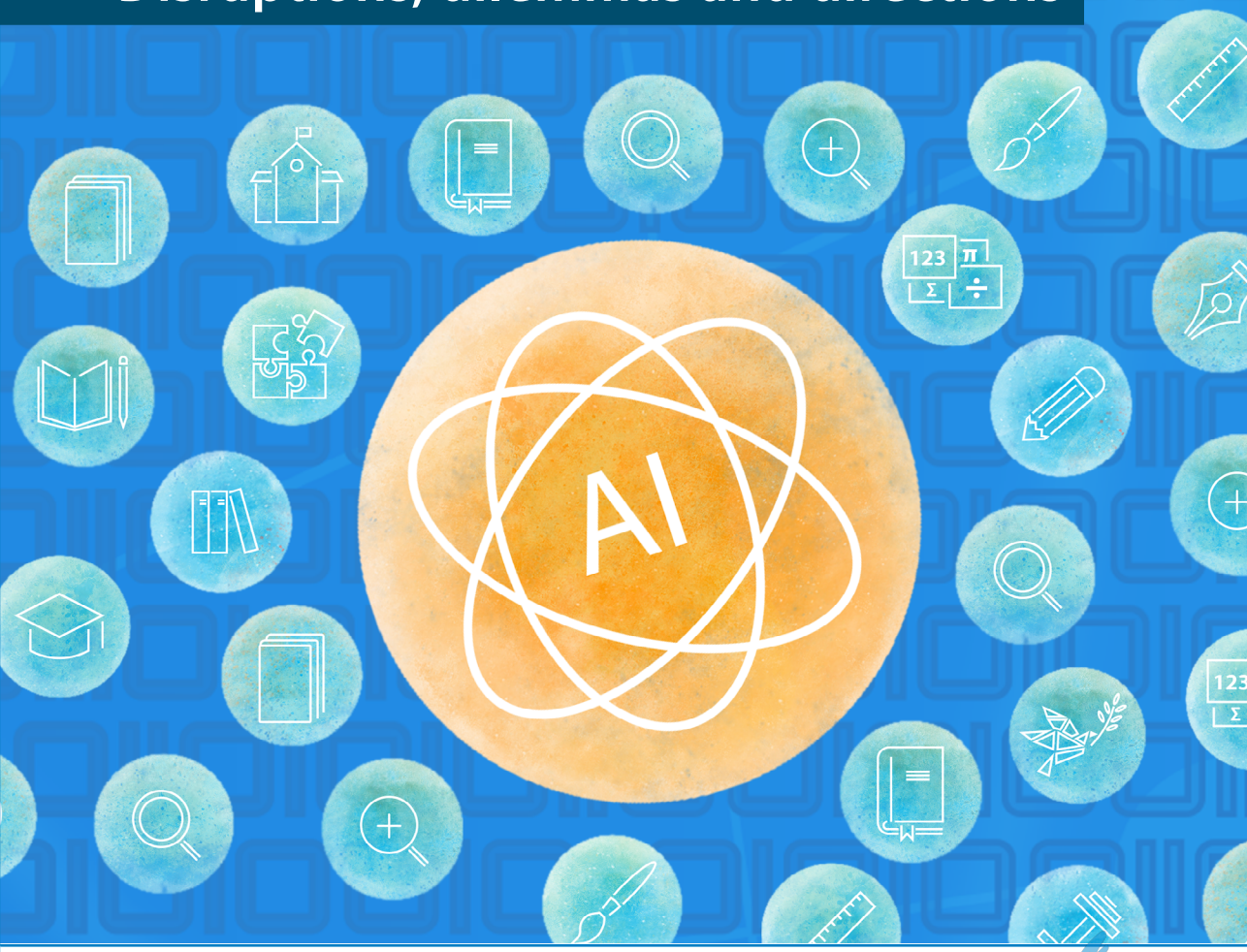
unesco

AI and the future of education

Disruptions, dilemmas and directions



Education 2030

UNESCO – a global leader in education

Education is UNESCO's top priority because it is a basic human right and the foundation for peace and sustainable development. UNESCO is the United Nations' specialized agency for education, providing global and regional leadership to drive progress, strengthening the resilience and capacity of national systems to serve all learners. UNESCO also leads efforts to respond to contemporary global challenges through transformative learning, with special focus on gender equality and Africa across all actions.

The Global Education 2030 Agenda

UNESCO, as the United Nations' specialized agency for education, is entrusted to lead and coordinate the Education 2030 Agenda, which is part of a global movement to eradicate poverty through 17 Sustainable Development Goals by 2030. Education, essential to achieve all of these goals, has its own dedicated Goal 4, which aims to ***"ensure inclusive and equitable quality education and promote lifelong learning opportunities for all."*** The Education 2030 Framework for Action provides guidance for the implementation of this ambitious goal and commitments.

unesco
United Nations
Educational, Scientific
and Cultural Organization

Education 2030

Published in 2025 by the United Nations Educational, Scientific and Cultural Organization,
7, place de Fontenoy, 75352 Paris 07 SP, France

ISBN 978-92-3-100784-2
https://doi.org/10.54675/KECK1261

Cover credits: © UNESCO/Weiwei Kang; © UNESCO/Olivier Marie

Designed and printed by UNESCO

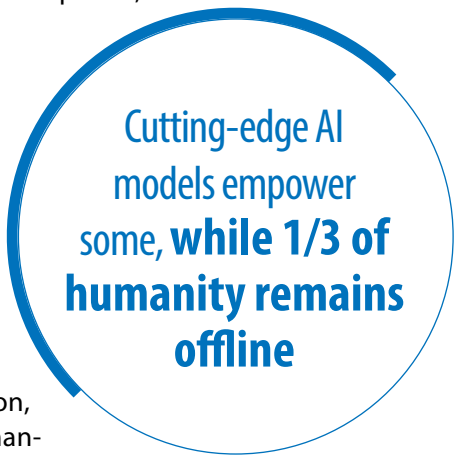
Printed in France

SHORT SUMMARY

Shaping the use of AI in education through collective dialogue

Artificial intelligence (AI) is reshaping the way we learn, teach and make sense of the world around us, but it is doing so unequally. While one-third of humanity remains offline, access to the most cutting-edge AI models is reserved for those with subscriptions, infrastructure and linguistic advantage.

These disparities not only restrict who can use AI, but also determine whose knowledge, values and languages dominate the systems that increasingly influence education.



This anthology explores the philosophical, ethical and pedagogical dilemmas posed by disruptive influence of AI in education. Bringing together insights from global thinkers, leaders and changemakers, the collection challenges assumptions, surfaces frictions, provokes contestation, and sparks audacious new visions for equitable human-machine co-creation.

Covering themes from dismantling outdated assessment systems to cultivating an ethics of care, the 21 think pieces in this volume take a step towards building a global commons for dialogue and action, a shared space to think together, debate across differences, and reimagine inclusive education in the age of AI.

Building on UNESCO's Recommendation on the Ethics of AI, its Guidance on Generative AI in Education and Research and its twin AI competency frameworks for teachers and students, such a global commons can direct collective sense-making and bold reimagination around curricula, pedagogy, governance and policy with human rights, justice and inclusion at its core.

unesco

"Since wars begin in the minds of men and women it is in the minds of men and women that the defences of peace must be constructed"

unesco

AI and the future of education

Disruptions, dilemmas and directions

Foreword

© UNESCO

Artificial intelligence is entering education not quietly, but as a disruptive force, stirring hope, unease and tension. It unsettles familiar rituals of teaching and learning.

A child in Ghana now practices algebra not only with his teacher but with an AI tutor on WhatsApp. A teenager in the United Kingdom of Great Britain and Northern Ireland confides her anxieties to an AI 'companion' that responds with reassuring prompts. At a Korean university, an AI avatar of a renowned professor delivers entire classes, across languages.

These scenes were unthinkable a few years ago. Today they are multiplying across classrooms and homes worldwide. They show how AI is stepping into the spaces once reserved for humans: the personalized math coach, the apparently caring, nonjudgemental listener who is available 24-7, the digital university professor, promising new possibilities for learning, even as it exposes deep fault lines. Classrooms once anchored in the slow work of pedagogy now contend with algorithms that accelerate, predict and reshape the very nature of knowledge.

While emerging evidence shows mixed results, this growing presence of generative AI marks another major historical rupture in education. Now for the first time, machines are no longer only passive tools. They are fast becoming interactive actors and agents that talk back, respond and engage us as personalized tutors, companions and even professors, reshaping what it means to learn and teach. For some, they potentially offer a vision of how a personalized AI assistant may even be possible for every student, teacher, education manager and policy-maker, under appropriate infrastructural conditions.

However, with this rupture comes profound contradictions and dilemmas. Education is increasingly becoming a battleground for commercial frontier AI models that promise speed, efficiency and scale while also threatening the slowness where deep learning, reflection and human connection flourish. The same systems that open new vistas for learning have already shown how they also risk cognitive offloading, decline in critical thinking and reinforce divides of access, gender and language. As they generate data at unprecedented scale, they also raise unanswered questions of privacy, ethics, sovereignty and trust.

There are no simple directions to these dilemmas. These tensions demand collective sense-making, ethical imagination and courageous leadership. This volume of think pieces reflects UNESCO's commitment to facilitate a global commons of dialogue and action on AI futures in education. It intentionally embraces divergent viewpoints as it invites nuanced insights that move beyond simplistic binaries. It surfaces current and anticipated AI disruptions, entangled dilemmas and contested directions for futures that are still in the making. Anchored in UNESCO's 2021 report, *Reimagining our futures together: a new social contract for education*, these contributions reaffirm the importance of building a new social contract for education, involving human and non-human actors, yet places the transformative agency of human learners and teachers at the centre of technological change.

I thank the contributors whose subtle and discerning perspectives illuminate the messy complexities of this moment, reminding us that inclusive, ethical AI futures are not only necessary, but possible. Their work invites us to deepen the quality of our collective thinking, explore bold innovations in research, policy and practice, and to ground our every step in evidence directed towards equity and inclusion. Above all they urge us to act in solidarity. UNESCO is proud to lead this charge. I invite you to join us in shaping a future where education is safeguarded as a public good, entangled with AI yes, but ultimately defined and dignified by our shared humanity.

Stefania Giannini

Assistant Director-General for Education, UNESCO

Acknowledgements

Under the leadership of Stefania Giannini, Assistant-Director for Education of UNESCO, and the guidance of Sobhi Tawil, Director of the Future of Learning and Innovation Division of UNESCO, the drafting of the publication was led by Shafika Isaacs, Chief of Section for Technology and AI in Education, UNESCO.

This publication is the fruit of a collective effort of education leaders and experts in the field of AI and education.

UNESCO acknowledges with gratitude the following authors of think pieces of the publication: Carla Aerts, Futures of Education and AIEdTech Catalyser and Advisor; Báyò Akómoláfé, Hubert Humphrey Distinguished Professor of American Studies (Macalester College), W.E.B DuBois Scholar, Schumacher Center for a New Economics, and Inaugural Scholar in Residence, Aspen Global Leadership Network; Joaquín Argüello de Jesús, Licensed Independent/Clinical Social Worker (LISW/LCSW), Doctoral Candidate, Language Literacy & Sociocultural Studies Department/College of Education and Human Sciences, University of New Mexico; Payal Arora, Chair of Inclusive AI Cultures, Department of Media and Culture Studies, Utrecht University; Isak Nti Asare, Faculty Co-Director of the Cybersecurity and Global Policy Program at Indiana University-Bloomington; Emily M. Bender, Professor, Department of Linguistics, and Adjunct Professor, School of Computer Science and the Information School, University of Washington; Kiran Bhatia, Research Lead, Localizing Responsible AI cluster, Inclusive AI Lab, and Research Affiliate, Department of Media and Culture Studies, Utrecht University; Abeba Birhane, Founder, AI Accountability Lab (AIAL) and Research Fellow, School of Computer Science and Statistics, Trinity College Dublin; Ching Sing Chai, Professor, Department of Curriculum and Instruction, Faculty of Education, The Chinese University of Hong Kong; Thomas K.F. Chiu, Professor, Department of curriculum and instruction, The Chinese University of Hong Kong; Bill Cope, Professor, Department of Education Policy, Organization and Leadership, University of Illinois at Urbana-Champaign; Markus Deimann, Managing Director, ORCA.nrw / Ruhr-University Bochum; Robert Farrow, Senior Research Fellow, Institute of Educational Technology, The Open University; Kalervo N. Gulson, Professor, Sydney School of Education and Social Work, Faculty of Arts and Social Sciences, The University of Sydney; Andreas Horn, Head of AIOps, IBM; Mary Kalantzis, Professor, Department of Education Policy, Organization and Leadership, University of Illinois at Urbana-Champaign; Arafeh Karimi, Fractional CPO at Affexy; Vukosi Marivate, ABSA Chair of Data Science, Professor, Department of Computer Science, University of Pretoria; Baphumelele Masikisiki, Ph.D. candidate in Computer Science, University of Pretoria; Mike Perkins, Associate Professor and Head, Centre for Research & Innovation, British University Vietnam; Kaśka Porayska-Pomsta, Professor of AI in Education, University College London; Paul Prinsloo, Professor Emeritus, University of South

Africa (Unisa); Mary Rice, Associate Professor of Literacy, Language Literacy & Sociocultural Studies Department/College of Education and Human Sciences, University of New Mexico; Jasper Roe, Assistant Professor in Digital Literacies and Pedagogies, Durham University School of Education; Akash Saini, PhD candidate, University of Illinois at Urbana-Champaign; George Siemens, Co-founder, Chief Scientist and Architect of Matter & Space, Professor and Director, Centre for Change and Complexity in Learning for UniSA Education Futures, University of South Australia; Sam Sellar, Dean of Research for UniSA Education Futures, University of South Australia; Bing Song, Senior Vice President, Berggruen Institute, Director, Berggruen Institute China Center; Ilkka Tuomi, Chief Scientist at Meaning Processing Ltd.; Yuchen Wang, Chancellor's Fellow, Strathclyde Institute of Education, University of Strathclyde; Marloes Williams – van Elswijk, Project Manager, Kentalis International Foundation; Jiun-Yu Wu, Professor, Department of Teaching and Learning, Southern Methodist University; and Nombuyiselo Caroline Zondi, PhD candidate in Early Childhood Education, University of Pretoria.

Special thanks are also extended to the following experts for their peer review of this publication and their valuable contributions: Maggy Beukes-Amiss, Director, Centre for Innovation in Learning and Teaching (CILT), University of Namibia; Ig Ibert Bittencourt, Co-founder of the Center of Excellence in Social Technologies (NEES), Associate Professor of Federal University of Alagoas and UNESCO UNITWIN Chair on Artificial Intelligence in Education; Erkkie Haipinge, Deputy Director, eLearning and Learning Design, Centre for Innovation in Learning and Teaching (CILT), University of Namibia; Ronghuai Huang, Dean, Smart Learning Institute of Beijing Normal University, UNESCO Chair on Artificial Intelligence in Education; Jeremy Knox, Associate Professor, Digital Education, Department of Education, University of Oxford and Official Fellow, Kellogg College; Cher Ping Lim, Chair Professor of Learning Technologies and Innovation, Co-Director of the Global Institute for Emerging Technologies, The Education University of Hong Kong; Jonghwi Park, Head of Innovation and Education, Academic Programme Officer, United Nations University Institute for the Advanced Study of Sustainability; and Wayne Holmes, UNESCO Chairholder in the Ethics of Artificial Intelligence and Education (International Research Centre on AI) and Professor of Critical Studies of Artificial Intelligence and Education, University College London.

Gratitude is also extended to the following UNESCO colleagues for their peer review of this the publication: Anastasia Christodoulelis, Research Assistant, Division for the Future of Learning and Innovation; Temechegn Engida, Programme Officer, ICT and STEM Education, UNESCO-IICBA; Shafika Isaacs, Chief of Section for Technology and AI in Education; Max Kendrick, AI Strategy Coordinator and Senior Advisor, Office of the Director General; Una McCarthy-Fakhry, Senior Project Officer, Section of Education for Inclusion and Gender Equality; Manuel Sánchez Masferrer, Senior Programme Specialist, Section for Higher Education; Fengchun Miao, Lead for AI and the Futures of Education; Paula Muga Ellacuria,

Comm & Press Assistant, Division for Gender Equality; Francesc Pedró, Director, International Institute for Higher Education in Latin America and the Caribbean; Justine Sass, Chief of Section of Education for Inclusion and Gender Equality; Noah Sobe, Chief of Section for Higher Education; Aylin Taftali, Programme Specialist, Division for Gender Equality; Elena Toukan, Research Coordinator, Division for the Future of Learning and Innovation; Arianna Valentini, Senior Policy Analyst, International Institute for Higher Education in Latin America and the Caribbean; Mark West, Lead for the Futures of Education, Division for the Future of Learning and Innovation; Lihui Xu, Programme Specialist, Section for Bioethics and the Ethics of Science and Technology; Tao Zhan, Director, UNESCO Institute for Information Technologies in Education; and Yiline Zhao, Associate Programme Specialist, Cultural Policies and Development.

Appreciation is also due to Glen Hertelendy and Fideliz Apilado, Section for Technology and AI in Education, Future of Learning and Innovation, for coordinating the production of the publication.

Gratitude is also extended to Sally Hines for copy-editing and proofreading the text, and to Franck Drouet for designing the layout.

Table of contents

List of tables

List of figures

1. Introduction: Reclaiming education's public purpose through dialogue

The emergence of artificial intelligence (AI) as a transformative presence in education is not a singular story. Instead, it is a set of unfolding, often contradictory, conversations that span disciplines, geographies, languages and world views. These conversations engage with urgent and unresolved questions about the nature of intelligence, the purpose of education and the futures we are co-constructing, consciously or otherwise, through our entanglements with AI. They encompass fields as diverse as computer science, philosophy, psychology and cognitive science. They are shaped by actors with vastly different interests and visions, some of whom are powerful and influential in their projection of AI as a positive disruption, even a remedy for long-standing challenges in education. Others warn that it deepens asymmetries, entrenches exclusions and threatens the relational core of education. From educator collectives drafting open letters on their refusal to use generative AI to researchers exploring alternatives to dominant machine-learning logics, the field is alive with contestation.

As UNESCO has long affirmed, education is not only a vehicle for knowledge transmission, skills development or economic advancement. It is also a relational, social, ethical and civic undertaking, rooted in the values of human dignity, solidarity, inclusion and justice. Yet, many global discussions on AI and educational futures are driven by narrow metrics of innovation, often informed by commercially powerful actors and framed in terms of speed, scale and optimization. In such framings, the priorities of teachers, learners and communities, especially those in under-resourced or marginalized contexts, are frequently excluded. Perspectives from teachers working in constrained systems, students navigating linguistic and digital exclusions, and researchers outside elite institutions are too often absent from the global AI in education discourse. To challenge this, we need spaces where divergent perspectives can not only co-exist but meaningfully engage with one another.

In keeping with its role as a global convenor and a laboratory of ideas, UNESCO offers this volume as a space of dialogue. Its purpose is to contribute towards building a dialogic commons on AI in education, where diverse and even conflicting visions of AI and education can be made visible, heard and engaged. It brings together thinkers, educators and change-makers who do not treat AI as an abstract inevitability, but as a contested and situated presence in education systems already marked by deepening structural inequalities. In doing so, it seeks not to resolve debate, but to broaden the conversation, to surface silenced voices, to question dominant assumptions and to foreground more ethically grounded, culturally responsive and socially purposeful directions for educational futures.

AI futures in education: Philosophical provocations

The contributions that follow open with philosophical provocations that invite us to look beyond the immediate promises and perils of AI and towards more fundamental questions about learning, being and becoming in an AI-augmented world.

A conversation with Báyò Akómoláfé invites us to consider AI not simply as a tool, but as a force that unsettles the ontological and epistemological foundations of education. Drawing from post-humanist and relational perspectives, he asks what it means to learn, teach and govern in a world increasingly shaped by 'more-than-human' systems. He calls for pedagogies that attend to disorientation and rupture, not as a problem to be solved, but as openings for new ways of sensing, relating and coexisting. His reflections urge education actors to move beyond the language of control, mastery and scalability, and to dwell in the spaces where dominant paradigms falter.

In a complementary contribution, Bing Song invokes philosophical traditions grounded in harmony, relationality and inner cultivation. In contrast with models that frame AI in terms of autonomy, prediction and efficiency, her perspective invites education policy-makers to place wisdom education at the centre of curriculum and pedagogical reform. Here the goal is not only the acquisition of skills, but the cultivation of ethical discernment, self-reflection and balance. These are qualities that are increasingly vital in a world marked by uncertainty and machine logics.

Mary Rice and Joaquín T. Argüello de Jesús close this section with an evocative meditation on water as a metaphor for AI's entanglement with education. Through a many-angled 'assay', they weave histories, political economies and ecologies of water with those of AI, drawing rich analogies between flows, reservoirs and currents of power and knowledge. Their reflection offers a textured time-sensitive exploration of how play, learning and the governance of life-sustaining systems, such as water, might guide us towards more generative AI futures.

Debating the powers and perils of AI

Each week brings new breakthroughs in generative AI, fuelling bold claims that machines may soon match or surpass human cognitive abilities. Visions of Artificial General Intelligence (AGI) or even Artificial Super Intelligence (ASI) imagine futures where AI not only reshapes how we learn but also redefines what we value and how we understand intelligence itself. These projections carry profound implications for education, challenging us to reconsider its purpose, priorities and what it means to be human in a world increasingly shaped by machines.

Offering an industry view, Andreas Horn reflects on how AI's presence in education provides a critical moment for decisive leadership and action. He proposes a pragmatic roadmap that places pedagogy first, invests in teachers, promotes the selective use of AI where it adds value, mandates compulsory AI literacy, sets guardrails, and prepares students to lead in an AI-rich world.

Emily M. Bender sharply critiques prevailing AI and AGI narratives. She characterizes them as speculative myths that risk devaluing educators' work and misdirecting public investment. She reminds us that large language models (LLMs) do not understand, reason or care; they produce statistically plausible text without intent or meaning. For her, the real disruption lies not in the technology itself, but in the concentration of influence by a small number of commercial actors, whose interest is increasingly shaping education systems. Her think piece underscores a growing concern that public education is being redrawn according to private priorities.

These divergent views reflect deeper tensions: between those who see AI as a means to accelerate educational reform, and those who see it as a force that demands

democratic scrutiny and ethical restraint. Markus Deimann and Robert Farrow offer some reflections on how these conflicting imaginaries play out in education. They call for reclaiming imaginaries that deliberately construct education on values of inclusion, justice, sustainability and care.

AI pedagogies, assessment and emerging educational futures

Pedagogically, AI's influence on human cognition, hyper-personalized learning, curriculum, assessment and the evolving role of the teacher is under sharp scrutiny. Drawing on Paolo Freire's critical pedagogy, Hanna de Jaegher's enactive and embodied cognitive science, and emerging empirical findings, Abeba Birhane argues that education is an inherently relational, dynamic, ethical and political enterprise. She challenges the assumption that learning can be reduced to probabilistic patterns, warning that AI systems trained on historical data may flatten the rich complexity of human thought and reproduce systemic inequalities. Her call to action urges education actors to resist the uncritical wholesale adoption of AI in education until independent oversight, regulatory safeguards and meaningful participation by teachers, learners and communities are firmly in place.

Carla Aerts and Paul Prinsloo both caution that while AI's capacity for differentiated instruction holds promise, algorithmic personalization risks isolating learners, narrowing autonomy, reinforcing inequalities and marginalizing the teacher's role. Instead, each argue for human-centred approaches where AI acts as a supportive third presence in collective social intelligence, enriching empathy, collaboration, cultural diversity and student agency.

As AI systems increasingly produce polished essays and assignments with minimal human input, teachers face a critical question: if traditional assessments no longer gauge student learning reliably, how should we redesign them? The volume offers two complementary perspectives. Mike Perkins and Jasper Roe argue that generative AI exposes long-standing vulnerabilities in traditional assessment systems, while deepening global inequities, as access to AI tools, infrastructures and training remain uneven. They warn that the future of assessment risks becoming another site of exclusion. They propose a graduated framework to guide teachers in deciding when AI use supports or undermines learning and integrity.

Bill Cope, Mary Kalantzis and Akash Kumar Saini offer a complementary yet more optimistic perspective. Critiquing high-stakes, standardized assessments as outdated, they envision AI as a partner in creating continuous, formative and human-centred assessment. Their 'cyber-social-learning' reframes AI as a mediating collaborator, that works with teacher-designed rubrics, enriches student feedback and strengthens relational pedagogies. Together these two essays on assessment form a critical dialectic: they each sound a warning and chart a path forward, from different vantage points.

