives, exceed our cognitive capacity in a growing number of domains?

Acknowledgements

This special issue was in part supported by funding from Australian Research Council (DP210100060, DP220101209), Economic and Social Research Council of the United Kingdom (ES/S015701/1), and Jacobs Foundation (CELLA 2 CERES) awarded to the first author. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Australian Research Council, Economic and Social Research Council of the United Kingdom, or Jacobs Foundation.

References

Carvalho, L., Martinez-Maldonado, R., Tsai, Y. S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence, 3*, Article 100053.

DARPA. (2022). DARPA's ANSR to improving trustworthy AI. *Defense Advanced Research Project Agency*. https://www.darpa.mil/news-events/2022-06-03.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*(4), 681–694.

Howard, S. K., Swist, T., Gasevic, D., Bartimote, K., Knight, S., Gulson, K., & Selwyn, N. (2022). Educational data journeys: Where are we going, what are we taking and making for AI? *Computers and Education: Artificial Intelligence, 3*, Article 100073.

Johnson, N., Zhao, G., Hunsader, E., Qi, H., Johnson, N., Meng, J., & Tivnan, B. (2013). Abrupt rise of new machine ecology beyond human response time. *Scientific Reports, 3*(1), 2627.

Kay, J., Bartimote, K., Kitto, K., Kummerfeld, B., Liu, D., & Reimann, P. (2022). Enhancing learning by Open Learner Model (OLM) driven data design. *Computers and Education: Artificial Intelligence, 3*, Article 100069.

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, Article 100074.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16).

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence, 3*, Article 100076.

Marcus, G., & Booch, G. (2023). AGI will not happen in your lifetime. Or will it? *The Road to AI We Can Trust*. January 23 https://garymarcus.substack.com/p/agi-will-not-happen-in-your-lifetime.
Marcus, G., & Davis, E. (2019). Rebooting AI: Building artificial intelligence we can trust. *Vintage*.
Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence, 3*, Article 100056.
Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence, 3*, Article 100070.
Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041.
Poquet, O., Kitto, K., Jovanovic, J., Dawson, S., Siemens, G., & Markauskaite, L. (2021). Transitions through lifelong learning: Implications for learning analytics. *Computers and Education: Artificial Intelligence, 2*, Article 100039.
Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., & Wellman, M. (2019). Machine behaviour. *Nature, 568*(7753), 477–486.
Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S., & de Laat, M. (2022). Human and artificial cognition. *Computers and Education: Artificial Intelligence, 3*, Article 100107.
Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence, 3*, Article 100075.
Taylor, J. E. T., & Taylor, G. W. (2021). Artificial cognition: How experimental psychology can help generate explainable artificial intelligence. *Psychonomic Bulletin & Review, 28*(2), 454–475.
Yazdanian, R., Davis, R. L., Guo, X., Lim, F., Dillenbourg, P., & Kan, M. Y. (2022). On the radar: Predicting near-future surges in skills' hiring demand to provide early warning to educators. *Computers and Education: Artificial Intelligence, 3*, Article 100043.

Dragan Gašević[*,1]
Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia

George Siemens[1]
Centre for Change and Complexity in Learning, University of South Australia, Australia
E-mail address: George.Siemens@unisa.edu.au.

Shazia Sadiq[1]
School of Information Technology and Electrical Engineering, The University of Queensland, Australia
E-mail address: shazia@itee.uq.edu.au.

[*] Corresponding author.
E-mail address: dragan.gasevi@monash.edu (D. Gašević).

[1] All the editors had the equal contributions in the preparation of this special issue.