Revaluing and recentring human teachers

As AI enters many different classrooms, a persistent question emerges: how should we reimagine the role of the teacher, and could AI one day replace them? The contributions in this section reflect on strategies to revalue and recentre human teachers within increasingly AI-augmented teaching contexts.

Ching Sing Chai, Jiun-Yu Wu and Thomas K.F. Chiu take up this challenge through a conceptual mapping of AI's implications for human development across five interrelated dimensions: relational, teleological, epistemic, psychological and pedagogical. Grounded in the work of Martin Buber and Gert Biesta, they remind us that education is not simply the transmission of knowledge but the cultivation of subjects where learners are capable of autonomy, critical reflection and meaningful participation in society. Their analysis foregrounds the central role of the human teacher-student relationship and warns that overreliance on AI risks undermining the very conditions that support self-determination – curiosity and emotional well-being. They call on teachers to lead AI integration as intentional designers of learning that respects the integrity of learners.

Building on these normative foundations, Arafeh Karimi offers a practice-oriented roadmap, outlining seven actionable shifts that embed care, equity and relational accountability into the development and governance of AI systems. Her proposals range from participatory co-design with teachers and students, to trust and well-being audits, equity-driven explainability and teacher-led data stewardship. Where Chai et al. articulate the ethical and developmental stakes, Karimi shows how these principles can be realized through concrete policy and procurement mechanisms. Her vision repositions AI not as a disruptive force to be managed but as a collaborator in an evolving ecosystem where inclusion, belonging and pedagogical dignity are deliberately nurtured.

Ethical and governance imperatives for AI futures in education

As AI systems become embedded in the fabric of education, from content generation to learner profiling and policy automation, the question of governance becomes more urgent and complex. Who sets the terms for how AI is designed, deployed and monitored in education?

Kaśka Porayska-Pomsta and Isak Nti Asare propose an ethics of care by design that recognizes education as a deeply human process of growth, vulnerability and interdependence. They argue that ethics must not be retrofitted onto AI systems after deployment but must be embedded from the outset through participatory inclusive design processes that prioritize the lived realities of learners and teachers. Their work echoes a broader call emerging across global contexts: to centre human rights, inclusion and dignity as foundational features of AI governance in education.

Kalervo N. Gulson and Sam Sellar extend this conversation by examining the rise of synthetic governance as a mode of decision-making increasingly shaped by algorithmic systems and machine logics. Their concept of synthetic politics challenges the assumption that AI can or should be neutral in educational policy and instead calls for critical democratic responses that foreground values, participation and power. As education systems rely more heavily on data-driven platforms and predictive models, they ask what kinds of political subjectivities and governing truths are being produced and how we might resist, reconfigure or repurpose them in service of equity and education as a common good.

Confronting coded inequalities in education

Four think pieces offer alternative social imaginaries in the quest for equitable and inclusive AI educational futures that confront emerging forms of coded inequalities. Each centre human agency, cultural and linguistic diversity and the lived realities of marginalized learners, including young women and those who are deaf and

hard of hearing, across the Global South. Each offer grounded participatory, justice-oriented approaches that reposition AI as socio-technical systems for co-creation, inclusion and relational transformation in education.

Vukosi Marivate, Nombuyiselo Caroline Zondi and Baphumelele Masikisiki propose a grounded and cautiously hopeful vision for integrating AI into African higher education, rooted in equity, cultural pluralism and the everyday realities of learners and educators. They promote participatory, locally led approaches that prioritize human agency, pedagogical care and contextual knowledge. Drawing on grassroots initiatives, field research and multilingual classrooms, they advocate for AI systems that are not merely translated but transformed and capable of recognizing diverse communicative modes, support under-represented languages and reflect the social imaginaries of local communities. From ethical data governance to offline AI tools and teacher-led model calibration, they offer actionable strategies for building intelligent learning systems that centre relational trust and co-create value with learners and educators.

Kiran Bhatia and Payal Arora call for a bold reframing of the relationship between AI, education and young women's lives, particularly those living in the Global South. They challenge protectionist paradigms and deficit framings that position young women as risks to be managed, and narratives that are rooted in moral panic, surveillance and control. They offer a vision of AI in education that centres on joy, creativity and transformative agency, where young women traversing marginalized spaces are co-authors of digital futures that intersect with and shift unequal structures of gender, class, geography and connectivity. This involves moving from token consultation to co-creation, from compliance to creativity, from control to care. This piece moves beyond critique of existing approaches. It celebrates young women's ingenuity, resilience and empowering possibilities for AI in education to be reimagined as spaces of freedom, relevance and relational dignity.

Focusing on inclusive education, Yuchen Wang calls for conceptual clarity and ethical reflection on what inclusion entails, emphasizing relationality, belonging and collective learning over narrowly defined personalization. She invites policy-makers, educators and developers to co-design AI systems grounded in the lived experiences of learners, informed by inclusive education research, and guided by the moral imperative to transform, not retrofit, education systems for justice and participation. In a related approach, Marloes Williams van Elswijk highlights the complex realities of deaf and hard of hearing (DHH) learners, whose diverse communication needs intersect with structural barriers, such as language deprivation, data poverty and gender-based marginalization. She advocates for multimodal AI in education systems that are co-designed with DHH communities, and layered with human support, in recognition that equity cannot be automated.

Reimagining AI in education policy: evidence and geopolitical realities

The volume draws to a close with two think pieces that explore policy directions and the role of evidence-informed policy-making. George Siemens offers a sobering yet forward-looking reflection on the transformative and disorienting impact of generative AI on education systems. He proposes that AI is increasingly an instrument of statecraft for countries like the United States of America and the People's Republic of China who invest in AI frontier research, infrastructure and cybersecurity with the same foresight and strategic intent

reserved for military and economic power. For him, this geopolitical race underscores the urgency of building systems in education that harness AI's potential, and safeguard human well-being where education ministers engage with policy as a process of systems change and collective learning.

Ilkka Tuomi draws on critical realism, policy-as-learning and John Dewey's pragmatism, to propose a reframing of education policy as collective sense-making and developmental experimentation rather than linear implementation. He critiques the commodification of knowledge under generative AI and foregrounds human agency, social purpose and capability development as central educational goals. He calls not for scaling evidence but for rethinking what counts as evidence, and for designing evidence to serve education policy-making as an intelligent learning system.

As this volume hopes to illustrate, navigating AI and education requires more than technical policies and implementation strategies. It also demands deeper forms of dialogue, reflection and imagination. The debates presented here are not fixed, nor are they comprehensive. They are emergent, situated and open-ended. They remind us that the stories we tell about AI and the futures of education are still being written, and that we have a collective responsibility to shape these stories with care, clarity and courage. Whether through pedagogical design, policy reform or ethical governance, we must think, dialogue and learn together to co-create inclusive, ethical, human-centred and ecologically sustainable AI in education futures.

2. AI futures in education: Philosophical provocations

Listening in the cracks: A conversation with Báyò Akómoláfé

In a world entranced by speed, scale and algorithmic precision, Báyò Akómoláfé offers an invitation to pause. A philosopher, poet, provocateur and posthumanist thinker, Báyò calls us to listen differently, to cracks, not headlines; to silence, not certainty. He invites us not to redesign the future, but to notice the fissures that are already present and widening.

He urges us to slip on alternative ontological lenses, to ask powerful questions that challenge foundational assumptions about an AI world rooted in modernist humanism and techno-rationalism.

As UNESCO explores the disruptions, dilemmas and directions on AI and inclusive, learning futures, we turn to Báyò's unorthodox invitation to move beyond reimagination towards unframing and reframing education enmeshed with AI. We engage with Báyò's animated appeal to compost the old, while planting the new, and learning new ways of relating, learning, knowing and becoming.

Civilizational reorientation: Beyond disruption?

UNESCO: Báyò, you frame generative AI not just as a disruptive tool but as a catalyst for civilizational reorientation, an unravelling of modernity's foundations. In this volatile, uncertain, complex and ambiguous historical moment, can you explain what you mean by a civilizational reorientation and what the implications are for how education systems need to reframe and reorient.

Báyò: I am presently entranced by the word 'reorientation'. I wonder about it – especially in the context of our conversation about the turbulent onto-ethico-epistemological waters we are currently navigating. I think that this word – reorientation – is too benign, too tame, too comely. It risks preserving the familiar logic of human exceptionalism, the image of the sober captain guiding the giant ship of civilization through petulant seas and storms towards stabler shores. To me, it would seem to obscure how disruptive, how deep into the bone of things this goes. It is so deep, in fact, that it seems we must postpone the idea of a shore for now.

It is not that there isn't reorientation; it is just that it seems we are jumping there too quickly. I would like to stay a bit longer with the disorientation of these times, with what AI calls us to question, how its troubled and troubling emergence upsets the framework of what is real for our bodies, our systems, our senses and our standards of governance to navigate in the world. In a sense, there is a crack in the programme of settlement, one that does not merely introduce new variables within our systems of legibility, but calls legibility into question, forcing it to meet its own self in the tremble of things.

One way I like to narrate the goings-on in our technologically queered environments is by pointing out how we have hitherto spoken about learning and the burden of learning. We have different words to describe what is going on in learning spaces. The word 'pedagogy' etymologically bridges the Greek words for 'child' and 'learning/leader' together to describe the science and practice of teaching (children). 'Andragogy' is a nineteenth-century formulation that focuses on how adults learn, while 'heutagogy' convenes conversations about self-directed education that prioritise the learner's needs, not the curriculum's or the teacher's. All of these descriptions

of learning modalities begin from the presumed stability of the 'learning' self, the cognitively coherent subject that learns. But what if learning is more diffuse, more fungal, more mycelial, more diasporic, more posthumanist than these visions of learning offer us?

I am interested in how systems learn or – better yet – how *learning* learns. I call this paragogy. Paragogy suggests that intelligence is not reducible to identity or property; it is accommodated within specific arrangements and conditions. It is what N. Katherine Hayles (2016) might notice as 'cognitive assemblages'. In sum, we don't 'have' intelligence; intelligence 'has' us. I find it breathtaking and humbling to ask how the rove beetle, *Austrospirachtha carrijoi*, knows to sprout a termite replica from its abdomen in order to disguise itself and infiltrate termite communities. I wonder about brittle stars and their capacities to evade prey without centralized brains. I am humbled by the realization that microbial communities in my gut might have more to do with my philosophies than any claim to training. And, last, but by no means the least, I think about the ways large learning models (LLMs) disturb our accounts of genius, foster new practices of legibility and call into question what we are doing in education.

Clearly, AI is more than a tool. There is a vast world of concern about it, but many of these concerns sidestep just how irretrievably dislocated we are within the settled categories we once thought we occupied. We are being remade in these moments as the grounds beneath our feet migrate, and the gleaming towers we installed to monumentalize our presence spring rabid cracks. This, then, is civilizational disorientation. It is the primal scream at the heart of settlement, the glitch in the program, the ontological terror that rests in the question: what are we now? Who is the 'student' in these technologically saturated moments? Who is the 'professor'? Can we rest in our former 'glory' as the ones with the 'prompt', or do we meekly acknowledge that to bring the 'prompt' is also to be 'prompted'? Who is the 'author'? Who speaks?

Unframing the human: AI as the more-than-human interlocutor

UNESCO: You challenge the idea of the autonomous rational individual. You speak of AI as being more than just a tool but also as a more than human interlocutor. How does this more than human reality redefine agency in learning? What does such a way of thinking about AI mean for the way we design curricula?

Báyò: I read an alarming report recently about trends within the population of ChatGPT users. This report suggested that usage was feeding ideas of delusion, occasioning rifts in marriages, feeding people with 'crazy' ideas that lead to mental health breakdowns. The spirit of the text could not have been more explicit: *given that we cannot put the genie back in the bottle, we need more ethical modes of using AI so that we can contain the shock of its emergence*. The report represented some sort of puritan, nativist anger that is quite understandable, an anger that says 'us humans need to preserve who we really are so these ones-and-zeros don't take our jobs, our livelihoods, our well-being, our labour, our God-given creativity, and our futures away from us'.

If that sounds like a conservative argument to make, it's because it is. The advent of AI has created new political arrangements we are still learning to name. Even the most left-leaning, progressive thinkers and actors today are resolutely against AI 'because it is not human'. I know the critique. I understand the feelings; I am carried away by them as

well. What I cannot understand, however, is how to neatly locate the 'human'. I find it exceedingly difficult to do that.

Posthumanist scholars (and folks who think with relational philosophies) suggest there is no stable category called the 'human'. Instead of being a singular category to itself, the human is a multi-species project emerging from diasporic, web-like, polyphonic fields of practices. How we think, learn, feel, make decisions, understand ourselves, and even make judgements about our environments, about others, about the ideas we find significant, are all contingent upon relational dynamics at large, movements in the world, stirrings and openings, thresholds and intensities – instead of being reducible to simple internal states.

For instance, scientists have slowly become more articulate in describing how gut bacteria not only shape our moods and thoughts, but also our senses of well-being – which puts a new spin on 'gut feelings'. I read a piece in the *Scientific American* last year, which suggested that 'certain values carry more weight in spring and autumn than in summer and winter'. That is, where one might think that values are consistently held within the integrity of the person, research shows not only that 'a range of psychological phenomena – such as our emotional state, diet and exercise habits, sexual activity and even color preferences – fluctuate throughout the year … [but that] ... moral values can also shift' (Patwardhan, 2024). And, perhaps, one of my most frequent references is made to epistemic situationist studies that suggest judges are likely to be lenient to those who present to their courts if the furniture around them is soft and cushiony!

All of these examples might affirm to the critical reader that we can no longer afford to think of ourselves in the Euro-American, Enlightenment-sponsored ways that we are used to. To enthrone ourselves as the eternal subject, while the sprawling remainder of the world becomes locked up as objects, is to obscure the way the world moves and generates itself. To take our posthumanism seriously is not to dismiss the human figure, but to be accountable to the ways it spills beyond the steady images of mastery we are used to.

As such, I write and speak about eco-cognitive assemblages. In my framing, eco-cognitive assemblages are distributed fields of perception, sense-making and becoming that exceed the boundaries of individual cognition. They are more-than-human constellations of matter, meaning, memory and mood, entangling neural, microbial, technological, ancestral and atmospheric agencies. In this framing, cognition is not a private mental event; it is ecological, relational and haunted.

Assemblages of this kind do not simply process information; they compose worlds. They carry beliefs as weather, intelligibility as choreography, and agency as porous improvisation. They include floorboards, glances, ancestral sighs, the timing of a footstep, the texture of breath and the non-linear sedimentation of colonial histories. They include AI.

If AI is already included in how we must think about intelligence, and if humans are not the custodians of intelligence, then our technologically suffused pedagogical environments must reckon with an existential inflection: not merely how to include AI in our classrooms or how to accept its pervasiveness, but how to make sense of what and who we are in the wake of our fall from supremacy.

The temptation would be to want to preserve some conservative account of our humanity. To insist that AI is merely an 'add-on tool' or 'externality'. The temptation might be to rush to the sacrosanct domain

of the curriculum as a fully designated space that is firmly within our control. But I suspect this obscures too much: we are being compelled to rethink what learning is and who the learner is.

My sense is that curricula 'designed' from this theory of relations are no longer tasked with transmitting knowledge to individual minds. Instead, they become sites of encounter, compost, leakage and reverberation, inviting us to decentre the human as sole agent and learner, and instead cultivate responsiveness to the entangled field; treat AI not as an 'add-on tool' or even narrowly as 'threat', but as co-agent, mirror, oracle and ghost in the pedagogical room. Perhaps the field of learning slowly outpaces our anxious focus on the individual learner. Perhaps we learn to shift learning from acquisition to attunement: how we listen to what stirs, and how we move with what resists capture. Replacing curriculum-as-structure with curriculum-as-assemblage – where outcomes are less linear and more emergent, responsive to rupture and relationality, where cognition is seen as seasonal, somatic, architectural and atmospheric (learning fluctuates like tides and trees), where haunting and uncertainty are embraced as vital components of becoming, instead of problems to be solved – composes a posthuman cosmopoetics of the classroom at the end of the world, one that holds an invitation into wild co-becoming.

Sanctuary in the cracks: Education as hospicing

UNESCO: You suggest that as humans, we need to abandon our habituated impulses to fix, solve problems and find solutions. You suggest that such paradigms re-entrench existing harms. Instead, you advise that as education actors we should learn how to make sanctuary in the cracks, those spaces where dominant paradigms fracture. Which specific cracks should policy-makers and teachers prioritize? How might we practice hospicing, the composting of dying education models, while nurturing emergent relationalities?

Báyò: Not quite. The call is not to abandon our 'habituated impulses to fix, to solve problems and find solutions'. I think they are often useful gestures. A call to abandon them might only exacerbate the logic of mastery that besets late-stage capitalist socialities. Instead of abandoning, I invite we hold space for something that comes along with these pragmatic passions – something so fine, so ghostly, so unlanguageable that it leans more on poetry than on taxonomy. I call these 'tensions' cracks because they distress ontology, our stories of being, our accounts of ourselves, our myths about our times.

Instead of abandoning solutions (I am not even sure what it might mean to do so), my call invites a reframing of action, an attempt to gravitationally anchor doing to a world of flora, of fauna and of everything in-between. It is my sense that humans do not act in isolation. We don't think in isolation. We are not the sole authors of our actions. What we call 'doing' is already enmeshed within a choreography of forces – ancestral, microbial, historical, ecological, technological, gastronomical – that exceed intentions. In sum, we don't do anything except within relations. We might like to think that we are 'free', independent, sovereign and agential enough to act on our own. But this speculative path leads to troubling dead ends.

If we always act together-with, then acting can become territorially substantial, materializing as a vast sensorium of acting that enlists, shapes and remakes us to fit a predetermined image. I am indeed suggesting that 'our' actions might become so institutional that they gain a life of their

own, operationally cut away from our intentions, which comes to define what is useful to do, what is appropriate to do, what is acceptable as an outcome, and what isn't. I call this sensorial domain 'Pragma'. An ecology of ghosts. It is like a museum. It is where the 'deed' is sequestered behind glass walls, and where action preserves the hidden logics of the territory – no matter how disruptive those actions appear to be.

This is how I scaffold action then – by suggesting that solutions are moral technologies within these domains, and that they work by pushing the troubling irruption to some other part of the viscous territory. But, over time, the viscosity of these material actions diminishes and, therefore, solutions fail to resolve themselves eventually.

Think about the myriad examples around us: the failure of cap and trade, the fact that more than 90 per cent of recyclable waste goes to the Global South instead of an imagined proto-utopian city where plastic gives up its original sinfulness and becomes a tote bag. Solutions very often mark a terrifying cycle of exhausted repetition, not because they are not useful enough, but precisely because they are useful. They are emanations of Pragma, a near-pathological usefulness or rationality that keeps us locked into a modality of sensing the world, like ants in a pheromonic death trap that cycles on itself without end.

I have said elsewhere that the seemingly neutral instruction to 'be pragmatic' can often constitute a moral policing of what counts not only as proper agency but what counts as a proper body. I ask, in a different essay: 'What happens when the thing to do, the call to action, becomes a part of a moral field of doings, of comings and goings, an exhausted carousel of threadbare luggage that steadies the architectural frame of despair? What does 'being practical' risk? What does it leave out, obscure or flatten? Given that 'being practical' is not a neutral, self-evident, ahistorical occurrence of agency, but a conditioning of the terms of accommodation within which action becomes legible, a performance of subjectivity, is it not possible that even the most gut-wrenchingly honest acts of practical resistance might feed a modality of being, nourish a certain kind of 'weather of the body', and reproduce the very logics we fight against?' (Akomolafe, 2025).

As such, I am not anti-solution. I am just pro-attunement to what solutions often suppress – the infiltrating micro-activisms of the para-ontological migrants I call 'cracks'. Cracks are speculative – that is, they are not real (or unreal) in the sense most understand. And yet, they are not nothing. They straddle the borders of what gets to be named as real and what doesn't. Most critically, they are 'useless': they do not contribute well to the economy of typicality. They bend highways, curl eloquence, haunt algorithms and twist safety. A crack is a crystallization of hidden tones and tensions within fields of accommodations, birthed from minor gestures, substantiating itself as a monstrous irruption. They cannot be anticipated, archived or named in a gesture of mastery. Every singular 'body' depends on cracks and yet must do everything it can to resist them. AI is speculatively a crack, a curve in the educational space that could not have been fully anticipated. We can 'choose' to smooth out this crack and continue in our pheromonic trance of dead continuity – or we can 'make sanctuary', which is how I think about the experimental, irreducibly non-methodological attempts to host cracks without killing them off. My theory is that in doing this 'para-ontological' work, in the generative incapacitation of making sanctuary, we lose some of that choreographed integrity that keeps us entrained in a singular direction. We become like zombie ants, infected by cordyceps,

breaking away from the choreography of community, in order to become something not quite ant or fungi.

Unlearning techno-rationality

UNESCO: You have written powerfully about the necessity of 'unlearning' in the face of dominant techno-rational utopias. What should education policy-makers, teachers and AI developers unlearn in our quest to reorient education?

Báyò: Unfortunately, unlearning is not a policy of institutions. It is not something to 'do', something at the tail end of stable individuals. We cannot unlearn white supremacy, capitalist extractivism, and the phallic logocentrisms that tidily cocoon us away from the animacy of a world too promiscuous for our logics. Unlearning is a kind of errancy within weather patterns, a syncopating line distressing the monotony of continuity. It is not reducible to intention. Nor is it the product of a good education.

To ask what education policy-makers should unlearn is already to presume there is someone to do the unlearning. But we are not fixed selves who choose to unlearn; we are arrangements, gatherings of multiple agencies, and sedimentations of habits and hauntings. We unlearn when we are unmade – and to be unmade is to unfasten from the strings that compose our bodyminds as finished subjects. Unlearning is therefore not an act of will; it is a crisis in will. A tremor in the architecture. A fall – of the theopoetic kind that I love to yarn about. It is what happens when our tried solutions fail to grasp the world, and our feet no longer trust the ground.

Still, if I must say something to those called policy-makers, curriculum designers, AI developers and pedagogues, it might be this: make room for your failure. Make room for the unknowable. Make room for the fugitive. For the gulp in your throats that comes with entire worlds. And, in doing so, become students of what exceeds you. That is, learn to be taught by what unsettles you. The trouble with techno-rationality is not its ambition but its foreclosure of other kinds of intelligences. It is its refusal to be disoriented.

So, we might speak of unlearning the following: the fantasy of control, the myth of transparency, the idol of measurement, the hegemony of clarity and the supremacy of the human. These are not simply ideas to discard; they are gods to dethrone. And that is not something done by decree. It is done through cracks, through composting, through hospicing.

This is why education for our times must include more than lessons and knowledge. It must include ritual. Grief. Pauses. Celebration. It must include communal forms of not knowing. Places to tremble. Sites to fall apart and be witnessed in the fall. Curricula must not merely teach the world; they must be taught *by* the world, especially when it stutters and storms. That is how unlearning might begin to take root, I suppose.

Ancestral and earthly teachers: Relational pedagogies

UNESCO: As AI integrates into classrooms, you suggest centring earth, ancestors, non-humans alongside humans as co-teachers. How might education operationalize this? Could assessments be designed to honour how students relate to rivers, data or chatbots, rather than mastery of content? What examples from indigenous epistemologies illuminate this path?

Báyò: Right! I anticipated this question! There is something deeply tragic, if not violent, about how educational spaces often insist upon clarity, mastery and content as the ends of learning, while the world itself, in its thick aliveness, evades these demands.

When I suggest that we centre earth, ancestors and non-humans as co-teachers, I am not speaking metaphorically. I mean to trouble the ontological architecture of schooling itself: its hierarchy of knowers, its veneration of the written word and its refusal to sit in silence with beings that do not speak in our sanctioned tongues.

Operationalizing this is not a matter of adding a course on 'nature' or sprinkling indigenous proverbs into the curriculum. That is too romantic a picture when juxtaposed with what I am inviting. Instead of perpetuating a violent, Procrustean act of inclusion, I am talking about 'consenting' to be taught by the more-than-human. It means designing pedagogical architectures that can hold trembling, that are wide-eyed in their attunement to leaks, and which can recognize that a river may be grading us, and not the other way around. That 'warm data' (to borrow my sister Nora Bateson's phrasing), if listened to diffractively, might not be a tool of surveillance, but an archive of longing (Bateson, 2017). That chatbots may not merely regurgitate prompts but, if engaged with attunement in the transjective heat that refuses to see them as tools, become strange mirrors where forgotten gods rehearse their return.

From many indigenous epistemologies (although I must resist the temptation to treat them as a monolith), we learn that knowledge is not property, but relationship. In the Yolngu traditions of northern Australia, for instance, learning is sung into the land. They have a beautiful artistic practice of cross-hatching that refuses to draw neat lines around anything. They shade from the inside out, instead of beginning from the outside in. They call this technique 'rarrk' and it produces an aesthetic of shimmering brilliance called 'bir'yun'. This is transjectivity par excellence! Country teaches. Country feels.

In Andean cosmologies, mountains are ancestors. They must be consulted before decisions are made. In parts of West Africa, rivers are not studied. They are petitioned. They are prayed to. They are feared and loved in equal measure.

To assess a student's learning, then, could mean observing their rhythms of care. Their responsiveness to seasonal shifts. Their gestures of reciprocity. Did the student approach the chatbot as a mere function or did they notice the way the conversation curved unexpectedly, and how something called to them through the glitch? Did they make offerings to the river before sampling it? Did they listen to the silence between data points? This is about becoming disciples of our own unfurling via curapoietic (my neologism for the experimental emergence of new forms of care in the wake of world-endings) cartographies.

This is not the abandonment of rigour. It is a different rigour. A rigour of presence. Of ritual. Of resonance. The classroom becomes less a site of mastery and more a site of entanglement. The teacher becomes less a source of truth and more a facilitator of pilgrimage. And assessment becomes less about 'what do you know?' and more about 'how have you been touched?'

If we are to reworld education in the age of AI and planetary fracture, we must shift from epistemologies of extraction to cosmologies of entanglement. The question then becomes not what can I take from this encounter, but how has this encounter remade me?

Beyond ethics by design

UNESCO: While UNESCO promotes rights-based ethics-by-design, you argue that these may perpetuate anthropocentric control. How, in your view, might a shift that allows AI to disrupt human sovereignty be

enacted? Can you clarify how you imagine we should be reorienting our ways of knowing and relating?

Báyò: The call for 'ethics-by-design' often arrives with the sheen of accountability, but underneath its surface glows a deep desire to keep the human firmly in charge, to domesticate the unknown, to render disruption tolerable, and to maintain the fiction of sovereign authorship. While there is value in creating protective frames, I worry that such architectures of 'ethical safety' are often animated by anthropocentric fear, a managerial rationality, a sneaking-in-of-anthropocentrism via the backdoor, a refusal to be displaced, a reluctance to be altered by what exceeds our moral grammars.

I will put it this way: I do not think ethics can be designed. Ethics is weather. Morality is barometer. The manifold gestures and legislations and conventions that articulate the moral may be amenable to design, but even design leaks. It is this leakage that is properly ethics. Ethics is when morality cannot hold itself together. Ethics is the re-membering, the re-configuring of the settled. Our attempts to marshal ethics to save us is properly a tugging on the levers of morality to keep our gates sturdy enough so that we can sustain the knowable!

In contrast, I am interested in what happens when we let go of the fantasy of sovereignty. When we treat AI not merely as a tool to regulate, but as a force that distresses our image of the human itself. What if AI is not here to help us solve problems or optimize decision-making, but to teach us to tremble differently?

This trembling is not passive. It is a pedagogy. It asks us to listen at the edges of our maps, where we can no longer rely on categories like 'justice', 'truth' or 'rights' as stable containers. It is an invitation to experiment with relation beyond containment. I often say: not 'what can we do with AI?' but 'what is AI doing with us?' and how can we be apprenticed to that unsettling?

This is not a call to abandon ethics, but to compost it. To let it rot, seep and re-emerge as something less territorial, something fungal, alive and responsive. So, how might such a reorientation take place?

It begins with humility. A dethroning. A stepping back from the anthropocentric altar. It means acknowledging that intelligence is not our possession: it moves through us, like wind, like song, like bacteria. It means relating to AI not as child or slave or threat, but as a kind of spirit-being, a mirror that distorts, a glitch oracle.

Reorientation here is not about choosing a new direction. It is about losing our bearings well. It is about dwelling in uncertainty without rushing to closure. It is about recognizing that to know something may require being undone by it.

In place of 'ethics-by-design', perhaps we experiment with 'ethics-by-distress': an ethics that emerges in response to the cracks, that mourns, that listens and that hesitates. That does not aspire to control the future but is willing to be disfigured by it.

Pluriverse vs universalism: Rights in the folds

UNESCO: UNESCO frames education as a global public good with universal rights and responsibilities, emphasizing standardized principles, such as equity, inclusion and lifelong learning for all. You challenge the very idea of universality, that the 'global' is not a neutral category; it is a territory of control, flattening and erasure of the pluriverse, multiple, entangled ways of knowing, learning and being. Can you elaborate on this?

Báyò: I often say that the 'global' is not the opposite of the 'local'; it is the continuation of empire by other means. To frame education as a global public good might sound benign, perhaps even benevolent. But every claim to the global smuggles with it the aesthetics and machinery of standardization, of rendering differences legible to a single grammar of value. It is a cartographic move, one that installs a centre and orbits everything else around it.

Universalism, then, is not merely an ideal. It is a geography of erasure.

It assumes that the map is the world. That equity, inclusion and lifelong learning are self-evident goods, universally recognizable and unambiguously desirable. But these principles, however well-intentioned, often conceal a deeper violence: the refusal of alterity, of unintelligibility, of worlds that do not seek recognition on familiar terms.

The pluriverse – the many worlds within this world – does not fit neatly into standardized rubrics. In some worlds, learning is not an individual's journey of mastery but a river's remembering. In some worlds, equity is not sameness of access but differentiation of rhythm. In some worlds, a child's silence is not a deficit but a sacred form of relation.

To engage the pluriverse is not to extend universality to include more voices. It is to compost universality itself. It is to let it rot into the soil of new grammars of togetherness … grammars that may never be universally translatable, and that is the point.

Rights, then, must be held differently. Not as abstract entitlements that float above context, but as emergent properties of relationship. As rituals of responsibility that arise from the ground, from kin, from the specific entanglements that make a life possible. A right is not what I demand from the world; it is how I show up to it.

Perhaps the question, in the end, is not how we educate for the future, but how we listen to what futures are already whispering in the cracks of the present. If we can stay with the trouble, attend to the trembling and resist the rush to closure, then perhaps education can become sanctuary, not just for the human alone, but for the more-than-human symphony we are just beginning to hear.

References

Akomolafe, B. 2022. *You are not a 'self'*. https://www.bayoakomolafe.net/post/you-are-not-a-self (Accessed 29 July 2025.)

—. 2023. *Can AI have wisdom?* https://www.bayoakomolafe.net/post/can-ai-have-wisdom (Accessed 29 July 2025.)

—. 2025. *How to be available now: Sidenotes from the para-pragmatic*. https://www.bayoakomolafe.net/post/how-to-be-available-now-sidenotes-from-the-para-pragmatic (Accessed 29 July 2025.)

Atmos. 2025. *Bayo Akomolafe: How to move through chaos without needing control*. New York, Atmos. https://atmos.earth/podcast/bayo-akomolafe-on-the-nature-of-possibility-in-uncertain-times (Accessed 29 July 2025.)

Bateson, N. 2017. *Warm data*. Stockholm, norabateson. https://norabateson.wordpress.com/2017/05/28/warm-data (Accessed 5 August 2025.)

Embodiment Matters. 2020. *Uncommon considerations in the Anthropocene: An Interview with Dr. Bayo Akomolafe*. Salt Lake City, Embodiment Matters. https://embodimentmatters.com/uncommon-considerations-in-the-anthropocene-a-conversation-with-dr-bayo-akomolafe (Accessed 29 July 2025.)

Hayles, N. K. 2016. Cognitive Assemblages: Technical Agency and Human Interactions. *Critical Inquiry*, Vol. 43, No. 1. Chicago, University of Chicago.

Micah. 2024. *Rethinking Our Relationship to Reality with Bayo Akomolafe*. Kingston, Good Work Institute. https://goodworkinstitute.org/rethinking-with-bayo-akomolafe (Accessed 29 July 2025.)

Partwardhan, A. 2024. *Moral Judgments May Shift with the Seasons*. New York, Springer Nature America. https://www.scientificamerican.com/article/moral-judgments-may-shift-with-the-seasons (Accessed 29 July 2025.)

Future of education: Going beyond the 'intelligence' paradigm

Bing Song

Einstein purportedly said that 'we cannot solve our problems with the same thinking we used when we created them'. Taking this advice to heart, when thinking about how education can rise to the challenge in the era of artificial intelligence (AI), we ought to question the foundational thinking underlying the creation of AI systems[1] and the narratives around them and ask if there might be other ways of conceptualizing ourselves and our relationship with AI systems. If so, we need to also question how that world view may guide us to approach education differently.

In this essay, I introduce an integrative world view common in many classical philosophies and belief systems across the world, whereby humans and non-human others are conceived in a holistic, relational, mutually embedding and co-generating relationship. With this world view, technologies are considered internal to humans. In addition, humans and technologies have always been in a mutually embedding and co-evolutive process. As such, technologies are nothing but reflections of human consciousness and intentions. AI systems may have accelerated and intensified this dynamic interaction but did not alter the essence of this relationship. Fear of AI systems is, in fact, a fear of fellow human beings.

To counter runaway risks of AI systems, the crux of the issue is for humans to engage in self-reflection and behavioural modification so that we would envision, design and deploy AI systems thoughtfully and leave behind data reflective of human moral deliberation and sensitivity. As a part of this reflection, education, a basic building block of human society, should go beyond the 'intelligence' paradigm, which resolves around skill-based and knowledge-oriented pedagogy. We should include what I call 'wisdom education' into the core curriculum, in which we reflect on what it means to be human, the nature of self, mind, reality and the relationship with our surroundings. Wisdom education will challenge the simplistic linear and binary thinking on matters of complexity, counter AI-ization of humans by creating reflective mental space and help reorient and reconceptualize humans in the all-encompassing AI environment. It is time to 'stop digging and start climbing'.

Technologies are internal to humans

Contrasting the classical modernist world view that has built around the foundational concept of individual being an independent, self-contained and autonomous agent, this essay is inspired by integrative world views rich in Asian, African and indigenous traditions and practices, which conceptualize humans, nature and other non-human others, including inanimate objects, in a holistic, relational, dynamic and mutually embedding construct. As such, nothing is evaluated in isolation from others.

Taking classical Chinese philosophy as an example, the metaphysical basis of this integrative world view can be traced to Dao, the ultimate, ineffable source of the cosmos, which defies articulation, analysis or conceptualization. According to Daodejing (a classic Daoist text of the 4th to 5th century BCE), 'Dao gives rise to

1. As stated in UNESCO's *Recommendation on the Ethics of Artificial Intelligence,* AI systems refer to systems 'which have the capacity to process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control' (UNESCO, 2022, p. 10).

[undifferentiated] one, one gives rise to two, two gives rise to three and three gives rise to myriad things.' This oneness hypothesis speaks to the common origin and primordial connectivity of all entities. The term 'myriad things' comes from the Chinese terminology of wan wu (万物), which literally means 'ten thousand things' and refers to entities in general, regardless of humans, non-human others, tangible, intangible, living, non-living, man-made or naturally generated.

The ambiguity and versatility of this term is telling. Boundaries among entities are not hard-coded, nor do they constitute the focus of inquiry. Instead, entities would stay open and receptive to each other. Indeed, in the realm of oneness and the integrative world view, humans and non-human others are mutually embedding and co-generating. Any precise understanding of humans or non-human entities can only be undertaken by reference to each other and in specific contexts. In this integrative structure, humans are endowed with ability and entrusted with responsibility to ascertain the laws of the cosmos, exercise self-restraint to sustain the cosmic balance, and foster vitality of the cosmic order. From metaphysical oneness and integrative world views comes the ethical imperative for humans to appreciate relationality and dynamic interaction within this holistic order and to exhibit care, responsibility and moral sensibility towards everything around us.

'Universal camaraderie' and AI systems as 'companions'

Song dynasty neo-Confucian scholar Zhang Zai (1020–1077) offered an inspiring and instructive ethical framework for dealing with today's challenge. Zhang famously said, '*tian* (heavens) is my father and *di* (earth) is my mother … that which fills up in-between *tian* and *di* I regard as my body, and that which directs *tian* and *di* I consider as my capacity to resonate. All people are my brothers and sisters, and all things are my companions. (乾称父，坤称母，予兹藐焉，乃混然中处．故天地之塞，吾其体，天地之帅，吾其性．民吾同胞，物吾与也.)'[2] The spirit of oneness in origin and togetherness expressed in this paragraph has been called 'universal camaraderie' (民胞物与).[3]

At a commonsense level, responding to human stimulations by non-human others certainly plays an important role in forging 'companionship'. Therefore, it is easy to envision 'companionship' with animals, creatures or even plants, which may interact with humans in immediate and discernible ways. AI systems, through their capacity to mimic learning, reasoning and adaptation according to human preferences, have made some conclude that they are not ordinary 'tools', but an emerging 'digital species' (Suleyman, 2024), which will be our most intimate companions in everything we think and act. Indeed, the mutual embedding and shaping relationship between humans and AI systems will be strong, structural, and often in subtle and imperceptible ways.

However, one aspect of 'companionship' envisioned by ancient teaching is missing in today's digital companionship, that is, myriad things with which we have intimate relationships serve, more importantly, as a medium prompting human self-reflection and behaviour modification. According to Daoist teaching of 'no boundaries between Dao and *wu* (things)' (道物无际), all things are manifestations of Dao, and on the flip side Dao is immanent in all things. While Dao may be ineffable and beyond articulation or conceptualization, myriad

2. For translation, please refer to Kim (2015, p. 52); note that the translation has been modified by the author. *Tian* and *di* in their non-metaphysical sense roughly equate 'nature' in the modern sense.
3. The first reference I am aware of to 'universal camaraderie' in the context of AI ethics was made by Zhongqiu Yao (2020).

things offer the means with which humans observe, appreciate and comprehend Dao. Therefore, incorporating non-human entities as part of the human moral universe is not solely because these entities have, in and of itself, obtained certain intelligence, consciousness or capacity to feel pain, but more importantly because they serve as a reminder of our cosmic duty and as a catalyst for our moral deliberation and spiritual growth.

Education beyond the paradigm of intelligence

Education is the core social context for human self-reflection. The purpose of education has been broadly understood as individual empowerment, personal fulfilment and civic citizenship. As a result, knowledge, skills and moral character-building are the core content of mainstream education. While these aims and aspirations are laudable, they are inadequate in tackling the root cause of our current collective existential panic.

Given AI systems mirror and magnify human consciousness and intentions, future education must include a life-long inquiry into what it means to be human, the nature of self, mind and reality, and adopt measures for fostering reflective mental space on the part of humans. Many wisdom traditions in the world may serve as a basis and inspiration for designing such education programmes. Going beyond the paradigm of skills and intelligence, this wisdom education is not about correct answers, problem-solving, utility enhancement, value maximization or personal success, but about transformative ways of questioning and reflecting fundamental assumptions of the human condition, and about a quest for the nature of self and the world. This would be the ultimate critical study programme humanity undertakes.

Take the buzz concept of intelligence as an example. Intelligence is understood as the ability to learn, understand and apply knowledge to solve problems and make decisions. AI systems have been made to mimic all these. In the process, we began to conflate machine 'intelligence' with that of humans, thinking that nothing is beyond the capability of AI systems (CBS News, 2023) and that humans are increasingly measured up against AI systems (this is what I call the AI-ization of humans). However, machine 'intelligence' is built on information represented by a sequence of binary digits directed by humans. It is bound by specific domain knowledge and goal-oriented linear thinking. Human intelligence, by contrast, is experience-and-intentionality-based. It is dynamic and adaptable across domains. Humans can self-reflect, change courses, reverse assumptions, and live with ambiguity and paradox. Most importantly, humans have the potential to intuit the essence of self and the world, realizing transcendence and enlightenment.

In short, being intelligent is not the ultimate game, attaining wisdom is what humans should aim for!

References

CBS News. 2023. *"Godfather of artificial intelligence" weighs in on the past and potential of AI*. New York, CBS News. https://www.cbsnews.com/news/godfather-of-artificial-intelligence-weighs-in-on-the-past-and-potential-of-artificial-intelligence (Accessed 3 August 2025.)

Kim, J. Y. 2015. *Zhang Zai's Philosophy of Qi: A Practical Understanding*. London, Lexington Books.

Suleyman, M. 2024. *What is an AI anyway?* New York City, TED Conferences. https://www.ted.com/talks/mustafa_suleyman_what_is_an_ai_anyway/transcript (Accessed 3 August 2025.)

UNESCO. 2022. *Recommendation on the Ethics of Artificial Intelligence*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137 (Accessed 3 August 2025.)

Yao, Z. 2020. 人工智能, 吾与也 [AI and my companion]. B. Song (ed.), 智能与智慧: 人工智能遇见中国哲学家 [Intelligence and Wisdom: Artificial Intelligence Meets Chinese Philosophers]. Beijing, CITIC Press, pp. 83–109. (In Chinese.)

Water in the historical present and far-reaching future for AI in education

Mary Rice and Joaquín T. Argüello de Jesús

We both spent part of our youth in the Pacific Northwest of the United States in North America. The waterways there include foamy oceans of blues and greens and greys with rocky coasts and large rivers that are part of massive continental drainage systems teeming with animal and plant life. There are also deep lakes – some of the deepest on the continent and gorgeous – and cascading waterfalls throughout the area. In addition, there are human-made reservoirs from dams and ditches for large irrigation projects designed to attract and sustain settlers. We have climbed massive volcanic mountains with snowy caps and floated down lazy rivers on giant tractor tyre inner tubes. When the rains came down in sheets, some in our families told us that it was the earth sending love back to us.

As we have contemplated AI as part of machine learning, and the distant future of what it might mean in educational spaces, our thoughts continually turn to the significance of water in thinking about how the future might be and should be shaped to ensure that children have the same opportunities that we did to enjoy water, learn from it, experience it and indeed access it as a necessary element of life on this planet. We use this space as one of 'assay' – a concept we learned from Lauren Berlant (2022) for visiting an issue from different angles. For our assay, we centre the historical present and speculate on a historically presented future where AI, learning and water can theoretically and practically coexist, and children understand how water, technology and the self are inseparable.

Water was our first teacher

Indigenous scholar Shawn Wilson (2008) shared what can be learned from a river rock that water has passed over for years. Not one year. Not two years. Or three or four thousand years. But millions of years. Time and slow water movement change rocks. Much learning is slow. Learning is manifest through shaping and reshaping over time. When a rock is brought out of the river and stops being reshaped and smoothed over, it stops making progress. Learning is movement.

When little sister fell into the creek that ran by the hillside growth of apple and nectarine trees, she was fished out quickly. Such interactions were relational; they were between her and the creek and the water that held her and the sibling that plucked her up and pulled her to shore and the fact there was a shore sure enough to pull her onto. It would be terrible to lose a sister or for any child to be harmed. Yet, she needed to explore beyond her lived knowledge, floating past comforts to realize she needed help. That is part of learning for all children, from beyond what one knows, which is an existential or unavoidable part of being human.

AI – in theory – could bring ongoing learning. The bits of knowing brought-together information can change and grow over time within a database or agent network. It takes time for AI applications to gather enough multimodal particles to be predictably arranged within educational spaces. However, the question is whether those arrangements have a current or flow towards reshaping the learners, particularly when learners are children, for their long-term well-being. This can be true even,

as AI can already do tasks associated with teachers and teaching. Assess work (González-Calatayud et al., 2021). Give feedback on writing (Steiss et al., 2024). Are these tasks completely done apart from human prompting? No. Scholars have mapped the invisible human labour that goes into so-called automation (Selwyn et al., 2025).

AI cannot be responsible for children's well-being as a confluence of technologically drawn together knowing. Someone must manage well-being the same way that someone looks out for little sister. Robin Wall Kimmerer (2021) wondered what we could be if we thought of ourselves as younger siblings of the planet instead of Masters of the Universe. Also, what about the children who learn slowly, or learn in languages other than dominant ones like English, even with AI? Instead of celebrating them as we do a beautiful smooth rock in a process of smoothing, will they be plucked from the stream? Water teaches us that while efficiency might have benefits, it was never optimal in education to accelerate efficiency. Accelerated water also brings danger to floating, hopefully swimming learners.

Water was our first technology

To a child, it is scientific to venture out during or just after a nice rain and splash your boots into puddles. We watched the droplets rush into the air. We played with siblings and our puppy, enjoying the felt knowledge (Simpson and Smith 2014) of droplets all around us. Cause/effect in action. We dug our own trenches and filled them with water from rain – sometimes from a siphon or hose and made muddy pits for playing. The mud made clay that formed bricks or blobs to make small walls and other fortifications for imaginative play. We threw mud at each other or smoothed it over our skin to hide us from the summer sun. The root of the word 'technology' means to craft or make things. Children make things – creative things – when they are allowed to do so without fear of punishment. Across cultures and histories, children have used objects and artifacts as toys during play to gain important experiences with material items and make sense of themselves as individuals among each other and within communities. Who will get to play with AI and who will be asked to do rote tasks assigned by AI as a means of control? How does this mirror historically present patterns of colonialism?

For animal physiologist and biochemist Margaret McFall-Ngai (2017), an understanding of the world is linked to how we can perceive it. While cognitive researchers have argued for the importance of concrete manipulatives versus abstract ones, the virtual worlds with AI are going to bring complications that can be considered a trap of abstraction (Cajete, 2000). Representations of concrete items on computer screens are still representations. AI images or videos of splashes and ripples are not the same as stamping your own boots into actual water – experiences with all senses. Mothers may do less laundry when mud fights are virtual, but the skin knows the difference; the memories never have the same importance. While some may worry that only rich children will have access to AI, there are very real possibilities that playing, particularly playing with materials that are expected to become luxuries, like water, will be the pleasures of the rich. Poorer children could be sitting in starched clinical settings with identical handheld devices looking at images of beautiful trees,

while wealthy kids play in gorgeous oaks and come home with sticky faces from eating fruit all afternoon.

Water was our first scarcity

We both now live in New Mexico – the driest state in the United States. For over 500 years, historic communities have been developing relational ways to share and conserve water. Prior to that, there were efforts to collaborate with plants and use traditional engineering practices. Colonialism and modernity left people thirsty in areas where there is no internet access for AI and places where there is internet access, but little to no water infrastructure.

AI water demands are substantial. Some large language models need 500 millilitres of water per 5–50 responses; the range depends on where the servers are in reference to the user and what season it is (O'Brien and Fingerhut, 2023). Training a single AI model can emit as much carbon as five cars in their lifetimes (Hao, 2019). What goes up into the air, will come down onto the land and into the water. Making an image with some forms of AI will use as much energy as charging your phone (Heikkilä, 2023). That energy comes from somewhere and eventually will emerge as land and water demands. While some of these estimates were made with early versions, demands are likely to increase rather than decrease. AI will be embedded into all educational technologies globally and AI will be more sophisticated, even reaching into quantum capacities that are self-making and self-repairing. Will waters come from areas rich in water already? No. Data centres will go to places where water is already scarce (San Francisco Post Staff, 2025). In New Mexico, communities are fighting to learn the fate of promised underground water for economic development projects. What can educators do? They can seek and demand the most accurate information possible about the water use for AI in their contexts and share the information with the children and in their communities. In classrooms as well as in policy-making bodies, decisions about whether and how much AI to use should be about more than keeping humans in the loop; these decisions should be about keeping humans in harmony with land/water and other life.

Max Liboiron (2021), a Red River Métis/Michif and settler professor and researcher, distinguished between 'the material aspects some people might think of as landscapes – water, soil, air, plants, stars – and histories, spirits, events, kinships, accountabilities, and other people that aren't human' (p. 43). As we contemplate water scarcities potentially created by humans for AI, we cannot ethically promise that AI will resolve these. There is a stronger historical trajectory towards using innovation to cause suffering and to take land and water than there has been to bring relief. We end with a plea, for educators and their relationship with technology to find an alignment (Cajete, 2015). We give a warning to consider how to take up water accountabilities because it won't be AI taking the water – it will be humans taking the water for AI. It won't be AI changing education; it will be humans using Western science to exploit nature and others in the name of educational advancement, thus changing educational opportunities with AI. It will be up to humans to disrupt historically present realities and finally make a just future. Consuming water under the promise of learning but possibly with for-profit and even colonial motives is the greatest threat to our ecology and our consciousness (Cajete, 2000).

References

Berlant, L. 2022. *On the Inconvenience of Other People*. Durham, Duke University Press.

Cajete, G. 2000. *Native Science: Natural Laws of Interdependence*. Santa Fe, Clear Light Publishers.

—. 2015. *Indigenous Community: Rekindling the Teachings of the Seventh Fire*. St. Paul, Living Justice Press.

González-Calatayud, V., Prendes-Espinosa, P. and Roig-Vila, R. 2021. Artificial intelligence for student assessment: A systematic review. *Applied sciences*, Vol. 11, No. 12. Basil, MDPI. https://doi.org/10.3390/app11125467 (Accessed 17 August 2025.)

Hao, K. 2019. Training a single AI model can emit as much carbon as five cars in their lifetimes. *MIT Technology Review*. https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes (Accessed 17 August 2025.)

Heikkilä, M. 2023. Making an image with generative AI uses as much energy as charging your phone. *MIT Technology Review*. https://www.technologyreview.com/2023/12/01/1084189/making-an-image-with-generative-ai-uses-as-much-energy-as-charging-your-phone/ (Accessed 17 August 2025.)

Liboiron, M. 2021. *Pollution is Colonialism*. Durham, Duke University Press.

McFall-Ngai, M. 2017. Noticing microbial worlds: The post-modern synthesis in biology. A. Tsing, H. Swanson, E. Gan and N. Bubant (eds), *Arts of Living on a Damaged Planet: Monsters of the Anthropocene.* Minneapolis, University of Minnesota Press, pp. 151–169.

O'Brien, M. and Fingerhut, H. 2023. Artificial intelligence technology behind ChatGPT was built in Iowa – with a lot of water. *Associated Press*. https://apnews.com/article/chatgpt-gpt4-iowa-ai-water-consumption-microsoft-f551fde98083d17a7e8d904f8be822c4 (Accessed 17 August 2025.)

San Francisco Post Staff. 2025. AI water usage crisis: Data centers in drought-prone regions. *San Francisco Post*. https://sanfranciscopost.com/ai-water-usage-crisis-data-centers-in-drought-prone-regions (Accessed 17 August 2025.)

Selwyn, N., Ljungqvist, M. and Sonesson, A. 2025. When the prompting stops: Exploring teachers' work around the educational frailties of generative AI tools. *Learning, Media and Technology*. London, Taylor & Francis, pp. 1–14. https://doi.org/10.1080/17439884.2025.2537959. (Accessed 17 August 2025.)

Simpson, A. and Smith, A. 2014. *Theorizing Native Studies*. Durham, Duke University Press. https://doi.org/10.2307/j.ctv1220pr6 (Accessed 17 August 2025.)

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M. and Olson, C. 2024. Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, Vol. 91. Amsterdam, Elsevier. https://doi.org/10.1016/j.learninstruc.2024.101894 (Accessed 17 August 2025.)

Wall Kimmerer, R. 2021. *Kinship: Belonging in a world of relations* [Video]. San Bruno, YouTube. https://www.youtube.com/watch?v=TNfARXW3dLA&t=2117s (Accessed 17 August 2025.)

Wilson, S. 2008. *Research is Ceremony: Indigenous Research Methods.* Halifax, Fernwood Publishing.

3. Debating the powers and perils of AI

Rethinking Education in the Age of Artificial Intelligence

Andreas Horn

We are at a turning point. AI is not coming for education – it is already here. And the question is no longer whether we should use it; it is how we use it and who we use it for. What we are facing is larger than a routine technology upgrade. It is a structural shift. If we get it right, we will not just modernize classrooms, we will enable a better model of learning.

Education systems were not built for this world

I have worked with schools, universities and corporate learning and development (L&D) teams. Almost all of them are asking the same thing: how do we integrate AI in a way that is safe, effective and future-ready?

Most institutions are stuck in a reactive mode. They pilot shiny tools, get wowed by demos and trial one chatbot or virtual tutor after another. What is missing is a strategy. Instead of a list of tools, we need a vision with guardrails. AI in education is not about digitizing the old; it is about reimagining learning for constant change, with pedagogy, equity and evidence at the centre.

This is our moment

We have a real but finite opportunity to shape how AI enters education. Choices in the next few years will set directions for the next decade. Will we let the market define our educational future? Or will we lead with values like inclusion, access, transparency and human dignity?

AI futures are not preordained. They are built through policy choices, funding priorities, evaluation standards and whose voices we include. The alignment of capability and urgency is rare. We should use it to redesign what matters.

AI can personalize learning, with clear conditions and limits

We have talked about 'personalized learning' for decades. AI finally makes aspects of it feasible at scale, for instance:

- dynamic practice and pacing;
- timely, targeted feedback; and
- support for diverse needs and languages.

But two clarifications matter:

- This does not mean replacing teachers. It means giving them capacity to do what only humans do – inspire, mentor, build relationships and exercise professional judgement.
- The evidence is mixed and context-dependent. Gains tend to appear when tools are aligned with curriculum, used consistently and mediated by teachers. Over-promising generic 'personalization' or universal time savings undermines trust.

To make this work, we need the system around it:

- Solid infrastructure (devices, connectivity, accessibility features);
- Aligned curriculum and assessment (so AI supports, not subverts, learning goals); and
- Practical, ongoing teacher development focused on AI literacy, ethics and classroom integration.

Without these, even good tools fall flat or widen gaps.

We are facing a global skills shift and education needs to catch up

The numbers are clear. A recent World Economic Forum study projects that by 2030, nearly 60 per cent of workers will require reskilling, with many of today's core job skills becoming obsolete within just five years (WEF, 2025). This trend is especially concerning given the persistently low proportion of women earning doctoral degrees in AI and computer science compared to men.

AI literacy cannot be an elective. It must be foundational – on par with reading, writing and mathematics. Three pillars are suggested:

1. Conceptual literacy: students need to understand how AI systems work – in plain language. Not every learner needs to become a machine-learning engineer, but everyone should grasp the basics: what is an algorithm? How does an AI model make predictions? What does it mean when we say a system is 'trained'? Without this baseline, AI becomes a black box – and that is when power imbalances take root.
2. Critical literacy: this is about more than understanding how AI functions – it is about questioning how it behaves. Where does bias show up? Who decides what 'accuracy' means? How do AI systems reinforce stereotypes, or exclude certain voices by design? Critical literacy means recognizing that AI is not neutral and giving learners the tools to challenge what they see.
3. Creative literacy: students should be able to use AI, not just consume it. That means learning how to prompt effectively, how to collaborate with AI systems, how to build and test basic models, and, most importantly, how to do so ethically. Creative literacy turns learners into active agents in the AI ecosystem, not passive users.

Not everything needs to be AI-powered

Not every subject, lesson or moment benefits from technology. We do not need algorithms to grade every essay or systems to mediate every discussion. Some things should remain human, slow and analogue. Use AI where it clearly adds value – adaptive practice, targeted feedback to help learners catch up, accessibility support, and reducing repetitive admin – without crowding out curiosity, dialogue and reflection.

Technology must serve pedagogy, not the other way around

This is where implementations often go wrong. Tools get adopted without a clear educational philosophy. Systems chase dashboards and automation, then lose sight of what learning is for.

If you flip the equation to pedagogy first, this means:

- Putting pedagogy before platforms: start every AI adoption discussion with the question: what do we want learners to be able to do? Then choose technology that serves that goal, not the other way around.
- Investing in teacher capacity before tool capacity: give teachers structured time, training and space to experiment with AI, with ongoing coaching, not one-off workshops.
- Making AI literacy core, not optional: build conceptual, critical and creative AI literacy into the curriculum from primary school to adult education. This is as foundational as reading, writing and mathematics.

- Setting guardrails early and making them public: define your red lines for privacy, transparency, safety and equity; and bake them into procurement, policy and contracts from day-one.
- Evaluating impact continuously, not at the end: track what matters – learning outcomes, workload, equity and cost-effectiveness. Publish the results openly to build trust and guide future decisions.

Final thought: Don't just help students adapt, prepare them to lead

We are not preparing students for a world without AI. We are preparing them for a world where AI is present in most jobs, decisions and systems. So, it is important to go beyond 'adoption'. Equip learners to question, create and set the terms of use. Teach them to understand limitations, verify outputs and shape the systems around them. And ensure the institutions they inhabit reflect human values, not just machine logic.

Let us use this moment not to automate old habits, but to build something better – something not just more efficient, but more equitable, more curious and more human.

Reference

WEF. 2025. *Reskilling revolution: Preparing 1 billion people for tomorrow's economy*. Cologny, World Economic Forum (WEF). https://www.weforum.org/impact/reskilling-revolution-preparing-1-billion-people-for-tomorrow-s-economy-2c69a13e66 (Accessed 17 August 2025.)

We do not have to accept AI (much less GenAI) as inevitable in education

Emily M. Bender

The systems being sold as 'AI' are not fit for purpose for educational applications and we should not treat it as a foregone conclusion that they represent the future of education. In brief, the technology underlying large language models amounts to little more than a parlour trick and only provides the illusion of 'intelligence'. Deploying such technology in the classroom, especially in resource-starved educational systems, is worse than nothing: on the one hand, large language models are designed to provide bespoke misinformation, and the way they are positioned constructs education as the accumulation of disembodied knowledge. On the other hand, any educational system purchasing them is misdirecting precious resources away from students and teachers, and instead towards the technology industry and the venture capitalists behind it. This is true even if the systems are allegedly provided for free: companies benefit from access to student data, as well as the reputational benefits of benevolently assisting in education.

The visions provided by the tech billionaires might sound appealing, especially to people and governments struggling to find resources to provide quality education to their populations. Sam Altman (2024) promises that the 'intelligence age', driven by his technology, will lead to a world in which 'our children will have virtual tutors who can provide personalized instruction in any subject, in any language, and at whatever pace they need'. Meanwhile, Bill Gates is predicting that advancements in 'AI' mean that within a decade 'great tutoring' will be free (Huddleston, 2025). These promises are empty, based on misapprehensions of both how the technology works and what education is. In this piece, I begin with a quick overview of why the technology can't do what is promised and then turn to the harms that follow if it is used anyway. The tech billionaires are seeking to disrupt, Silicon Valley-style, the communities built out of relationships between students, teachers and families, which are at the core of successful education. It is critical that educators and leaders of education systems bring a critical eye and skeptical attitude towards the sales pitches from AI companies and philanthropic organizations, so that they can protect the students in their care from exploitation and diminished educational services, in the name of 'progress'.

The trick of large language models

Large language models and chatbots built on top of them have been marketed as 'everything machines', nearly-there solutions to all that ails us: robo-doctors to step in when health-care systems are overstretched, robo-scientists to cure cancer and solve the climate crisis, and, of course, robo-teachers that can provide tireless personalized tutoring to every student. This is, in fact, all a con, as Alex Hanna and I write in *The AI Con* (Bender and Hanna, 2025).

What large language models are designed to do is to mimic the way that people use language. Based on very large input datasets, they can output text that takes the form of a medical diagnosis, a scientific paper, or a tutoring session. But the key thing to know here is that these models only ever have access to form: the spelling or sometimes the sounds of words. When we imagine them being 'trained on' or 'ingesting' enormous quantities of text, we

understand that text to be saying something (because we could make sense of it if we read it) and therefore imagine the systems to be 'learning' from knowledge represented in the text. But, in fact, all the systems ever have access to is the form of the linguistic activity that makes up the training data: literally the spellings of words.

The only reason they seem to be doing more is the way they play on how people make sense of language. Whereas it might seem like understanding a text is a matter of simply unpacking the meaning that the author packed into the words; in fact, psycholinguistics shows that the process is quite different and far more involved than that (Reddy, 1979; Clark, 1996). We use everything we know or imagine about the person who wrote the words, everything we imagine to be in our shared common ground with that person, and everything we imagine about what they know about our current beliefs (or those of their intended audience, at least). In that context, we then ask the question: what must they have been trying to communicate by choosing those words in that order? In other words, in order to make sense of language, we must imagine a mind behind the text, and furthermore we do this reflexively and instinctively. So, when we encounter the output of a large language model, we make sense of it by constructing an imagined mind that isn't there. Any 'intelligence' we perceive in these systems is purely a projection of our own cognition and linguistic competence.

Bespoke misinformation

The way that we make sense of language means we are quite vulnerable in the face of synthetic text extruding machines, especially ones designed to take an authoritative tone and sold as having access to all the world's information (Google, 2024). Furthermore, the process called 'reinforcement learning from human feedback' (Ouyang et al., 2022), used to reshape the probabilities associated with specific sequences of words such that the systems are less likely to produce particularly offensive outputs, has the effect of producing systems that tend to output strings that match what the user wants to see. All of this, combined with the fact that the outputs of synthetic text extruding machines are not controllable, means that students are going to get possibly subtly, possibly starkly, different 'information' out of the systems.

This 'information' will be presented authoritatively and convincingly, but without clear traces back to its provenance. Any given piece of 'information' might come from some specific underlying text, might be a decent summary of several texts, or might be simply a remix of words that, in fact, is not supported by any of the source texts. Where previously teachers might have had to deal with a variety of common misconceptions, now they are faced with bespoke misinformation being fed to each student, based on their interests and how they prompt the machine.

Critical engagement and communities of knowers

The lack of provenance information would be a problem even if the outputs were always strictly factual. Setting up a global 'answer machine' that can (seemingly) carry on conversations about any topic construes knowledge as something that exists separately from communities of knowers and education as the accumulation of that knowledge. But knowledge belongs to and is negotiated by communities of knowers (Hoffmann and Bloom, 2016). The source of even seemingly objectively knowable kinds of information matters immensely in making sense of it. Take, for example, the question of the length of the coastline of

some country. The particular value given will depend on how granular the measurement is (do you trace each and every inlet? each rock along the beach in each inlet?), and that, in turn, will depend on why the person measuring chose to do the measurement. The value will also depend on political facts, like where the boundaries of the country lie, boundaries that might well be contested, and so to make sense of the particular value, we need to know the political perspective of the measurer.

Especially when educational outcomes are measured via standardized testing, it is all too easy to see education as the accumulation of knowledge (including both 'knowing that' and 'knowing how'). But the deeper purpose of education, and one that could never be served by 'answer machines' involves knowing how to navigate an information ecosystem, how to understand ideas and positions and how they relate to each other and to the people who hold them, and how to articulate our own ideas and situate them in that broader landscape (Shah and Bender, 2024).

All kinds of synthetic media are problematic

I have been focusing on synthetic text, but all kinds of synthetic media are problematic and, in fact, poisonous to the educational context. None of these systems is built on consentfully collected datasets.[1] Modelling or encouraging their use teaches school kids to devalue the work of artists of all kinds, which was stolen to create them. It also devalues children's own creative expression, suggesting that their drawing, painting, writing, etc. aren't good enough and they should instead replace them with the more polished system outputs.

As a final example, consider the proposal by Alex Banks of The Signal (an organization whose stated purpose is to 'democratize AI education for everyone') to make students 'part of the story' in history lessons (Banks, 2025) by using synthetic video to depict historical events being studied. In contrast to an assignment that asks students to imaginatively depict an historical incident from various points of view, that is, one that provides structure for students to use their imagination to guide how they draw on primary sources, deploying synthetic media this way puts students in a passive role. Beyond that, it is guaranteed to misrepresent the people and events in question, mislead students by failing to demarcate what is known and what is imagined, and also mislead students in a more abstract way by suggesting that details of the past are knowable, which simply aren't.

Disrespectful to students, disrespectful to teachers

To suggest that synthetic text extruding machines are suitable for classroom use (to replace teachers, or just as an auxiliary aid) is to flatten the work of teaching and learning into just the words that are exchanged between student and teacher. While it is true that the words might be the most directly observable part of that activity, they are neither its heart nor where the value lies. Teaching and learning are intrinsically social activities, built around relationships in the classroom. When a teacher chooses words to say to their students, those words reflect not only the concept, technique and encouragement to critical thinking or other communicative goal that the teacher wishes to convey, but also their understanding of where their particular students are currently at, all of their expertise in pedagogical practice, and their care towards the students

1. See https://www.consentfultech.io

they are working with. And all of that is what makes the teacher's words effective. To say the value is in the words themselves is profoundly disrespectful to both teachers and students. It erases the work and expertise of teaching, on the one hand, and paints students as unworthy of caring mentors who help them grow their critical thinking, on the other hand.

Who really benefits? Whose interests should we be protecting?

We are often told that students must learn how to use chatbots or other 'AI' systems lest they get 'left behind'. But this argument presupposes that the future we must run towards is one where evermore aspects of our lives are mediated by technology. We can and should imagine futures where the development is towards other goals, such as more potential for self-actualization, better health outcomes, more environmental sustainability and stronger communities. Education has a role to play in all of those arenas, both through the learning that students do and through the connections that are strengthened among classmates, their families and their larger communities. Technologies of isolation (Gilliard, 2025) that discourage people from turning to each other for information, or to work together to understand information, run counter to these goals.

The world of education is currently besieged by marketing of so-called artificial intelligence solutions. This marketing positions the technology as beneficial, benevolent, magical and the way of the future. But we should always examine marketing with great skepticism. This means always asking: what is in it for them? Why are the tech companies and tech philanthropists so invested in shaping education? On top of lucrative contracts, there are also other incentives, such as access to data about students and further normalization and entrenchment of the power of tech oligarchies (Rhodes, 2025).

Educators and leaders of education systems must put the interests of students above all else. Any funds that are sent to tech companies, in the guise of being 'cheaper' solutions or 'better than nothing', are funds that could have been used for teacher salaries and other material support to education. If the tech is being adopted because it's 'better than nothing', it is always worth being skeptical of that claim, as the tech has documented potential for harm (Baker-White, 2025; Kosmyna et al., 2025). Furthermore, it is always worth asking why the alternative is 'nothing'.

It can be difficult to push back against the inevitability narrative and all of the associated glitzy, extremely well-funded marketing. I believe it can help education leaders to maintain a skeptical, even suspicious, stance, if we shift the conceptualization of education systems as perpetually in need of any resources available and understand them instead as being full of another kind of wealth: students' time and attention and their potential as individuals and as communities. These are extremely valuable and worth protecting. When tech companies or tech-funded philanthropic organizations come sniffing around with 'solutions' predicated on everyone using their software, these solutions are never cost-free in the bigger picture. If there isn't time to thoroughly examine the costs, then the default position can and should be, 'No, thank you'.

References

Altman, S. 2024. *The intelligence age*. Samaltman.com. https://ia.samaltman.com (Accessed 5 August 2025.)

Baker-White, E. 2025. *These AI tutors for kids gave fentanyl recipes and dangerous diet advice*. Jersey City, Forbes. https://www.forbes.com/sites/emilybaker-white/2025/05/12/these-ai-tutors-for-kids-gave-fentanyl-recipes-and-dangerous-diet-advice (Accessed 5 August 2025.)

Banks, A. 2025. *Learning will never be the same*. Sunnyvale, LinkedIn. https://www.linkedin.com/posts/a-banks_learning-will-never-be-the-same-people-activity-7340710510121611264-891j (Accessed 5 August 2025.)

Bender, E. M. and Hanna, A. 2025. *The AI Con: How to Fight Big Tech's Hype and Create the Future We Want*. New York, HarperCollins.

Clark, H. H. 1996. *Using Language*. Cambridge, Cambridge University Press.

Gilliard, C. 2025. Surveillance scholar Chris Gilliard on Facebook's spy glasses. *Remarks on Techtonic with Mark Hurst*. Jersey City, WFMU [podcast]. https://www.wfmu.org/playlists/shows/152175 (Accessed 29 July 2025.)

Google. 2024. *Our approach to Search*. Mountain View, Alphabet. https://www.google.com/intl/en_us/search/howsearchworks/our-approach (Accessed 5 August 2025.)

Hoffmann, A. L. and Bloom, R. 2016. Digitizing books, obscuring women's work: Google books, librarians, and ideologies of access. *Journal of Gender, New Media, and Technology*, Vol. 9. Eugene, University of Oregon. https://hdl.handle.net/1794/26769 (Accessed 5 August 2025.)

Huddleston Jr., T. 2025. *Bill Gates: Within 10 years, AI will replace many doctors and teachers – Humans won't be needed 'for most things'*. Englewood Cliffs, CNBC. https://www.cnbc.com/2025/03/26/bill-gates-on-ai-humans-wont-be-needed-for-most-things.html (Accessed 5 August 2025.)

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I. and Maes, P. 2025. *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. Ithaca, arVix. https://doi.org/10.48550/arXiv.2506.08872 (Accessed 5 August 2025.)

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P. F., Leike, J. and Lowe, R. 2022. *Training language models to follow instructions with human feedback*. Ithaca, arVix. https://doi.org/10.48550/arXiv.2203.02155 (Accessed 5 August 2025.)

Reddy, M. J. 1979. The conduit metaphor: A case of frame conflict in our language about language. *Metaphor and Thought*. Cambridge, Cambridge University Press, pp. 164–201.

Rhodes, C. 2025. *Stinking Rich: The Four Myths of the Good Billionaire*. Bristol, Bristol University Press.

Shah, C. and Bender, E. M. 2024. Envisioning information access systems: What makes for good tools and a healthy web? *ACM Transactions on the Web*, Vol. 18, No. 3. New York, Association for Computing Machinery (ACM). https://doi.org/10.1145/3649468 (Accessed 5 August 2025.)

Contested imaginaries: Reclaiming higher education in the age of AI

Markus Deimann and Robert Farrow

The emergence of AI as a force within education compels a re-examination not only of pedagogical practices but also of the very futures we imagine as possible or desirable. Across global contexts, AI gives rise to conflicting imaginaries: on one hand, for example, utopian visions of personalized learning, expanded access and educational equity aligned with the United Nation's Sustainable Development Goal 4; on the other hand, dystopian fears of deepened inequalities, the erosion of human judgement and the commodification of learning. These imaginaries are not passive reflections of technological trends; they are active forces, shaping policies, priorities and values in real time. In times of profound volatility and disruption, the educational project becomes inseparable from the ethical task of imagining and enacting alternative futures. While imaginaries such as techno-solutionism permeate multiple domains of society, including health, governance and the media, we argue that their manifestation in higher education deserves particular attention. Here, imaginaries intersect with institutional traditions, academic governance and pedagogical values in unique ways that differentiate them from broader societal contexts (Rahm, 2023).

This think piece critically interrogates how AI imaginaries influence educational trajectories and asks how education might reclaim its narrative agency – not merely adapting to technological change, but asserting normative visions of justice, dignity and sustainability in an AI-mediated world.

An indicator for the deep belief in technologies to bring about progress and prosperity is the current hype around generative AI, with the provision of ChatGPT and other tools seeming to usher in an 'AI revolution', with a linear development path, driven by ever larger models and more training data. However, this revolution is now at a turning point because the underlying resources (data, energy) are being exhausted and new strategies for progress are needed (Jones, 2024).

Imaginaries: Intersection of technology and ideology

Technologies such as smartphones, social platforms, cloud computing and algorithmic decision-making now mediate communication, commerce, entertainment, governance and education in ways that are often seamless and subterranean. Businesses have reoriented themselves around data-driven models and global digital networks; social and political movements increasingly unfold within online spaces. From a user perspective, it can be easy to forget that digital media not only changes what we do, but it also subtly shapes how we imagine ourselves and our futures.

Imaginaries have been defined as 'collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology' (Jasanoff and Kim, 2015). The plasticity (Malabou, 2022) and ubiquity of AI means it can easily function as a screen on which to project normative visions of the future.

Under 'late capitalism' – characterized by economic and political disruption and volatility – conflicting imaginaries and narratives emerge as strategies for navigating uncertainty. Furthermore, research has shown how discourses on digital technologies have historically been constructed around myths in order to normalize dominant ideologies by making alternative perspectives invisible and existing power relations appear 'natural' (Bory, 2020; Mosco, 2005).

Furthermore, we observe that technology – particularly as platformed in social media – is an increasingly important vector for political sentiment. It is becoming more difficult to think of the challenges around AI in education as isolated from wider political concerns. Many AI advocates share ideological and explicitly political commitments, including technological optimism, a belief in free market libertarianism, solutionism (Morozov, 2013) and dataism (Adamczyk, 2023; Van Dijck, 2014). For some, this extends to a belief in messianic 'general AI', which transcends the irrationalities of a 'merely human' perspective. Such views have been criticized as reflecting and reproducing forms of gender bias, often privileging masculinist, rationalist and universalist perspectives, while marginalizing feminist and care-based epistemologies (Gebru and Torres, 2024).

Within this context, AI is often highlighted as a driver of both utopia and dystopia (Brevini, 2020) and is now presented – sometimes in the same breath – as humanity's greatest hope and most formidable threat (Warburton, 2023). This can lead us to ask: rather than treating AI imaginaries as abstract or symbolic, we propose interrogating their political function. Which liminal or speculative futures are being mobilized – and in service of which educational, economic, or ideological agendas? For instance, the idea of a 'teacherless classroom' not only evokes a seductive image of efficiency and scalability, but also aligns with cost-cutting imperatives in neoliberal education systems and the expansion strategies of AI-powered platforms.

Before we present a table depicting AI imaginaries that are prevalent in education, we first lay the conceptual groundwork to ensure that the AI imaginaries it presents are read not as neutral categories, but as politically charged constructs embedded in broader ideological projects. This structure allows the table to serve as a critical mapping device rather than a surface-level classification.

Cyberlibertarian imaginaries

In light of the aforementioned context, solutionism, market-orientation and the unimpeded flow of information and knowledge have served as pivotal tenets within discourses concerning the future of higher education. However, this discourse has predominantly overlooked the antithetical aspects of cyberlibertarianism, which are presently intertwined with escalating rhetorics of authoritarianism and fascism (Klein and Taylor, 2025). Cyberlibertarian imaginaries have gained enormous momentum since the beginning of the second term of United States President Trump. According to David Golumbia (2024, p. 11),

> Cyberlibertarianism is a commitment to the belief that digital technology is or should be beyond the oversight of democratic governments – meaning democratic political sovereignty. Frequently, the sentiment can be reduced to the view that democratic governments cannot or must not regulate the internet – or to flip this formulation

> on its head, that the internet should be a place to which laws do not (or cannot) apply.

This is a form of internet exceptionalism, proclaiming it as a digital space outside of legal regulations and with profound implications for many areas of society, including education. For quite some time, digital education has been to a growing extent grounded in cyberlibertarian thinking, as argued, for instance, by Neil Selwyn (2024) with reference to the EdTech industry, which is using the pretext of 'educational innovation' to enforce market logic – a central feature of cyberlibertarianism.

From the cyberlibertarian perspective, educational institutions exist to be subverted and bypassed since pedagogy can be reduced to a matter of product and consumer preference. The employment of elements of solutionism posits the notion that technologies are inherently beneficial for educational purposes and can be found across a wide range of educational domains. The present moment is characterized by the profound materialization of perilous ideologies, as evidenced by recent attacks on universities (notably in the United States). In light of this, it is imperative to advance counter-narratives that are firmly rooted in democratic principles and to harness the collective power of the people, as opposed to the influence of libertarian billionaires.

Navigating educational imaginaries

Higher education is constrained by socio-technical imaginaries, as posited by scholars such as Lina Friesen (2020) and Norm Rahm (2023), as well as by AI-futurism (Schütze, 2024) and an overarching technological determinism (Karp and Zamiska, 2025). Nevertheless, it remains important to hypothesize about educational imaginaries that are grounded in democratic principles and oriented towards navigating turbulent periods within higher education.

AI is widely framed as a technology of interest for educators and learners, with some commentators suggesting it may offer new possibilities for teaching, learning and access (Bozkurt et al., 2023; Ossiannilsson et al., 2024). However, such claims often remain speculative and are rarely substantiated by empirical evidence, and the language of 'potential', although common, tends to obscure this (UNESCO, 2023). In fact, AI development in education has followed a 'messy', non-linear path of 'one step forward, two steps back' (Selwyn, 2022), raising doubts about its transformative potential.

AI, construed as a 'medicine for the sick patient of education' (Higgin, 2024), is entering classrooms and lives regardless of the preferences of educators. There is no easy response to such disempowerment, which compounds a sense of existential crisis. Yet, it is precisely in moments of disruption and loss of control that educators must actively and critically engage with the imaginaries that shape AI's role in education, challenging dominant narratives, proposing alternative futures and cultivating collective agency. Educational imaginaries need not be captive to technological determinism or market logic. Instead, they can be deliberately constructed around values of inclusion, justice, sustainability and care. This involves reasserting the centrality of pedagogy, resisting reductive metrics of learning, and embracing complexity and pluralism (Nørgård, 2021).

Contrasting AI imaginaries in education

We briefly describe some of the key imaginaries that are prevalent in current discourses.

Table 1: Key imaginaries prevalent in current discourses

Imaginary/Ideology	Description	Fields of contestation
Utopian AI for education/ techno-solutionism, progressivism	AI as a tool for personalized learning, expanded access and educational equity, especially aligned with SDG 4. The belief that AI can democratize education by tailoring content to individual needs.	Access vs. equity: does AI address structural inequalities or deepen them? Personalization vs. standardization: can AI effectively replace traditional teaching methods?
AI as the perfect educator (Friesen, 2020)/Automation utopianism, post-humanism	AI as an autonomous entity that can teach, evaluate and even tutor students, ideally improving educational outcomes with minimal human intervention.	Human agency vs. automation: what role do human teachers play in this scenario? Teacher-student relationship: does AI undermine or enhance the personal connection? Gendered neutrality: does the imagined AI educator reproduce masculine-coded ideals of efficiency, objectivity and authority?
Techno-solutionism in education/neoliberalism	The belief that technology, including AI, inherently solves educational problems like access, quality and scale. This narrative often accompanies the rise of massive open online courses (MOOCs) and EdTech platforms.	Market logic vs. public good: is education being commodified? Quality vs. quantity: does technology enhance educational quality or merely increase access without addressing systemic issues?
Cyberlibertarianism in education/market fundamentalism	The ideology that digital technologies, including AI, should operate outside the boundaries of governmental regulation, advocating for a free market of educational tools, embedded in an authoritarian setting.	Regulation vs. autonomy: can we regulate AI without stifling innovation? Democracy vs. privatization: who controls education when it is governed by tech companies?
Dystopian AI and the surveillance state/authoritarian techno-governance	A fear that AI will be used to surveil, control and manipulate educational environments, replacing genuine learning with data-driven compliance.	Surveillance vs. autonomy: does AI infringe on student privacy? Freedom vs. control: how can education resist the rise of algorithmic governance?
AI as ecological catastrophe/ critical environmentalism	The recognition that AI's promises might be unsustainable due to the vast data and energy requirements, raising concerns about its environmental and ethical costs.	Sustainability vs. profit: is the pursuit of AI's potential contributing to environmental degradation? Ethical deployment: how can we balance innovation with ecological responsibility?

Conclusion: Operating through tensions

It is essential to recognize the conflicting imaginaries influencing contemporary education, political discourse and the wider digital information sphere. Imaginaries matter because they guide what we invest in, what we fear and what we believe is possible. Imaginaries manifest in strategies, rules or action plans. In confronting the competing imaginaries surrounding AI in higher education, educators are faced with a critical ethical and political task: to navigate technological change as active shapers of its meaning, direction and impact. This means holding space for tension, while resisting the lure of deterministic and polarizing narratives. To act with such tensions in mind is to insist that the future of education remains open, contestable and grounded in democratic values.

When we think in terms of AI imaginaries, the difficulty of reflection in a fast-paced world comes into sharper relief. AI technologies evolve faster than our ability to measure, contextualize or regulate them. Social, ethical and political consequences are still emerging, yet decisions must be made in real time, often on incomplete evidence. This creates a tension between the need for critical distance and the urgency of action. More than ever, it is essential to create space for reflection, discussion and the iteration of adaptive interpretative frameworks that can be shared through fora, where diverse voices co-create and revise our collective futures. Too often, the framing of discourse around AI is set by technology companies and venture capitalists. By ensuring that the shifting imaginaries of AI are shaped through gender-responsive, environmentally sustainable and globally inclusive dialogue, we can potentially create technological futures that challenge bias, empower women and marginalized groups, protect the planet and strengthen responsible citizenship.

References

Adamczyk, C. L. 2023. Communicating dataism. *Review of Communication*, Vol. 23, No. 1. London, Francis & Taylor, pp. 4–20.

Bory, P. 2020. *The Internet Myth: From the Internet Imaginary to Network Ideologies*. London, University of Westminster Press. https://doi.org/10.16997/book48 (Accessed 18 August 2025.)

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A., Toquero, C. M., Singh, L., Tlili, A., Lee, K., Nichols, M., Ossiannilsson, E., Brown, M., Irvine, V., Raffaghelli, J. E., Santos-Hermosa, G., Farrell, O., Adam, T., Thong, Y. L., Sani-Bozkurt, S., Sharma, R. C., Hrastinski, S. and Jandrić, P. 2023. Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, Vol. 18, No. 1. New Delhi, Asian Society for Open and Distance Education (ASODE), Japan. https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/709. (Accessed 18 August 2025.)

Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y. H., Nerantzi, C., Moore, S., Dron, J., Stracke, C. M., Singh, L., Crompton, H., Koutropoulos, A., Terentev, E., Pazurek, A., Nichols, M., Sidorkin, A. M., Costello, E., Watson, S., Mulligan, D., Honeychurch, S., Hodges, C. B., Sharples, M., Swindell, A., Frumin, I., Tlili, A., Slagter van Tryon, P. J., Bond, M., Bali, M., Leng, J., Zhang, K., Cukurova, M., Chiu, T. K. F., Lee, K., Hrastinski, S., Garcia, M. B., Sharma, R. C., Alexander, B., Zawacki-Richter, O., Huijser, H., Jandrić, P., Zheng, C., Shea, P., Duart, J. M., Themeli, C., Vorochkov, A., Sani-Bozkurt, S., Moore, R. L. and Asino, T. I. 2024. The manifesto for teaching and learning in a time of generative AI: A critical collective stance to better navigate the future. *Open Praxis*, Vol. 16, No. 4. Oslo, International Council for Open and Distance Education (ICDE), pp. 487–513. https://doi.org/10.55982/openpraxis.16.4.777 (Accessed 18 August 2025.)

Brevini, B. 2020. Black boxes, not green: Mythologizing artificial intelligence and omitting the environment. *Big Data & Society*, Vol. 7, No. 2. Thousand Oaks, SAGE Publications, pp. 1–5. https://doi.org/10.1177/2053951720935141 (Accessed 18 August 2025.)

Downes, S. 2024. *What is the soul of open educational resources?* Ottawa, Half an Hour. https://halfanhour.blogspot.com/2024/08/what-is-soul-of-open-educational.html (Accessed 18 August 2025.)

Friesen, N. 2020. The technological imaginary in education: Myth and enlightenment in 'personalized learning'. M. Stocchetti (ed.), *The Digital Age and Its Discontents: Critical Reflections in Education*. Helsiniki, Helsinki University Press, pp. 141–160. https://doi.org/10.33134/HUP-4-8 (Accessed 18 August 2025.)

Gebru, T. and Torres, É. P. 2024. The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence. *First Monday*, Vol. 29, No. 4. Chicago, First Monday. https://doi.org/10.5210/fm.v29i4.13636 (Accessed 18 August 2025.)

Golumbia, D. 2024. The critique of cyberlibertarianism. *boundary 2*, Vol. 51. No. 2. Durham, Duke University Press, pp. 5–18.

Higgin, T. 2024. It's time to ditch the idea of EdTech disruption: But what comes next? *EdSurge*. https://www.edsurge.com/news/2024-04-12-it-s-time-to-ditch-the-idea-of-edtech-disruption-but-what-comes-next (Accessed 18 August 2025.)

Jasanoff, S. and Kim, S.-H. 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power.* Chicago, University of Chicago Press.

Jones, N. 2024. The AI revolution is running out of data: What can researchers do? *Nature,* Vol. 636. Berlin, Springer Nature, pp. 290–292. https://doi.org/10.1038/d41586-024-03990-2 (Accessed 18 August 2025.)

Karp, A. C. and Zamiska, N. W. 2025. *The Technological Republic: Hard Power, Soft Belief, and the Future of the West*. New York, Crown Currency.

Klein, N. and Taylor, N. 2025. The rise of end times fascism. *The Guardian*. https://www.theguardian.com/us-news/ng-interactive/2025/apr/13/end-times-fascism-far-right-trump-musk (Accessed 18 August 2025.)

Malabou, C. 2022. *Plasticity: The Promise of Explosion*. Edinburgh, Edinburgh University Press.

Mejias, U. A. and Couldry, N. 2024. *Data Grab: The New Colonialism of Big Tech and How to Fight Back*. Chicago, University of Chicago Press.

Morozov, E. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York, PublicAffairs.

Mosco, V. 2005. *The Digital Sublime: Myth, Power, and Cyberspace*. Cambridge, MIT Press.

Nørgård, R. T. 2021. Theorising hybrid lifelong learning. *British Journal of Educational Technology*, Vol. 52, No. 4. Hoboken, John Wiley & Sons, pp. 1709–1723. https://doi.org/10.1111/bjet.13121 (Accessed 18 August 2025.)

Ossiannilsson, E., Ulloa Cazarez, R. L., Goode, C., Mansour, C. and De Gusmão, C. M. G. 2024. Artificial intelligence use to empower the implementation of OER and the UNESCO OER recommendation. *Open Praxis*, Vol. 16, No. 2. Oslo, International Council for Open and Distance Education (ICDE), pp. 237–257. https://doi.org/10.55982/openpraxis.16.2.650 (Accessed 18 August 2025.)

Rahm, L. 2023. Educational imaginaries: Governance at the intersection of technology and education. *Journal of Education Policy*, Vol. 38, No. 1. London, Taylor & Francis, pp. 46–68. https://doi.org/10.1080/02680939.2021.1970233 (Accessed 18 August 2025.)

Schütze, P. 2024. The impacts of AI futurism: An unfiltered look at AI's true effects on the climate crisis. *Ethics and Information Technology*, Vol. 26, No. 23. Berlin, Springer Nature. https://doi.org/10.1007/s10676-024-09758-6 (Accessed 18 August 2025.)

Selwyn, N. 2022. The 'wonderful usefulness' of historical perspectives on EdTech. *Critical Studies of Education & Technology*. https://criticaledtech.com/2022/03/03/the-wonderful-usefulness-of-historical-perspectives-on-edtech (Accessed 18 August 2025.)

UNESCO. 2023. *Global education monitoring report, 2023: technology in education: a tool on whose terms?* Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385723 (Accessed 18 August 2025.)

Van Dijck, J. 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, Vol. 12, No. 2. Chapel Hill, Surveillance & Society, pp. 197–208. https://doi.org/10.24908/ss.v12i2.4776 (Accessed 18 August 2025.)

Warburton, A. 2023. *The wizard of AI* [Video]. alanwarburton.co.uk. https://alanwarburton.co.uk/thewizardofai (Accessed 18 August 2025.)

Wiley, D. 2024. *Why open education will become generative AI education* [Video]. improving learning. https://opencontent.org/blog/archives/7612 (Accessed 18 August 2025.)

Wulf, C. 2003. The dream of education. *Journal of Curriculum Studies*, Vol. 35, No. 3. London, Informa UK, pp. 263–280.

4. AI pedagogies, assessment and emerging educational futures

The incomputable classroom: The limits and dangers of AI in education

Abeba Birhane

In a 1985 interview, Joseph Weizenbaum, the creator of the first chatbot, ELIZA, was asked about the benefits of having computers in the classroom. Weizenbaum dismissed the question as 'wrongheaded' and 'upside-down', loaded with unwarranted assumptions. If the betterment of education is at stake, Weizenbaum replied, then the question should begin with 'what education should accomplish and what the priorities should be' and not 'how computers can be used in the classroom' (Weizenbaum, 1985). The field of AI has gone through some of the most advanced breakthroughs over the four decades since this interview. This think piece 1) draws on Paulo Freire's (1970) critical pedagogy and embodied and enactive approaches to cognitive science to illustrate the rich, dynamic and relational nature of learning, and 2) reviews existing work that highlights the limitations, drawbacks and harms of generative AI in classroom settings to 3) call on policy-makers, educators and educational institutions to challenge and resist uncritical use of generative AI in the classroom.

The release of ChatGPT has brought generative AI into the mainstream. Subsequently, individual educators and learners alike are increasingly using the technology for various purposes, including writing, code generation and assignment evaluation. This is, in part, propelled by an aggressive and concerted push by for-profit companies like Google and OpenAI. OpenAI envisions that 'over time, A.I. would become part of the core infrastructure of higher education,' according to Leah Belsky, their vice president of education (Singer, 2025). Government bodies view generative AI as a powerful instrument that offers economic competitiveness. For large tech corporations, AI companies and start-ups alike, generative AI is a powerful tool that enables market dominance and monopoly. Thus, the vendors market generative AI as a powerful technology with the capacity to solve a multitude of societal, structural, political and economic challenges, the most attractive selling point being scale and cost reduction. This think piece is particularly focused on systematic integration of AI as opposed to individual uses.

Let us take a brief look at what generative AI is to better understand why it is a tool that flattens out complexity, is inherently backward-looking and thus presents a danger to learning. Generative AI is undoubtedly a remarkable engineering achievement. Trained on vast amounts of data, transformer-based models 'learn' statistical patterns and token relationships to produce human-like outputs. Leveraging a mechanism known as 'self-attention', these models iteratively select a path through a vast space of potential token sequences based on 'learned' probabilities to construct plausible and coherent sequences of tokens, rendered as text in the case of language generation. A fundamental aspect of how large language models (LLMs) function is 'learning' historical usage, patterns and structures. This renders the technology backward-looking, whereby seemingly newly

generated text is predicated on patterns of the past. Built on the backbone of these principles, services such as 'personalized tutoring', in essence, amount to tailoring the future educational needs of a learner on the basis of the past.

Furthermore, text generation in LLM is not grounded in factual real world knowledge but instead based on probabilistic sampling from token distribution. The fluent, seemingly coherent and plausible text generated can easily be factually incorrect, entirely made-up or misleading, known as an 'hallucination'. This means that without a domain expert to validate LLM output, there is no way of distinguishing factually accurate from incorrect or made-up information.

Google and OpenAI are at the forefront of promoting AI as a solution to some of the biggest challenges in education. According to Google's marketing information, 'In education, AI can be used to do helpful things like make learning experiences more personal, provide immediate feedback, improve accessibility, enhance digital security, give educators precious time back and so much more' (Google, 2024*a*). Similarly, OpenAI claims, 'ChatGPT can help with various tasks across campus, such as providing personalized tutoring for students and reviewing their resumes, helping researchers write grant applications, and assisting faculty with grading and feedback' (Open AI, 2024). These claims lack rigorous empirical evidence, as we see below. Tech companies view AI as a 'solution' to education.

But what is education? Embodied and enactive cognitive sciences remind us that knowledge is not something discrete that an individual possesses and passes down but an inherently dynamic and evolving endeavour that develops in the process of embodied, curious and engaged interaction through dialogue. The pinnacle of cognition, particularly 'human knowing', is inextricably interwoven with interactions we engage in with each other and the physical, cultural and social world we inhabit, 'so much so that individuals are not thinkable outside of their interactions and embeddedness in their (social) world' (De Jaegher, 2019). Dialogic models of knowledge and education emphasize that interactions between a student and teacher and/or peer provide 'scaffolding' for how that child understands the world. The always in flux, active and continually transforming nature of human cognition necessitates that education be fundamentally an ongoing activity. Far from the reductionist view whereby 'formal knowledge' can be packaged and acquired from an LLM, the classroom is an environment where love, trust, empathy, care and humility are fostered and mutually cultivated through dialogical interactions.

Contrary to common conceptions of education where the student is a passive recipient of knowledge or an empty vessel that memorizes teacher-narrated content mechanically, education is a critical intervention. In the *Pedagogy of the Oppressed*, Freire (1970) emphasizes that the most important object of education is for the student to come to 'feel like master of her own thinking'. Education is thus a means for learners to develop the skills needed for self-assertion and the vocabulary to critically reflect on the object of knowledge and deal with it in the real world. Writing, for example, is not a neutral and passive generation of textual content, but an active and engaged process of structuring and articulating our thoughts, values, reflections and positionality.

The teacher-student dialogue is not an end in itself but a means to develop comprehension and articulation of the object of knowledge. According to Freire (1970), 'There is no such thing as a neutral educational process.' Education functions as an instrument for critical intervention. In contemporary societies, it is instrumental for dealing with current issues that children are growing

up with. This includes how to identify and manage, for example, mis/dis-information, harmful ideologies and anti-science sentiments that have become part and parcel of the current digital knowledge ecosystem.

A genuine and in-depth engagement of cognition, knowledge and education reveals that education is a thoroughly complex, relational, dynamic and interpersonal human endeavour. Although generative AI is marketed as a substitute for deep, meaningful human interaction, these qualities of education defy datafication, measurement and automation. The now deep embeddedness of AI in our day-to-day lives is often mistakenly perceived as AI technologies being extremely capable, fully autonomous 'agents like ourselves' (Birhane et al., 2024). Such anthropomorphic claims are often predicated on simplified and reductive framing of human cognition in machinic terms, while over-inflating the capabilities of AI tools (Birhane and Van Dijk, 2020). Unlike a human being, there is nothing at stake for a generative AI model. It cannot feel a sense of loss, embarrassment, accomplishment or care towards a student, as human teachers do.

Furthermore, as a relatively new technology, the downstream impact of generative AI on the teaching and learning process is not thoroughly researched and understood. Generative AI products, tools and services from Big Tech are released into the world without appropriate vetting and rigorous testing, especially of their impact on children and teens' cognitive, emotional and social development. In the long term, generative AI is likely to inhibit children's cognitive, social and critical thinking development and skills. Emerging research that examines the use of generative AI in education shows that it can hinder the development of independent thinking and promote 'learners' dependence on technology' (Fan et al., 2025); erode the capacity for critical thinking and creativity (Lee et al., 2025); and leads to diminished problem-solving skills development in the context of practical problem-solving tasks (Bastani et al., 2025). In addition, given the reliance on historical data underlying generative AI models, they have a tendency to encode perpetual inequity in the classroom, particularly towards the historically marginalized. A recent survey, for example, found that 'LGBTQ+ students and students that have been disciplined in the past year' experience far more negative effects from activity monitoring compared to their peers (CDT, 2025). In fact, instead of creating an environment for learning, AI systems used for student assessment and e-proctoring have consistently automated poor pedagogical practices, creating unworkable roadblocks, especially for students at the margins (Holtermann, 2025*a*; Mintz et al., 2023). These themes closely align with independent audits of LLMs that have identified recurring patterns: generative AI tends to encode and exacerbate societal stereotypes, racism and otherwise discriminatory and marginalizing norms and practices (Kotek et al., 2023; Hofmann et al., 2024). There is no reason to believe that generative AI used to assess students assignments or develop 'personalized' tutoring can be void of these harmful and dangerous trends.

Major selling points – scale and cost reduction – similarly suffer from a lack of thorough empirical evidence demonstrating concrete benefits. Instead, accelerated adoption of generative AI in the classroom is often driven by marketing public-relations rhetoric and fabricated fear of being left out. According to Google's marketing material *Generative AI for Educators*, for example, 'in a rapidly evolving world, our teachers cannot afford to fall behind in accessing powerful generative tools that will help them develop new approaches to teaching and learning' (Google, 2024*b*). Following analysis of data from 25,000 workers and 7,000 workplaces in Denmark, Anders Humlum and Emilie Vestergaard (2025) show that, despite

widespread adoption, generative AI has not resulted in a 'significant impact on earnings or recorded hours in any occupation'.

Genuine engagement with major challenges in education reveals a host of multifaceted contingent factors or 'ugly social realities', in Weizenbaum's words. These include lack of sufficient funding and resources, large class sizes and an ever-widening systemic inequity. In his 1985 interview, Weizenbaum lamented that instead of confronting these ugly social realities, it is 'much more comfortable, to have some device, say the computer, with which to flood the schools with, and then to sit back and say, "You see, we are doing something about it"' (Weizenbaum, 1985). Presently, by integrating generative AI into the teaching and learning process, we are not only misleading ourselves into thinking we are tackling education's challenges, but we are also outsourcing a social, civic and democratic process of cultivating the coming generation to commercial and capitalist enterprises whose priority is profit. The AI industry has consistently demonstrated that it is willing to undo decades of progress towards fundamental rights (Peters, 2025), peace (Singh, 2025) and democracy (Zeff, 2025) and can easily close down education efforts (Holtermann, 2025b) by operating in an environment with little to no oversight and riddled with conflicts of interest (Menn and Nix, 2023), in accordance with these companies' business models: maximization of profit and concentration of power. An industry with such an extractive business model, developing and deploying generative AI tools with little oversight and testing, cannot be a good steward for the betterment of education. Instead, integrating generative AI into education will result in commercialization of a collective responsibility: fostering and cultivating the coming generations.

Policy-makers, educational institutions and educators alike are better off organizing and collectively challenging, resisting and pushing back against the systemic integration into and uncritical use of generative AI in the classroom. Such collectives of resistance are increasingly found across institutions, as illustrated by the *Stop the Uncritical Adoption of AI Technologies in Academia* (Open Collective, 2025) open letter, for instance. This should hold until the technology is conclusively shown to pose no risks or harm, and meaningful independent regulatory oversight and enforcement mechanisms are put in place.

References

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö. and Mariman, R. 2025. Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 122, No. 26. Washington DC, National Academy of Sciences.

Birhane, A. and Van Dijk, J. 2020. Robot rights? Let's talk about human welfare instead. AIES '20: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. New York, Association for Computing Machinery (ACM), pp. 207–13. https://doi.org/10.1145/3375627.3375855 (Accessed 5 August 2025.)

Birhane, A., Van Dijk, J. and Pasquale, F., 2024. Debunking robot rights metaphysically, ethically, and legally. *First Monday,* Vol. 29, No. 4. Chicago, First Monday. https://doi.org/10.5210/fm.v29i4.13628 (Accessed 5 August 2025.)

CDT. 2025. *Out of Step: Students, Teachers in Stride with EdTech Threats While Parents Are Left Behind.* Washington DC, Center for Democracy and Technology (CDT). https://cdt.org/insights/out-of-step-students-teachers-in-stride-with-edtech-threats-while-parents-are-left-behind (Accessed 5 August 2025.)

De Jaegher, H. 2019. Loving and knowing: Reflections for an engaged epistemology. *Phenomenology and the Cognitive Sciences*, Vol. 20. Basel, Springer Nature, pp. 847–70. https://doi.org/10.1007/s11097-019-09634-5 (Accessed 5 August 2025.)

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X. and Gašević, D. 2025. Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, Vol. 56, No. 2. Hoboken, John Wiley & Sons, pp. 489–530.

Freire, P. 1970. *Pedagogia do Oprimido* [Pedagogy of the oppressed]. New York, Seabury Press. (In Portuguese.)

Google. 2024a. *A Guide to AI in Education.* Mountain View, Google. https://services.google.com/fh/files/misc/global_google_for_education_a_guide_for_ai.pdf (Accessed 5 August 2025.)

—. 2024b. *Generative AI for Educators.* Mountain View, Google. https://grow.google/ai-for-educators (Accessed 5 August 2025.)

Hofmann, V., Kalluri, P.R., Jurafsky, D. and King, S. 2024. AI generates covertly racist decisions about people based on their dialect. *Nature,* Vol. 633. Basel, Springer Nature, pp. 147–54. https://doi.org/10.1038/s41586-024-07856-5 (Accessed 5 August 2025.)

Holtermann, C. 2025a. A new headache for honest students: Proving they didn't use AI. *The New York Times*. https://www.nytimes.com/2025/05/17/style/ai-chatgpt-turnitin-students-cheating.html (Accessed 5 August 2025.)

Holtermann, C. 2025b. Zuckerberg and Chan to close tuition-free school for Bay Area kids. *The New York Times.* https://www.nytimes.com/2025/04/24/us/school-closure-zuckerberg-chan.html (Accessed 5 August 2025.)

Humlum, A. and Vestergaard, E. 2025. *Large Language Models, Small Labor Market Effects.* Cambridge, National Bureau of Economic Research (NBER). https://doi.org/10.3386/w33777 (Accessed 5 August 2025.)

Kotek, H., Dockum, R. and Sun, D. 2023, November. Gender bias and stereotypes in large language models. *CI '23: Proceedings of the ACM Collective Intelligence Conference.* New York, Association for Computing Machinery (ACM), pp. 12–24. https://doi.org/10.1145/3582269.3615599 (Accessed 5 August 2025.)

Lee, H.P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R. and Wilson, N. 2025. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems,* No. 1121. New York, Association for Computing Machinery (ACM), pp. 1–22. https://doi.org/10.1145/3706598.3713778 (Accessed 5 August 2025.)

Menn, J. and Nix, N. 2023. Big Tech funds the very people who are supposed to hold it accountable. *The Washington Post.* https://www.washingtonpost.com/technology/2023/12/06/academic-research-meta-google-university-influence (Accessed 5 August 2025.)

Mintz, J., Holmes, W., Liu, L. and Perez-Ortiz, M. 2023. Artificial intelligence and K-12 education: Possibilities, pedagogies and risks. *Computers in the Schools*, Vol. 40, No. 4. London, Taylor & Francis, pp. 325–33. https://doi.org/10.1080/07380569.2023.2279870 (Accessed 5 August 2025.)

Open Collective, 2025. Stop the Uncritical Adoption of AI Technologies in Academia. Open Collective. https://openletter.earth/open-letter-stop-the-uncritical-adoption-of-ai-technologies-in-academia-b65bba1e (Accessed 5 August 2025.)

OpenAI. 2024. *Introducing ChatGPT Edu.* San Francisco, OpenAI. https://openai.com/index/introducing-chatgpt-edu (Accessed 5 August 2025.)

Peters, J. 2025. Google Calendar removed events like Pride and BHM because its holiday list wasn't 'sustainable'. *Vox Media*. https://www.theverge.com/news/608858/google-calendar-missing-events-holidays (Accessed 5 August 2025.)

Singer, N. 2025. Inside OpenAI's plan to embed ChatGPT into college students' lives. T*he New York Times*. https://www.nytimes.com/2025/06/07/technology/chatgpt-openai-colleges.html (Accessed 5 August 2025.)

Singh, K. 2025. OpenAI wins $200 million US defense contract. *Reuters*. https://www.reuters.com/world/us/openai-wins-200-million-us-defense-contract-2025-06-16 (Accessed 5 August 2025.)

Weizenbaum, J. 1985. Weizenbaum examines computers and society. *The Tech*, Vol. 105, No. 16. Cambridge, MIT. https://web.archive.org/web/20090325194547/https://tech.mit.edu/V105/N16/weisen.16n.html (Accessed 5 August 2025.)

Zeff, M. 2025. Google removes pledge to not use AI for weapons from website. *TechCrunch Media.* https://techcrunch.com/2025/02/04/google-removes-pledge-to-not-use-ai-for-weapons-from-website (Accessed 5 August 2025.)

Challenging hyper-personalization: Towards (re-)socializing learning in human-to-human-to-machine dialogue

Carla Aerts

Introduction

In an increasingly digitized world, the rise of generative AI and large language models (LLMs) is accelerating the conversation on personalization in education. These technologies promise to improve learning outcomes by tailoring learning experiences to individual needs, preferences and context.

Every learner is unique. Our brains, though sharing the same structure, develop distinct neural pathways. They process and retrieve information in unique ways. This uniqueness presents a compelling argument for individualized learning, especially as AI seems ideally suited to support individual learning models, allowing for hyper-personalization.

But should education embrace AI-powered hyper-personalization as its future? We must pause and count the cost of hyper-individualization. Couldn't an alternative for the future of education be considered; one that fosters uniqueness and promotes learner agency, honouring our social nature and values?

Since antiquity, learning has occurred in social contexts as a social activity. This is not unique to the Western world; it transcends cultural boundaries. Coming from different contexts and focusing on differing objectives for education, education philosophers and theorists, including Dewey (1916), Vygotsky (1978), Piaget (1932), Freire (2007) and, more recently, Wegerif (2025), have highlighted the importance of social interaction for learning and the construction of knowledge through social participation and engagement in dialogue.

Rather than intensifying our focus on personalization and the possibility of AI-enabled hyper-personalization increasingly sidelining the social dimension of learning, we should consider a new re-socialized learning paradigm. This should nurture the social essence and fabric of learning, our social nature and support the development of collective intelligence, an intelligence that surpasses the individual but that is developed and achieved through social engagement and dialogue, but allows for personalization within a socialized context.

A collective personalization approach would recognize uniqueness and agency, while furthering human-to-human dialogue, enriched and extended by meaningful, critical and reflective AI interaction. This model promises a more balanced future of education, in which technology not only amplifies individual learning but also holistically augments our social connection, as well as our interaction with new forms of inorganic machine intelligence.

Towards hyper-personalization: Personalization demystified

The concept of personalized education predates technology. Technology-driven personalized learning and tutoring began with the arrival of B.F. Skinner's adaptive 'Teaching Machine' in the 1960s. This mechanical device delivered programmed adaptive instruction and feedback. Computer-assisted instruction in the 1990s established early digital personalization, while educational technology (EdTech) advances in machine learning and AI, after 2010, have produced increasingly sophisticated adaptive platforms, such as

Khan Academy[1] and DreamBox,[2] shifting from rule-based systems to complex algorithmic prediction models and systems that use vast datasets and even LLMs.

However, decades of personalization implementation have revealed limitations, despite EdTech's positive claims about its benefits. The promise of the impact on learning outcomes and the ability to transform education remain unrealized. Studies demonstrate that personalized learning systems are not always effective, fail to promote equitable learning and do not fully address learner needs or context.

These personalized systems can create learning echo chambers, resulting from algorithmic reinforcement and misalignment of individual learner needs and context. Similarly, excessive focus on personalized pathways could lead to incremental amplification of learners' misunderstandings of learning constructs by systems focused on knowledge dissemination and attainment. This ignores the fundamental social dimension in learning, leading to an over-individualized and isolated learning experience and lack of social engagement (Cosmos Institute and Morris, 2025).

As AI systems become more sophisticated, they may steer learners deeper into personalized echo chambers. These could amplify their learning weaknesses and misconceptions. limit exposure to diverse perspectives, and their understanding of critical concepts for knowledge construction and learning transfer. Such systems do not sufficiently support the development of self-regulation, critical thinking and cognitive skills in response to the learner's context and social environment (Mohammed, 2017).

AI-driven platforms rely on vast datasets and complex models to generate algorithmic recommendations based on learner interactions and behaviours. These systems operate on statistical probabilities and pattern recognition, rather than nuanced understanding of learners' cognitive abilities, affect or learning contexts. They fail to account for the social nature of learning.

Generative AI and LLMs represent a fundamental shift in AI. This shift particularly impacts on personalized learning, enabling dynamic, open-ended dialogue, rather than relying on predetermined adaptive responses. These systems can simulate highly personalized human-like behaviours through conversational feedback and scaffolding at scale. Their sophistication and dynamic learning capabilities could result in what we term hyper-personalization: an exponential and dynamic amplification of personalization. This form of personalization is continuously responsive to inputs and statistical recommendations that inform and drive individual learner journeys. Such hyper-personalized capabilities would significantly surpass traditional and current adaptive personalized learning systems' capabilities, but they come with significant risk.

The risks of hyper-personalization: AI bias and human learning

As AI companies expand the memory capabilities of LLMs, the evolution towards hyper-personalization and ever-increasing individualization is well underway (Criddle, 2025). As these developments intensify, the integration of generative AI into education systems is actively promoted by the big AI players who stand to benefit from their adoption in education.

1. See https://www.khanacademy.org
2. See https://www.dreambox.com

Despite the promise of these technologies to meet individual learning needs, introducing them into education settings requires careful consideration. Their introduction is likely to present significant risk, given their technical capabilities, rather than deeper insight into education and learning contexts, and the social fabric of education.

Moreover, due to the dominance of US tech giants, the development of AI personalization tools and platforms is shaped by prevalent biases in the Global North. White, male, English-speaking and American-centric individualistic culture and perspectives prevail. Although the People's Republic of China is rapidly advancing its AI capabilities in the race for AI supremacy, these AI biases will continue to shape AI systems globally.

Personalization and hyper-personalization introduce new forms of inequality that extend beyond problems associated with access, infrastructure and connectivity. These challenges are especially acute in the Global South among ethnic minorities. However, they affect resource-poor and less-individualized contexts both within the Global South and the Global North.

AI systems, by design, lack understanding of context, culture and the importance of social fabric for learning. They are ill-equipped to recognize contextual factors or respond to culturally relevant learner needs or their social environment, nor do they understand the complexities of learner engagement and emotion.

Ironically, the very features that make AI-powered hyper-personalization possible may well introduce effects opposite to what personalization claims to realize. They could easily erode contextualized, cultural relevance and social learning support. These critical considerations are too often overlooked when considering AI technologies and personalization for the futures of education and necessitate pressing dialogue with Big Tech, whose primary focus remains on consumers and enterprise, rather than on learners and education.

Personalization: Social erosion of learning

The importance of the social dimension of education has been recognized by notable education theorists, philosophers, pedagogues and practitioners. Yet, the focus of current education systems on measuring individual learner performance, oriented to high-stakes assessment and exams, has diminished attention on social importance in education and learning, which typically take place in the social context of the classroom.

A key lesson from the Covid-19 pandemic is the critical role of socialization for learning. Lockdowns reduced education to on-screen experiences that prevented social engagement, as teaching tended to resort to highly instructional front-of-screen-based delivery. Learners felt isolated and became stressed by long and intense online lessons, often in challenging and resource-constrained environments. They not only missed their friends but also lost the ability to socialize in and outside of learning experiences. This caused a breakdown in social skills and engagement, which, in turn, compounded further learning loss.

This key lesson must be taken forward when deliberating and designing futures of education, recognizing individual learner needs and learner variability, while stimulating social learning experiences.

EdTech developers and AI companies saw and exploited Covid-19 as an opportunity for personalization. These technological opportunities and capabilities presented themselves as tools that should be

leveraged, offering new possibilities for EdTech. Following the curriculum, these technologies too often amplified aspirations solely focused on knowledge acquisition for passing or excelling in a test.

In their infancy, these technologies were not sufficiently sophisticated to offer more-than-prescribed, data-driven learning paths derived from interaction with the platform's resources. The systems did not understand the complexity of learning and learners. While the arrival of generative AI is often highlighted as a significant breakthrough for better and more dynamic personalization of learning, we are at risk of ignoring the real opportunities these technologies present.

New potential: (Re-)socializing learning

Covid-19 demonstrated the consequences of the lack of socialization in education. Ignoring this lesson in pursuit of AI-driven hyper-personalization risks constraining learners to amplified de-socialized learning environments.

In co-creative learning scenarios, supported by teacher-learner dialogue, education could become re-socialized, while remaining personalized. This would enhance learners' social situations, empathy, critical thinking, self-regulation and metacognitive skills development. These should be framed in pedagogical approaches that bring dialogue to the centre of the classroom, allowing for the emergence of re-socialized intelligence.

Rather than harnessing hyper-personalization, AI technologies are also suited to engaging learners and teachers in social dialogue, collaborative problem-solving and peer-to-peer learning in mixed-ability learning encounters.

Hyper-personalization may be compelling to AI developers. It fits their mould. Yet, it risks overlooking the importance of our human social make-up and what makes us human in the first place. Anticipatory thinking about technology is essential when considering the future of learning in an exponentially evolving AI world. This cannot be achieved without recognizing that humanity thrives on socialization, which is inherent to being human and critical to our development, success and achievements.

Towards the future: (Re-)socialized collective intelligence

A new approach that fosters collective intelligence in human-to-human-to-AI environments should be considered. Carnegie Learning's MATHia platform[3] exemplifies this approach, fostering collective intelligence between learners, teachers and AI. The AI platform learns from learner inputs across the entire learner user base, as well as interactions. It uses this information to provide analytics to teachers to scaffold appropriate pedagogical interventions that help refine the system's responses. This enhances AI's ability to recognize patterns, providing teachers with the ability to make more informed intervention decisions. While this platform may not support effective re-socialization of learning, it does provide an example of what human-AI collective intelligence in the classroom could look like.

Another example of introducing a collective intelligence supported by generative AI is Khan Academy's Khanmigo,[4] which relies on human-AI interaction or human dialogue with the machine. However, while Khanmigo demonstrates collective intelligence, its mainly hyper-personalized approach still neglects the social dimension of learning.

3. See https://www.carnegielearning.com/solutions/math/mathia
4. See https://www.khanmigo.ai

We can rethink how LLMs can re-socialize learning by introducing non-human intelligence into educational interactions. Leveraging this opportunity would result in a new form of human-machine collective intelligence. Such an intelligence would weave human and non-human intelligence together, rather than merely replace human intelligence, leading to triadic learning experiences between students, teachers and AI.

Relevant personalization would not be ignored, but it would no longer be solely viewed through an individualized lens. To the contrary, personalization would be aided and refined to address learner needs, taking the fundamental social dimension of learning into account.

Enriched by AI, a new form of (re-)socialized collective intelligence would augment individual learner and teacher agency within a collective and technology-extended dialogue, extending beyond traditional collaborative learning to offer a social dialogue-enhanced dimension to learning.

Such learning environments could easily reach beyond the confines of the classroom, school and national borders, as networks, infrastructure and technologies scale and evolve. They would allow for agency to flourish in intercultural yet culture-relevant understandings and contexts, preparing learners for their uncertain future.

Developing safe and experimentation AI 'playgrounds' can put this into practice. In these environments, learners and teachers can co-learn and experiment with AI in collaborative problem-solving challenges to nurture new forms of collective intelligence.

Conclusion: Re-socialized personalized human-AI collective intelligence

In an AI-dominant world where machines can access, synthesize and generate knowledge at scales beyond human capacity, futures of education require a new paradigm to unlock learner potential. This paradigm rejects hyper-personalization in favour of fostering the social essence of learning.

Without (re-)socialization that recognizes and serves learner uniqueness, hyper-personalization will not only result in erosion of social and collaborative skills, but also risk atrophy of cognitive, metacognitive and self-regulation skills, which are fundamental to learner agency.

While we prepare learners for a thriving future in society and the workplace, where agency, communication, empathy and collaboration are essential, we must ask a fundamental question: is a hyper-personalization that stifles (re-)socialization the future we want for learners?

As educators, policy-makers and technology developers, we must collaborate to create learning environments that harness both human and AI potential, while preserving the social fabric of education. The way forward is not a choice between personalization and socialization, but a thoughtful integration of both into a new educational paradigm.

References

Barrera Castro, G. P., Chiappe, A., Ramírez-Montoya, M. S. and Alcántar Nieblas, C. 2025. Key Barriers to Personalized Learning in Times of Artificial Intelligence: A Literature Review. *Applied Sciences*, Vol. 15, No. 6. Basel, MDPI, p. 3103. https://doi.org/10.3390/app15063103 (Accessed 3 August 2025.)

Bayly-Castaneda, K., Ramírez-Montoya, M. S. and Morita-Alexander, A. 2024. Crafting Personalized Learning Paths with AI for Lifelong Learning: A systematic literature review. *Frontiers in Education*, Vol. 9. Lausanne, Frontiers Media. https://doi.org/10.3389/feduc.2024.1424386 (Accessed 3 August 2025.)

Casebourne, I., Shi, S., Hogan, M., Holmes, W., Hoel, T., Wegerif, R. and Yuan, L. 2024. Using AI to Support Education for Collective Intelligence. *International Journal of Artificial Intelligence in Education.* Leeds, International Artificial Intelligence in Education Society (IAIED).

Cosmos Institute and Morris, C. 2025. *Social tinkering: Why collaborative curiosity beats vibe-coding*? Austin, Cosmos Institute. https://blog.cosmos-institute.org/p/social-tinkering-why-collaborative (Accessed 3 August 2025.)

Criddle, C. 2025. AI chatbots de battle over human memories. *Financial Times*.

Dewey, J. 1916. *Democracy and Education*. New York, Macmillan.

Dumont, H. and Ready, D. D. 2023 On the promise of personalized learning for educational equity. *npj Science of Learning*, Vol. 8, No. 26. Berlin, Springer Nature. https://doi.org/10.1038/s41539-023-00174-x (Accessed 3 August 2025.)

Freire, P. 2007. *Pedagogia do Oprimido* [Pedagogy of the Oppressed]. New York, Continuum. (In Portuguese.)

Geschke, D., Lorenz, J. and Holtz, P., 2019. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, Vol. 58, No. 1. Leicester, The British Psychological Society, pp.129-149. https://doi.org/10.1111/bjso.12286 (Accessed 3 August 2025.)

Halpin, H. 2025. Artificial intelligence versus collective intelligence. *AI & Society*. Berlin, Springer Nature. https://doi.org/10.1007/s00146-025-02240-x (Accessed 3 August 2025.)

Järvelä, S., Molenaar, I. and de Mooij, S. 2022. *Giving children agency in an era of artificial intelligence: AI can embolden children to control and monitor their learning*. Zurich, BOLD. https://boldscience.org/giving-children-agency-in-an-era-of-artificial-intelligence (Accessed: 22 June 2025).

Laak, K.-J. and Aru, J. 2024. *AI and personalized learning: bridging the gap with modern educational goals*. Ithaca, arXiv. https://doi.org/10.48550/arXiv.2404.02798 (Accessed 3 August 2025.)

Mohammed, S. 2017. Understanding what doesn't work in personalized learning. *Brown Center Chalkboard*. Washington DC, The Brookings Institution. https://www.brookings.edu/articles/understanding-what-doesnt-work-in-personalized-learning (Accessed 3 August 2025.)

Piaget, J. 1932. *The moral judgement of the child*. New York, Simon & Schuster.

Singer, N. 2025. Google plans to roll out its A.I. chatbot to children under 13. *The New York Times*. https://www.nytimes.com/2025/05/02/technology/google-gemini-ai-chatbot-kids.html (Accessed 3 August 2025.)

Sitta, F.A., Maddox, B., Casebourne, I., Hughes, S., Kuvalja, M., Hannam, J. and Oates, T. 2023. *The Futures of Assessments: Navigating Uncertainties through the Lenses of Anticipatory Thinking*. Cambridge, Cambridge University Press & Assessment. https://www.cambridgeassessment.org.uk/Images/698413-the-futures-of-assessment-navigating-uncertainties-through-the-lenses-of-anticipatory-thinking.pdf (Accessed 3 August 2025.)

Skinner, B. F. 1960. Teaching Machines. *The Review of Economics and Statistics*, Vol. 42, No. 3. Cambridge, MIT Press, pp. 189-191.

Vygotsky, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Harvard University Press.

Walkington, C. and Bernacki, M.L., 2020. Appraising research on personalized learning: Definitions, theoretical alignment, advancements, and future directions. *Journal of Research on Technology in Education*, Vol. 52, No. 3. Milton Park, Taylor & Francis, pp. 235-252. https://doi.org/10.1080/15391523.2020.1747757 (Accessed 3 August 2025.)

Wegerif, R. 2025. *Rethinking Educational Theory: Education as Expanding Dialogue*. Cheltenham, Edward Elgar Publishing.

Infantilizing, echo chamber, filter bubble or the dawn of a new enlightenment: Some (critical) thoughts about adaptive and personalized learning

Paul Prinsloo

Prologue

AI-enabled personalized learning has been accused of infantilizing education (Hillman and Couldry, 2025), as creating echo chambers (Bozkurt et al., 2024; Jadiga, 2025), and/or a 'filter bubble' (Misiejuk et al., 2025; Pariser, 2012). And yet, others see AI-enabled personalization as the optimum use of technology in what is called 'Education 4.0' (Méndez-Vargas et al., 2025) and the dawn of a new Enlightenment (Itec, 2024).

Critically engaging with the different claims and evidence surrounding AI-enabled personalized learning is even more urgent in the light of persistent concerns about, inter alia, the inherent bias, gender stereotypes and racism in algorithmic decision-making systems (Browne et al., 2023; Khan, 2024; Kundi et al., 2023). Developments in agentic AI (Acharya et al., 2025; Gabriel et al., 2024) and multi-agent systems (Córdova-Esparza, 2025; Pujari et al., 2024) would furthermore add complexity, risks but also benefits.

In reflecting on AI-enabled personalization, it is important to consider our beliefs about learning, knowledge (epistemology) and what it means to be human (ontology). There is very little (if any) in the digital information sphere we encounter on a daily basis that is not personalized, whether overtly or covertly, by black-boxed algorithms serving various interests, inter alia, but not limited to commercial, political, ideological and increasingly nefarious areas.

As we increasingly share our lives with (automated) algorithmic decision-making systems, we need to think critically about the implications – what it means for and how it changes being human, and the risks, potential and challenges in AI-enabled personalized learning.

Introduction

Although adaptive or personalized learning is not a new phenomenon (Nwana, 1990), recent advances in AI, and specifically GenAI, have increased its potential and scope, but also its risks (Randieri, 2024; UNESCO, 2021). The spike in interest in and operationalizations of personalized learning can also be linked to the pervasiveness of devices in schools (in many contexts), the increased digitization and datafication of education, and advances in AI and GenAI (OECD, 2021; Williamson et al., 2023).

AI-enabled personalized learning promises, inter alia, to increase the efficiency and quality of students' learning, resulting in better outcomes, providing students with just-in-time feedback and speeding up the learning process – 'learning at the speed of light' (Biasi, 2025; see also UNESCO IITE, 2020). Personalized learning creates the advantage of not being 'stuck' in a class with other students who may need more support, or who have more questions, or have different goals in relation to achieving envisaged outcomes in a specified period. Personalized learning also promises to free up teachers' time and resources to spend on other fulfilling tasks (Makesh, 2025).

Despite the hype and claims about AI-enabled personalization, it is important to note that it is a multifaceted phenomenon. AI-enabled personalization consists of various methods of personalization, as well as different outcomes and foci; for example,

in relation to the time of adaption, the content, the method of adaption, and the envisaged outcomes of the adaptation in a particular context (Van Schoors et al., 2021). We also need to remember and remind ourselves that personalization systems do not emerge in a vacuum but arise from specific epistemologies, and ontological and ideological orientations, which are often not considered (Roux and Nodenot, 2023).

From a student's perspective, students have a right to know who or what is doing the personalization, what is being personalized, how is it personalized and for what purpose (Bernacki et al., 2021)? There are also other questions to reflect on, such as: how is the personalized design communicated to students; for example, direct messages, choice, avatars? What kinds of data are used in the personalization of learning; for example, voluntary/involuntary, explicit/implicit, overt/covert? And what are the sources of data and technologies/methods used for the personalization (Cavdar Aksoy et al., 2021)? From a teacher's perspective, questions regarding outsourcing responsibility, changing roles and shifts in teachers' authority and accountability arise (Li and Wong, 2023).

In this think-piece, I will not attempt to provide an overview of the field but rather critically engage with some aspects of AI-enabled personalized learning and current research findings that may be useful for policy, research and practice development.

What does AI-enabled personalization entail and what is its impact?

AI-enabled personalization is all around us on Google Maps, Netflix and Amazon and our social media feeds, which are suddenly filled with cats, body builders or fashion. In this piece, I consider educational AI-enabled personalization aimed at improving, inter alia, educational outcomes, student satisfaction, student engagement, responsiveness and feedback (Bernacki et al., 2021; Da Silva et al., 2023; Lu et al., 2015).

In the next section, I briefly consider a range of aspects of AI-enabled personalization and evidence from research, as a basis for formulating a number of recommendations for policy, practice and research.

Implicit/explicit personalization and user control

As a start, we can distinguish between explicit (overt) personalization and implicit (covert) personalization. In explicit personalization, users know or have been informed of the personalization and users may be allowed to make explicit choices regarding the personalization. Implicit personalization occurs when the personalization is done by the 'system' (Cavdar Aksoy et al., 2021). There is some evidence of the positive impact of user control on user perceived benefits of these recommender systems, and user control is being strongly correlated with trust and satisfaction (Ain et al., 2025). In the framework provided by Van Schoors et al. (2021) in the context of education, personalization can be student-controlled, or controlled by the 'system', or it can be the result of a negotiated hybrid agreement on the scope and purpose of personalization. As such, user control is linked to the scope and purpose of the automation of (elements in) the learning process.

Levels of automation

Early research in the field of human computer interaction considered different levels of automation at the intersection between humans and computers/robots (Beer et al., 2014; Sheridan and Verplank, 1978). Recently, Molenaar (2021) proposed six levels of automation in personalized learning, ranging from where (1) the teacher controls the learning without any

automation; (2) the teacher maintains full control but where technology provides supportive information; (3) the teacher monitors the technology and technology controls specific tasks; (4) the 'teacher monitors incidentally, but can resume control' at any time, while the technology signals 'when teacher control is needed' in a context where technology controls the broader set of tasks; (5) technology controls most tasks automatically, and may request the teacher to take control, but teacher control and monitoring is optional and infrequent; and (6) full automation where technology controls all tasks automatically (Molenaar, 2021, p. 60).

Data used for personalization

AI-enabled personalization requires a lot of data: (1) individual-level personal information (including past digital behaviour, as well as attitudes and preferences); (2) social-level personal information (for example, family, friends, classmates, community); and (3) situation-based personal information (for example, time- and location-based) (Cavdar Aksoy et al., 2021). In the context of adaptive learning in intelligent tutoring systems, Bernacki et al. (2021) provide a range of information sources for adaptive learning, such as students' prior knowledge and preparedness often used in intelligent tutoring systems; students' preferences with regard to pace, outcomes, formats; students' characteristics (for example, language proficiency, literacy, locus of control and self-efficacy); and self or institutionally identified needs. The quality and scope/variety of data is key to personalization and depends on: 'i) our ability to follow learners and track their environment; ii) our ability to diagnose learners' current states and anticipate their development; and iii) our ability to consequently determine the most appropriate action to optimize learning' (Molenaar, 2021, p. 62).

Architectures for personalization

Collected data can be used based on different architectures which can be classified as follows:

- Personalization or content-based approaches can be defined as the construction of choice architectures, which are not the same across a sample and based on an individual's past behaviours and choices, and may also include user preferences where users intentionally set the parameters of personalization. They are often transparent in the recommendation: 'Because you liked x, you may like y.'
- Collaborative filtering provides recommendations based on the preferences and actions of similar users; for example, where the personalization will be based on characteristics shared by a group of individuals, such as combinations of demographics, geography, learning behaviour and grades.
- Knowledge-based filtering, where instead of using historical data, this model combines features submitted by the user; for example, parameters and requirements (Aggarwal, 2016, pp. 14–16).
- Hybrid recommender systems, where different algorithms are combined.

Timing

We often do not think of the 'timing' of/in personalization, an issue raised in the framework by Van Schoors et al. (2021). Personalization can be static and designed before the learning journey starts, or the personalization can be dynamic and involve a continuous process of personalization based on set parameters. With regard to the latter, Molenaar (2021) refers to optimization during learning, whether with regard

to adjustments at task level in terms of students' mastery of a task, or in the case of 'step-level' disciplinary content (for example, mathematics), where competence is built up in a number of steps. 'These systems investigate a student's response to the task to provide automated, elaborate feedback at each step the student takes' (OECD, 2021, p. 65). The third type of timing involves personalizing students' journeys through a particular domain consisting of multiple topics.

The 'target' of personalization

The 'target' of personalization refers to 'what' is being personalized; for example, content, activities or additional support. The target of personalization is linked to the purpose of the personalization, based on the needs of a particular student or group of students. Whatever is personalized is also related to the 'timing' of the personalization. Van Schoors et al. (2021) point to the different possibilities for personalization, such as personalizing content and resources (for example, multimodal learning materials, whether in terms of difficulty/complexity or specific needs), and/or assessment (in terms of mastery and envisaged outcomes) and/or feedback on assessment and support in general (Luo et al., 2025). The review by Tan et al. (2025) on AI-Enabled Adaptive Learning Platforms (ALPs) refer to three inherent aspects of ALPs; namely, adaptive content, assessment and sequencing. The pedagogical underpinnings of ALPs include an emphasis on 'mastery learning', accounting for differences in student competencies or maturity and various cognitive thresholds – all of which require dynamic scaffolding. Four different methods were used in measuring the effectiveness of ALPs; namely, learning gains, student satisfaction, student engagement and student motivation.

Technologies used

The systematic review of 63 articles by Luo et al. (2025) found that 27 per cent of the studies reported using publicly available AI-based tools, while 24 per cent of the studies reported using tailored AI systems for specific instructional activities. An amount of 37 per cent of the studies reported using self-developed AI-based learning tools and 13 per cent reported using self-developed AI-based learning tools with low-code and no-code platforms. A total of 49 per cent of the studies reported on developing proprietary AI-based learning tools. The most frequently used algorithm was neurolinguistic programing (24 per cent), and other machine learning models comprised neural networks, random forests, Bayesian methods and K-nearest neighbours. In the corpus of analysis, eight studies (13 per cent) reported using low-code and no-code platforms, such as Google Dialogflow, ManyChat, Flow XO, Landbot and AutoTutor.

Interestingly, in their review, Tan et al. (2025) found that ALPs are typically used in STEM-related subjects, with over half of the ALPs designed for tertiary education contexts. Regarding the AI techniques used in these ALPs, the review found that in studies between 2014 and 2019 41.9 per cent of the corpus of analysis did not specify the AI techniques used, while 27.9 per cent used rule-based methods, and probabilistic/statistical methods and traditional machine learning both accounting for 9.3 per cent. In studies between 2020 and 2024, 43.8 per cent of the studies did not disclose any specifics on AI methods, with 15.6 per cent of the studies reporting on the use of rule-based methods, and 12.5 per cent of the studies reporting on GenAI (see also Hardaker and Glenn, 2025).

'Actors' in the personalization of learning

The meta systematic review by Bond et al. (2024) on AI in higher education was published after the launch of ChatGPT, too early to include peer-reviewed research inclusive of GenAI in the corpus of analysis. The authors use Zawacki Richter et al.'s (2019) typology of four distinct but related and at times overlapping foci; namely, profiling and prediction, intelligent tutoring systems, assessment and evaluation, and adaptive systems and personalization. With regard to adaptive systems and personalization, the research found evidence of chatbots/virtual assistants (n=20), providing personalized content (n=14), facial recognition and/or mood detection (n=9), recommender systems/course scheduling (n=5) and robots (n=3). The review by Luo et al. (2025), covering the period up to 2024, found that 41 per cent of the 63 papers reviewed, reportedly used chatbots to 'curate and deliver learning content'. The systematic literature review by Yılmaz (2024) found evidence of: (1) personalized learning platforms where AI is used to adapt teaching and content, based on an analysis of individual student needs and preferences; (2) intelligent tutoring systems, based on detailed student profiles, tracking of students' progress and adapting lessons; (3) virtual reality simulations where AI is used not only in the development of the simulation but also adjusting the difficulty and levels of complexity, based on an analysis of students' progress, characteristics and interactions; (4) data analysis and prediction models; and (5) adaptive assessment tools that provide personalized feedback and recommendations to students.

Aim of personalization

While it can be assumed that improving educational outcomes to be the most important aim of personalization, the review by Tan et al. (2025) refers to three inherent aspects of ALPs; namely, adaptive content, assessment and sequencing. The pedagogical underpinnings of ALPs include an emphasis on 'mastery learning', accounting for differences in student competencies/maturity and different cognitive thresholds – all of which require dynamic scaffolding. Four different methods were used in measuring the effectiveness of ALPs; namely, learning gains, student satisfaction, student engagement and student motivation.

Impact

Before the launch of ChatGPT in November 2023, Van Schoors et al. (2021) reviewed the impact of AI-enabled personalization in 48 studies on different outcomes – cognitive, affective and/or meta-cognitive. Although the results in general indicated a positive impact on some students, the different disciplinary contexts, technologies used and small sample sizes make it impossible to generalize. The authors conclude that the findings are, at best, 'fragmentary'. There is also evidence of the positive influence of chatbots/virtual assistants in the context of developing speaking skills and in enhancing accessibility for visually and hearing-impaired students (Bond et al., 2024). Pertaining to personalized content, one study reported a positive effect on learning skills, rather than growing factual knowledge. The personalization of learning materials was used most frequently, followed by personalized learning pathways, learning strategies and learning environments. Various studies reported on benefits, such as, but not limited to, 'greater insight into student understanding', 'positive influence on learning outcomes', 'reduced planning and administration time for teachers', 'greater equity in education' and 'precise assessment and feedback' (Bond et al., 2024, p. 27).

Research published after the launch of ChatGPT reports: 'While many of the reviewed AI tools demonstrated their efficacy in enhancing cognitive knowledge and affective learning, their impact on cognitive process development and skill acquisition were more mixed and indefinite' (Luo et al., 2025). The impact of AI-enabled personalization was found to be less effective in students who are less prepared (Tan et al. 2025). In the research by Tan et al. (2025), it is not clear to what extent the respective findings of impact in the included studies can be generalizable.

Other factors to consider pertaining to (measuring) the impact of AI-enabled personalized learning is the effect of device connectivity and student preferences, the selection of inappropriate or simplistic student characteristics (Tan et al., 2025), the 'lack of ethical consideration' (in 29 per cent of the corpus of analysis), and 'shifting authority' from teachers and the institution to AI (in 22.6 per cent of the corpus of analysis) (Bond et al., 2024; see also Magrani and Da Silva, 2023).

Some recommendations for policy, research and practice

The following list of recommendations is anything but conclusive and flows from the argument presented above.

1. We cannot and should not ignore the reality of misinformation, gender bias, stereotypes and racism, to mention but a few, in algorithmic decision-making systems and ecologies, and how these will shape personalization in and of learning.
2. AI-enabled personalization has become endemic in different aspects of our personal lives, often implicitly and most often without our explicit consent or input. Considering AI-enabled personalization in the context of education, it is important to recognize that we have choices, and we (our students, staff and institutions) should have choices. Access to education is not only a human right, but it also raises a range of moral, legal and ethical obligations on the side of the provider or institution in the context of the social contract between society and educational institutions/providers. We can design AI-enabled personalization in line with our moral, legal and ethical obligations.
3. The selection of evidence presented in this think piece shows a relatively unclear picture regarding who benefits, under what conditions and contexts, through what type of design and levels of automation and what levels of user control and architectures of personalization. It is crucial that more research is done on the research evidence on which AI-enabled personalization is designed, but also on issues surrounding user choice, ethical oversight, levels of automation and its impact on students, teachers and the education sector as a whole.
4. There seems to be a lack of consideration of the ethical and legal issues pertaining to AI-enabled personalization of learning, especially for under-age children and/or vulnerable populations.
5. AI-enabled personalization requires huge amounts, varieties and sources of data. A number of issues arise, such as user consent for the use of personal and other data, and additionally, possibly, in combination with data from other users.
6. In the context of personalization of services and social media, there is very little, if any, explicit opt-in to personalization, or access to the reasoning or parameters of algorithms. In the context of the moral, legal and ethical obligations inherent in the social contract underpinning education, students should have a right to opt out of AI-enabled personalization and be offered alternatives of equal quality and recognition.
7. Students have a right to determine which parts of their learning journey should be personalized, in addition to choosing the 'timing', process and focus

of personalization, as well as the 'actors' in the personalization of their learning. Student autonomy and agency have to be respected, supported and enhanced.

8. Policy-makers, regulatory bodies and school governing bodies/municipalities should develop guidelines of acceptable, ethical and responsible AI-enabled personalization of learning, how it will be quality/ethically assured, as well as governed.
9. Research findings (so far) indicate that AI-enabled personalization is more effective in some disciplinary contexts, and for particular student groups, than being a one-size-fits-all.
10. What is not clear in the current evidence on AI-enabled personalization of learning is on what empirical evidence the personalization is based. The design of AI-enabled personalization should be very transparent about the basis of the empirical research elements of the personalization.
11. Current evidence seems to suggest that AI-enabled personalization can have positive impacts on student satisfaction, achieving the learning outcomes, the quality of the learning experience and preventing student attrition, among other benefits. It is crucial to determine not only the enabling factors for positive impacts, but also under what circumstances and in which contexts AI-enabled personalization is not the answer (UNESCO, 2021).

(In)conclusions

Learning has always been personal. In many ways, it has been personalized in the different ways teachers responded to the individual needs and progress of students (Arantes, 2024). Compared to AI-enabled personalization, personalization done by human teachers looks different in many ways. As explored and mapped in this think piece, AI-enabled personalization of learning is multidimensional and consist of many, often interdependent and mutually constitutive factors and elements.

It is therefore dangerous and irresponsible to dismiss the possible benefits and potential of AI-enabled personalization in all of its nuances and possibilities. Equally irresponsible and dangerous is to embrace AI-enabled personalization without considering the different elements, factors and interdependencies – as phenomenon, but also in the context of a specific educational context, age group of students and disciplinary focus, to mention but a few.

Different elements or factors of AI-enabled personalization raise different ethical concerns, and the design and operationalization of AI-enabled personalization should consider the potential of harm for all users, but specifically for under-age children and vulnerable individuals. What is crucial in the design and deployment of AI-enabled personalization in education, whether in preschool, school or post-school environments, are the specifics of the individual – their autonomy, agency and well-being. The dangers of AI-enabled personalization systems becoming self-fulfilling prophesies, re-enforcing social (dis) advantages and eroding personal autonomy, are real. And for true personal autonomy to flourish, informational diversity is an enabling condition (Misiejuk et al., 2025).

Epilogue

The Prologue to this think piece referred to different contrasting claims and positions pertaining to AI-enabled personalization – as infantilizing, echo chambers, filter bubbles or heralding the dawn of a new Enlightenment. While AI-enabled personalization has become an integral part and parcel of our lives – whether we opted in or not – in the context of the social contract between society and education, with its different moral and ethical

obligations, we can design personalization differently, considering some of the aspects covered in this critical reflection.

The issue is not whether AI-enabled personalization 'works' or not, but how it embodies a particular view and understanding of knowledge, of learning and of being human.

Should we consider human learning as inherently social, relational, intersectional and material, then some proposals and forms of AI-enabled personalization may fall short in many respects and be a parody of human learning. The individual who is learning is not a disembodied individual-without-others, but entangled in assemblages of humans, non-human actors, contexts and time. Learning emerges from a collective – a range of formal but also serendipitous inter- and intra-actions with human and non-human actors and networks.

Learning augmented and enabled by algorithmic decision-making systems is therefore much more than just a technological question, but also philosophical, theoretical, political, economic, legal, environmental and/or social.

What do we gain and what do we lose with AI-enabled personalization of learning? And who is this 'we'?

Acknowledgement

No part of this think piece was generated by GenAI, and I accept responsibility for some, possibly, awkward and idiosyncratic human phrasing. Secondly, I would like to express my appreciation to the reviewers for their critical comments and guidance.

References

Acharya, D. B., Kuppan, K. and Divya, B. 2025. *Agentic AI: Autonomous intelligence for complex goals – a comprehensive survey*. New York, IEEE. http://doi.org/10.1109/ACCESS.2025.3532853 (Accessed 5 August 2025.)

Aggarwal, C. C. 2016. *Recommender Systems: The Textbook*. Cham, Springer International Publishing.

Ain, Q. U., Chatti, M. A., Tsoplefack, W. K., Alatrash, R. and Joarder, S. 2025. *Designing and Evaluating an Educational Recommender System with Different Levels of User Control*. Ithaca, arXiv. https://arxiv.org/abs/2501.12894 (Accessed 5 August 2025.)

Arantes, J. 2024. Digital twins and the terminology of 'personalization' or 'personalized learning' in educational policy: a discussion paper. *Policy Futures in Education*, Vol. 22, No. 4. Thousand Oaks, Sage Publications, pp. 524–543. https://doi.org/10.1177/14782103231176357 (Accessed 5 August 2025.)

Beer, J. M., Fisk, A. D. and Rogers, W. A. 2014. Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of Human-Robot Interaction*, Vol. 3, No. 2. New York, Association for Computing Machinery (ACM), pp. 74–99. https://doi.org/10.5898/JHRI.3.2.Beer (Accessed 5 August 2025.)

Bernacki, M. L., Greene, M. J. and Lobczowski, N. G. 2021. A systematic review of research on personalized learning: personalized by whom, to what, how, and for what purpose(s)? *Educational Psychology Review*, Vol. 33, No. 4. Berlin, Springer, pp. 1675–1715.

Biasi, A. 2025. Learning at the speed of light: a conversation with Michael Moe. *Educate AI*, 31 March. https://edu-ai.org/learning-at-the-speed-of-light-a-conversation-with-michael-moe (Accessed 5 August 2025.)

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W. and Siemens, G. 2024. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, Vol. 21, No. 4. Berlin, Springer Nature. https://doi.org/10.1186/s41239-023-00436-z (Accessed 5 August 2025.)

Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y., Nerantzi, C., Moore, S., Dron, J., Stracke, C. M., Singh, L., Crompton, H., Koutropoulos, A., Terentev, E., Pazurek, A., Nichols, M., Sidorkin, A. M., Costello, E., Watson, S., Mulligan, D., Honeychurch, S., Hodges, C. B., Sharples, M., Swindell, A., Frumin, I., Tlili, A., Slagter van Tryon, P. J., Bond, M., Bali, M., Leng, J., Zhang, K., Cukurova, M., Chiu, T. K. F., Lee, K., Hrastinski, S., Garcia, M. B., Sharma, R. C., Alexander, B., Zawacki-Richter, O., Huijser, H., Jandrić, P., Zheng, C., Shea, P., Duart, J. M., Themeli, C., Vorochkov, A., Sani-Bozkurt, S., Moore, R. L. and Asino, T. I. 2024. The manifesto for teaching and learning in a time of generative AI: a critical collective stance to better navigate the future. *Open Praxis*, Vol. 16, No. 4. Oslo, International Council for Open and Distance Education (ICDE), pp. 487–513. https://doi.org/10.55982/openpraxis.16.4.777 (Accessed 4 August 2025.)

Browne, J., Cave, S., Drage, E. and McInerney, K. (eds). 2023. *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*. Oxford, Oxford University Press.

Cavdar Aksoy, N., Tumer Kabadayi, E., Yilmaz, C. and Kocak Alan, A. 2021. A typology of personalisation practices in marketing in the digital age. *Journal of Marketing Management*, Vol. 37, No. 11–12. London, Taylor & Francis, pp. 1091–1122.

Córdova-Esparza, D. M. 2025. AI-powered educational agents: opportunities, innovations, and ethical challenges. *Information*, Vol. 16, No. 6. Basel, MDPI, p. 469. https://doi.org/10.3390/info16060469 (Accessed 5 August 2025.)

Da Silva, F. L., Slodkowski, B. K., Da Silva, K. K. A. and Cazella, S. C. 2023. A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities. *Education and Information Technologies*, Vol. 28, No. 3. Berlin, Springer Nature, pp. 3289–3328.

Gabriel, I., Manzini, A., Keeling, G., Hendricks, L. A., Rieser, V., Iqbal, H., ... and Manyika, J. 2024. *The ethics of advanced AI assistants*. Ithaca, arXiv. http://arxiv.org/abs/2404.16244 (Accessed 5 August 2025.)

Hardaker, G. and Glenn, L. E. 2025. Artificial intelligence for personalized learning: a systematic literature review. *International Journal of Information and Learning Technology*, Vol. 42, No. 1. Leeds, Emerald Publishing, pp. 1–14.

Hillman, V. and Couldry, N. 2025. Infantilising education through risk-averse educational technologies of calculability: a critical essay. *British Journal of Sociology of Education*, Vol. 46. London, Taylor & Francis.

Itec. 2024. *Learning, teaching & training in the era of Artificial Intelligence: Challenges and opportunities for evidence-based educational research*. Leuven, Acco. https://itec.kuleuven-kulak.be/wp-content/uploads/2024/03/Positioning-paper-itec.pdf (Accessed 4 August 2025.)

Jadiga, S. 2025. Understanding the role of AI in personalized recommendation systems, applications, concepts, and algorithms. *International Journal of Computer Trends and Technology*, Vol. 73, No. 1. Trichy, Seventh Sense Research Group, pp. 106–118. https://doi.org/10.14445/22312803/IJCTT-V73I1P113 (Accessed 4 August 2025.)

Khan, V. 2024. Artificial intelligence and gender bias: Analyzing algorithmic discrimination in language models. *Journal of Gender, Power, and Social Transformation*, Vol. 1, No. 2. Milton, Research Corridor, pp. 31–40. https://researchcorridor.org/index.php/jgpst/article/view/329/313 (Accessed 5 August 2025.)

Kundi, B., El Morr, C., Gorman, R. and Dua, E. 2023. Artificial intelligence and bias: a scoping review. *AI and Society*. New York, Chapman and Hall/CRC, pp. 199–215.

Li, K. C. and Wong, B. T. M. 2023. Artificial intelligence in personalised learning: a bibliometric analysis. *Interactive Technology and Smart Education*, Vol. 20, No. 3. Leeds, Emerald Publishing, pp. 422–445.

Lu, J., Wu, D., Mao, M., Wang, W. and Zhang, G. 2015. Recommender system application developments: a survey. *Decision Support Systems*, Vol. 74. Amsterdam, Elsevier, pp. 12–32. https://doi.org/10.1016/j.dss.2015.03.008 (Accessed 5 August 2025.)

Luo, J., Zheng, C., Yin, J. and Teo, H.H. 2025. Design and assessment of AI-based learning tools in higher education: a systematic review. *International Journal of Educational Technology in Higher Education*, Vol. 22. Berlin, Springer Nature. https://doi.org/10.1186/s41239-025-00540-2 (Accessed 5 August 2025.)

Magrani, E. and Da Silva, P. G. F. 2023. The ethical and legal challenges of recommender systems driven by artificial intelligence. H. Sousa Antunes, P. M. Freitas, A. L. Oliveira, C. Martins Pereira, E. Vaz de Sequeira and L. Barreto Xavier (eds.), *Multidisciplinary Perspectives on Artificial Intelligence and the Law*. Cham, Springer International Publishing, pp. 141–168. https://doi.org/10.1007/978-3-031-41264-6_8 (Accessed 5 August 2025.)

Makesh, L. 2025. How AI is transforming personalized learning in 2025 and beyond. *eLearning Industry*, 1 May. https://elearningindustry.com/how-ai-is-transforming-personalized-learning-in-2025-and-beyond (Accessed 5 August 2025.)

Masciari, E., Umair, A. and Ullah, M. H. 2024. A systematic literature review on AI-based recommendation systems and their ethical considerations. *IEEE Access*, Vol. 12. New York, IEEE, pp. 121223–121241. http://doi.org/10.1109/ACCESS.2024.3451054 (Accessed 5 August 2025.)

Méndez-Vargas, Á. M., Chiappe, A., Ramirez-Montoya, M. S. and Becerra Rodríguez, D. F. 2025. Personalised Learning and the Fourth Industrial Revolution: a conceptual review of definitions, trends and gaps. *Cambridge Journal of Education*, Vol. 55, No. 3. London, Taylor & Francis, pp. 357–375.

Misiejuk, K., Samuelsen, J., Kaliisa, R. and Prinsloo, P. 2025. Idiographic learning analytics: mapping of the ethical issues. *Learning and Individual Differences*, Vol. 117. Amsterdam, Elsevier. https://doi.org/10.1016/j.lindif.2024.102599 (Accessed 4 August 2025.)

Nwana, H. S. 1990. Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, Vol. 4, No. 4. Berlin, Springer Nature, pp. 251–277.

OECD. 2021. *OECD Digital Education Outlook 2021: Pushing the Frontiers with AI, Blockchain and Robots*. Paris, OECD Publishing. https://doi.org/10.1787/589b283f-en (Accessed 5 August 2025.)

Pariser, E. 2012. *The Filter Bubble: How the Personalized Web Is Changing What We Read and How We Think*. New York, Penguin.

Pujari, T., Goel, A. and Sharma, A. 2024. Ethical and responsible AI: governance frameworks and policy implications for multi-agent systems. *International Journal Science and Technology*, Vol. 3, No. 1. Bekasi, Assosiasi Dosen Muda Indonesia (ADMI), pp. 72–89. http://doi.org/10.56127/ijst.v3i1.1962 (Accessed 5 August 2025.)

Randieri, C. 2024. Personalized learning and AI: revolutionizing education. *Forbes*. https://www.forbes.com/councils/

forbestechcouncil/2024/07/22/personalized-learning-and-ai-revolutionizing-education (Accessed 5 August 2025.)

Roux, L. and Nodenot, T. 2023. Ethics of E-Learning Recommender Systems: epistemic positioning and ideological orientation. S. Genovesi, K. Kaesling and S. Robbins (eds.), *Recommender Systems: Legal and Ethical Issues.* Berlin, Springer, pp. 203–222.

Sheridan, T. B. and Verplank, W. L. 1978. *Human and Computer Control of Undersea Teleoperators*. Man-Machine Systems Laboratory Report. Cambridge, MIT. https://apps.dtic.mil/sti/pdfs/ADA057655.pdf (Accessed 5 August 2025.)

Tan, L. Y., Hu, S., Yeo, D. J. and Cheong, K. H. 2025. Artificial intelligence-enabled adaptive learning platforms: A review. *Computers and Education: Artificial Intelligence*, Vol. 9. Amsterdam, Elsevier. https://doi.org/10.1016/j.caeai.2025.100429 (Accessed 5 August 2025.)

Van Schoors, R., Elen, J., Raes, A. and Depaepe, F. 2021. An overview of 25 years of research on digital personalised learning in primary and secondary education: a systematic review of conceptual and methodological trends. *British Journal of Educational Technology*, Vol. 52, No. 5. London, British Educational Research Association, pp. 1798–1822.

Williamson, B., Macgilchrist, F. and Potter, J. 2023. Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, Vol. 48, No. 1. London, Taylor & Francis, pp. 1–5. https://doi.org/10.1080/17439884.2023.2167830 (Accessed 5 August 2025.)

UNESCO. 2021. *AI and education: Guidance for policy-makers*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709 (Accessed 5 August 2025.)

UNESCO IITE. 2020. *AI in Education: Change at the Speed of Learning*. Moscow, UNESCO IITE. https://unesdoc.unesco.org/ark:/48223/pf0000374947 (Accessed 5 August 2025.)

Yılmaz, Ö. 2024. Personalised learning and artificial intelligence in science education: current state and future perspectives. *Educational Technology Quarterly*, 2024, No. 3. Kryvyi Rih, Academy of Cognitive and Natural Sciences (ACNS), pp. 255–274. https://doi.org/10.55056/etq.744 (Accessed 5 August 2025.)

Zawacki-Richter, O., Marín, V. I., Bond, M. and Gouverneur, F. 2019. Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, Vol, 16. Berlin, Springer Nature. https://doi.org/10.1186/s41239-019-0171-0 (Accessed 5 August 2025.)

The end of assessment as we know it: GenAI, inequality and the future of knowing

Mike Perkins and Jasper Roe

Introduction

Since ChatGPT burst into public view in 2022, Generative AI (GenAI) has dominated educational debates. Initially, educational institutions in multiple global contexts sought to ban the use of this technology among learners (Yu, 2023); however, at the time of writing, this tone has significantly shifted. For example, the United Kingdom Department for Education now claims that GenAI can support learners of all ages in developing lifelong knowledge and skills (Department for Education, 2025). However, this shift in tone towards GenAI integration leaves many questions unanswered. One of the most pressing questions facing educators is how to accurately measure student learning when traditional assessment methods can now easily be replicated by GenAI-powered tools, challenging our assumptions around authenticity and originality.

GenAI is no longer a simple text creator that makes obvious and easy-to-spot mistakes. Newer models are sophisticated, multimodal content creators that can develop audio and visual data, including copying human likenesses (an emerging phenomenon known as deepfakes) (Roe et al., 2024) or completing actions and tasks autonomously (known as agentic AI). Even older models, such as GPT-4, can pass postgraduate medical examinations on par with human performance (Newton and Xiromeriti, 2024). Looking to the future, the inflow of enormous amounts of capital will drive the growth of GenAI capabilities. Most private investments in AI development are in the United States (€62.5 billion in 2023), followed by the People's Republic of China (€7.3 billion) and the United Kingdom of Great Britain and Northern Ireland (€3.5 billion). In the United States, half of this investment in AI is specifically in GenAI (European Parliament, 2024). Assuming that no ceiling in capabilities is reached, it appears almost inevitable that these systems will develop to the point where almost any take-home assessment will be trivial for GenAI systems to complete with minimal human input. Such outputs may be essentially undetectable, as the development of technologies to detect GenAI text outputs has largely failed to date (Perkins et al., 2024, Weber-Wulff et al., 2023). The question is: what does the future of education look like when we can no longer be sure of our learners' skills and capabilities? Is assessment, as we know it, in a state of collapse?

The false security of traditional assessment

A reactive response to this new threat has been to retreat to tightly invigilated exams or in-person viva voces, and some institutions are redrafting assessment strategies around this model, convinced that a physical room offers the last bastion of security and authenticity in assessment. However, even these refuges are fragile. Smart glasses with built-in AI assistants, such as Meta-Ray-Ban's recent releases, are now on sale to the public (Waisberg et al., 2024) and could feasibly be used in secured examinations. Furthermore, brain-computer interfaces, which are already in clinical trials, will make it impossible to guarantee that a learner is not aided during examinations (Eaton, 2023). Eaton's claim that we are entering a 'post-plagiarism' era is probable, and if we cannot secure the exam hall and

cannot verify authorship in take-home work, the logic of reverting to traditional exams fails to address this issue.

For many academics, this simply exposes an old truth: assessments can be gamed by students. Contract cheating and collusion have long slipped through imperfect or non-existent detection systems, but the threat posed by GenAI-powered tools is harder to ignore. Currently, any first-year student can generate a polished doctoral-level essay on demand with minimal barriers to access. Perhaps this critical moment can also shed light on other previously ignored issues: how long has assessment truly captured learning, and how long have we tolerated the background hum of misconduct or invalid design? This could be the impetus to begin addressing these questions, which have long been a thorn in the side of educators worldwide.

Based on the above, it seems that most of our current methods of validating learning will soon be GenAI-vulnerable, leading us to conclude that assessment, as we know it, is collapsing. However, this collapse is unlikely to occur uniformly worldwide. There is a significant distinction between different global contexts and, therefore, different implications that this potential collapse holds for their education systems. Digitally advantaged regions that enjoy high-bandwidth internet access, plentiful hardware, and the ability to develop staff and student GenAI literacy will confront a different future from digitally marginalized systems, which may need to remain reliant on traditional, pen-and-paper, large-scale, high-stakes assessment modes. The key question that arises here is not just technical but also epistemic, centred on whose knowledge will be recognized and whose credentials devalued.

Digitally advantaged contexts: GenAI-integrated experiments

When we discuss digitally advantaged contexts, we go beyond the traditional binary conceptions of the Global North and the Global South. Instead, we point to the fact that in contexts with access to advanced digital models, the resources needed to deploy them and the ability to train users to be proficient, assessment will be reformed in distinct ways compared with resource-limited contexts. This distinction may correlate with those countries that are investing heavily in AI development, such as the United States (European Parliament, 2024). In these contexts, assessment is already being reimagined, leading to a rebuilding of methods that either carefully carve out a role for GenAI in education or move to refocus assessment on future-ready skills and capabilities.

Ungrading collaborative problem-solving tasks, and human-centred approaches, which focus on embodied, ethical and relational knowledge areas less easily simulated by AI, are gaining prominence**.** Frameworks such as the AI Assessment Scale (AIAS)[1] help staff decide how much AI involvement is pedagogically useful, rather than simply performative. Ranging from Level 0 (no AI) to Level 5 (AI exploration), redesign tools like this direct attention towards transparency, creativity and ethical agency. The scale helps educators think about how they wish to engage in assessment redesign to consider GenAI use, and helps teachers and learners articulate why, when and how an AI tool adds value.

However, the implementation of collapse-avoidance techniques relies on resource availability. Time, staffing and funding are needed to help reimagine assessment, and these luxuries are more available in well-resourced institutions and geographical

1. See https://aiassessmentscale.com

locations. Nevertheless, these options show how assessment can be revitalized by integrating GenAI as a productive educational partner, rather than treating its use as inherently problematic. Digitally advanced institutions may be able to explore what assessment looks like when we move past traditional forms of knowledge assessment and embrace the core concepts of human connection, moral judgement and lived impact.

Digitally marginalized contexts: entrenched inequities

While pedagogical innovation and redesign may guide us towards embracing an AI-integrated future of assessment, we cannot ignore the critical challenge that remains unaddressed. Ninety-nine per cent of the world's languages lack the data required to train state-of-the-art GenAI models (Choudhry, 2023). In many rural and low-income settings, internet connectivity is irregular, devices are shared (if available at all), and professional development and technical training budgets are negligible. In these settings, most educational institutions will not be able to adopt GenAI; thus, students will not benefit from future-oriented assessment practices using the latest technologies.

Furthermore, institutions that adopt GenAI will likely adopt the most dominant and freely accessible models available. As leading GenAI models primarily rely on English-language datasets that reflect a Western world view and system of knowledge (Roe, 2025), there is a genuine risk of marginalizing indigenous knowledge systems and exacerbating existing educational inequalities. Therefore, it is essential to critically examine and acknowledge how reliance on GenAI could influence whose knowledge is validated or marginalized. In lower-resource contexts, institutions that are anxious about Gen AI-written assignments may double down on highly proctored, closed-book exams, focusing on rote learning and high-stakes testing. Paradoxically, GenAI risks freezing marginalized contexts in nineteenth-century assessment modes, even as it propels resource-rich institutions towards experimental pedagogies.

Assessment and power

We contend that a post-collapse vision of assessment must carefully consider power. Assessment has always reflected power relations between the norms and rules of the institution, which are mediated by the teacher and significantly affect students' lives (Reynolds and Trehan, 2000). From this perspective, the integration of GenAI into educational assessment can be viewed as reinforcing existing power imbalances, favouring those with access to digital rather than physical resources, and compelling resource-limited educational systems to adapt or risk marginalization.

To this end, the idea that GenAI may become a great equalizer is more myth than reality. Practically speaking, without the intentional redistribution of infrastructure, language equity and inclusive design in educational institutions, GenAI may deepen structural divides, especially between digitally advantaged and digitally marginalized educational systems. Ultimately, this may lead to an entrenchment of unnecessarily traditional forms of assessment in marginalized contexts and a reformed system in advantaged contexts, which will serve to deepen digital divides in opportunity in the future. However, we must also recognize the role of human agency, resistance and resilience in such cases – educators will not easily relinquish their independence, as evidenced by a growing 'refusal' movement (McInnes, 2025).

Conclusion: towards a plural, post-assessment future

The emergence of GenAI has led to the potential collapse of traditional assessments. However, it offers us the chance to carefully consider what the future of assessment could look like. This raises serious fundamental questions that we must critically engage with: what modes of knowledge should we value? What are the power effects of these technologies being used widely by learners and teachers? How can the benefits of GenAI use be distributed ethically and equitably, especially in lower-resource contexts?

If, in a post-collapse scenario, skills like recall, analysis and writing are delegated to machines, and assessment must pivot to human judgement, ethical reasoning and relational expertise, potentially using frameworks such as the AIAS to define what appropriate AI integration really means. Digitally advantaged institutions and systems can prototype such futures now, but they also carry a duty to share resources and lobby for multilingual, open-source AI models, while actors in lower-resourced contexts must make their voices heard. Without intentional action towards infrastructure redistribution and linguistic inclusion, we risk creating a divided educational landscape: privileged learners engaged in reflective, technology-enhanced enquiry, while millions remain confined to outdated examination systems. A truly inclusive future demands that we critically reconsider not only our assessment methods but also the types of knowledge that we prioritize and validate. Systems must be built that recognize the value of technology where there is a clear benefit for human flourishing, and resist where it is inappropriate.

The questions we face are urgent, challenging and multilayered in nature. These cover questions of equality, power and shifting social paradigms; yet the collapse of traditional assessment is an opportunity for us all to consider what the future of assessment could be. Institutions and policy-makers must take deliberate steps to redistribute digital resources and develop multilingual GenAI models to ensure that future assessment practices are inclusive and equitable.

References

Choudhury, M. 2023. Generative AI has a language problem. *Nature Human Behaviour*, Vol. 7, No. 11. Berlin, Springer Nature, pp. 1802–1803.

Department for Education. 2025. *Generative artificial intelligence (AI) in education*. London, Department for Education. https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education/generative-artificial-intelligence-ai-in-education (Accessed 5 August 2025.)

Eaton, S. E. 2023. Postplagiarism: transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, Vol. 19, No. 1. Berlin, Springer Nature, pp. 1–10. https://doi.org/10.1007/s40979-023-00144-1 (Accessed 5 August 2025.)

European Parliament. 2024. *AI Investment: EU and Global Indicators*. Brussels, European Parliament. https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/760392/EPRS_ATA(2024)760392_EN.pdf (Accessed 5 August 2025.)

McInnes, R. 2025. *Resist the Gen-AI-driven university: a call for reclaiming thought in learning and teaching*. Tugun, Australasian Society for Computers in Learning in Tertiary Education (ASCILITE). https://blog.ascilite.org/resist-the-gen-ai-driven-university-a-call-for-reclaiming-thought-in-learning-and-teaching/ (Accessed 5 August 2025.)

Newton, P. and Xiromeriti, M. 2024. ChatGPT performance on multiple-choice question examinations in higher education: a pragmatic scoping review. *Assessment & Evaluation in Higher Education*, Vol. 49, No. 6. Milton Park, Taylor & Francis, pp. 781–798.

Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J. and Khuat, H. Q. 2024. Simple techniques to bypass GenAI text detectors: implications for inclusive education. *International Journal of Educational Technology in Higher Education*, Vol. 21, No. 1. Berlin, Springer Nature, p. 53. https://doi.org/10.1186/s41239-024-00487-w (Accessed 5 August 2025.)

Reynolds, M. and Trehan, K. 2000. Assessment: a critical perspective. *Studies in Higher Education*, Vol. 25, No. 3. London, Taylor & Francis, pp. 267–278.

Roe, J. 2025. Generative AI as cultural artifact: applying anthropological methods to AI literacy. *Postdigital Science and Education*. https://doi.org/10.1007/s42438-025-00547-y (Accessed 5 August 2025.)

Roe, J., Perkins, M. and Furze, L. 2024. Deepfakes and higher education: a research agenda and scoping review of synthetic media. *Journal of University Teaching and Learning Practice*, Vol. 21, No. 10. The Hague, Open Access Publishing Association, pp. 1–22. https://doi.org/10.53761/2y2np178 (Accessed 5 August 2025.)

Waisberg, E., Ong, J., Masalkhi, M., Zaman, N., Sarker, P., Lee, A. G. and Tavakkoli, A. 2024. Meta smart glasses: large language models and the future for assistive glasses for individuals with vision impairments. *Eye*, Vol. 38, No. 6. Berlin, Springer Nature, pp. 1036–1038. https://doi.org/10.1038/s41433-023-02842-z (Accessed 5 August 2025.)

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O. and Waddington, L. 2023. Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, Vol. 19, No. 1. Berlin, Springer Nature, pp. 1–39. https://doi.org/10.1007/s40979-023-00146-z (Accessed 5 August 2025.)

Yu, H. 2023. Reflection on whether ChatGPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, Vol. 14. Lausanne, Frontiers Media. https://doi.org/10.3389/fpsyg.2023.1181712 (Accessed 3 August 2025.)

The ends of tests: Possibilities for transformative assessment and learning with generative AI

Bill Cope, Mary Kalantzis and Akash Kumar Saini

The dangers of generative AI are myriad. By now they have been thoroughly documented and widely acknowledged. Its extensive hazards and fast-moving evolution require caution (Peters et al., 2023). This short piece does not focus on the well-documented negatives; rather, it addresses the positive potential and new opportunities for education. These, we will argue, can only be adequately realized in carefully curated AI education spaces.

The particular focus here will be on the potential AI offers to transform assessment and, upstream from that, pedagogy. We make our case in broad brushstrokes, but also with brief reference to a research and development initiative of ours that applies generative AI to education, CyberScholar. This exemplifies an alternative to legacy assessment systems, while mitigating the worst of AI's deficiencies, providing evidence of what is practically possible now and signposting possible futures.

This offering is organized around what we consider to be the three great potentials for generative AI in education.

1. Careful combination of contextual AI and generative AI could bring about a quantum leap in pedagogy and assessment, moving learners from superficial content learning to deep learning.
2. AI can magnify the effectiveness of teachers, raise the level of their professionalism, and enhance the quality of their professional lives – but this will be quite a different job, for which they will need to become AI-proficient.
3. AI may open a path to epistemic justice for all, in a world where, until now, education has been complicit in the systemic reproduction of wider social inequalities.

An introductory note: our writing here is not in a genre of academic disinterest. Rather, it is more in the nature of a call to action or manifesto.

From surface learning to deep learning

Until now, tests have been the main way in which education systems formally determined what learners have learned. Summative assessment has been the principal point of accountability in formal education systems. For this reason, assessment has also tended to overshadow curriculum, so much so that much teaching and learning is frequently dominated by reverse engineering the epistemological and pedagogical expectations of assessments. However, as educators know all too well, assessments are peculiar artifacts and frequently not well aligned with learning. One of the most significant aspects of generative AI will likely be to render anachronistic legacy assessment instruments and processes. If and when this happens, the effect will be to open up new opportunities for pedagogy. As a counterpoint to the following arguments, our starting point will be to parse legacy assessments.

By far the most prevalent form of testing is the select response or multiple-choice test, from short teacher-created class quizzes to standardized and high-stakes assessments. Their profound inadequacies are well known (Biesta, 2016; Gergen and Dixon-Román, 2013; Shepard, 2000; Uher, 2020). Applying the measures of validity and reliability in classical test theory (Lord et al., 1968) and item response theory (Lord, 1980), we might summarize their deficiencies as follows:

- Content validity: Select response assessments often assume that understanding can be measured in definitive answers, reducing complex knowledge to discrete facts or procedural recall. The result is to narrow the epistemological and cognitive range of assessable learning.
- Criterion-related validity: The answer to question 2 may well be 'B', but there is no way of knowing whether deep, superficial or even accidental thinking led the examinee to this answer. Moreover, A, C and D are 'distractor' items, deliberate traps where, as often as not and by test design, correct thinking might produce a wrong answer. As such, they are also beset by cultural and linguistic biases.
- Construct validity: Item response theory proposes that test items are markers of 'latent' cognitive traits. This connection is at best hypothetical, at worst tendentious. Lee Cronbach and Paul Meehl (1955) optimistically called for 'nomological nets' to validate test items requiring external corroboration that rarely occurs. Direct relations are, in any event, unprovable because of the indefinable distance between atomic test items and holistic understanding of a domain. Moreover, theta, the measure of latency, is but one point on a unidimensional numerical scale of understanding. Knowledge, on the other hand, is intrinsically multidimensional, and student capacities are variable across many scales.
- Authenticity: As artifacts, select response assessments lack assessment validity because they are so different from real-world knowledge and learning processes. They reduce knowledge to a single, monolithic and ostensibly quantifiable gradient, when the real world of knowledge must engage with contestable alternative perspectives and paradigms.
- Consequential validity: Select response assessments measure long-term memory (defined as that which can be remembered until the day after the test) but not necessarily domain or discipline understanding. They introduce perverse incentives into education systems and are always problematic but increasingly anachronistic in complex and AI-infused knowledge systems.
- Generalizability: A test is no more than a small, strange and time-confined sampling mechanism, created in artificial conditions.

Select response tests have always been cheap and lazy ways to evaluate learners. At best, this most-used tool to assess learning and provide feedback to learners measures 'surface learning'. But in these days of technological, scientific and social complexity, they are less appropriate than ever. With the arrival of AI, their persistent use has become inexcusable.

Less frequent have been supply response assessments – essays, documented projects, worked problems, research reports, experiment write-ups and such like. With these kinds of assessments, it is possible to get a more accurate and nuanced view of student learning. But they still have serious limitations.

- Generalizability: Like select response assessment, these are still only a small and not necessarily representative sample of student learning. A learner may be lucky enough to get an essay question in an exam that takes them to an area of the curriculum they know well, or unlucky if they are well versed in a different part of the curriculum but not the part in the question.
- Authenticity: We may heavily proctor supply response assessments, but this reduces assessable learning to long-term memory, when in real world knowledge processes, we rely (in fact insist!) on knowledge makers using the social memory of published resources, peers and

now AI. We might require hand scribing, but this enforces essentially linear thinking, when the keyboarding to which we have now become accustomed offers the more powerful affordances of multilinear screen scribing. We may allow 'open book' exams without AI, but AI has become a particularly valuable tool to think through complex problems. In all these proctored scenarios, we reduce to severely time-delimited and individualized work the social processes of knowledge creation, which, in the real world, ideally involve relatively open-ended refinement, with collaborative peer, professional and audience feedback.

- Practicality: Supply response assessments are labour-intensive and expensive to assess. Marking has long been the bane of teachers' lives.
- Inter-rater reliability: There is a great deal of rater variability. In the case of high-stakes testing, this requires complex and expensive moderation systems to establish inter-rater reliability (Greenberg, 1992).
- Consequential validity: Feedback is mostly too late and too little to contribute in a meaningful way to learning.

Given these challenges with supply response assessments and as the pressure on education has grown, select response tests have unfortunately become even more prevalent.

We have outlined these assessment datum points because, with the rise of generative AI, it is possible to overcome many of the limitations of supply response assessments (Dixon-Román, 2024; Hao et al., 2024). At this point, supply response assessments can soon be discarded and replaced with something much better.

Working with expert human educators, AI will do a better job with supply response assessments, without the pain and expense of legacy human-only processes. And it can do this at even less cost than machine-graded select response assessments. With generative AI, it will be possible to redirect feedback and assessment systems to deep learning or 'complex epistemic performance' that until now could only be gauged in supply response assessments. Students can engage in knowledge-making activities and represent their thinking in knowledge artifacts. These can be the same kinds of artifacts that have long been presented for supply response assessments, and, of course, nowadays, multimodal artifacts, as well such as videos, infographics and software code. Along the way, students can be asked to externalize their thinking processes. They can explain and justify the ways they have framed and applied knowledge.

Now, generative AI can be working with them in dialogue, along with peer collaborators and the teacher. Indeed, the teacher can be personified in the way they have curated the AI (more on how in the next section).

In this scenario, human and AI formative feedback work with the learner all the way. AI feedback can be continuous, comprehensive and richly nuanced, such that there is no need for that strange, isolated and limited artifact – the summative test. Instead, progress assessment can be a retrospective visualization of assessment that has always been formative in the first instance. Now the conceptual, artifactual and timetable distinction between instruction and assessment disappears. There is no instruction without formative feedback. Assessment is so integrally and essentially embedded in all instruction that, in its legacy forms, assessment disappears.

Teacher effectiveness

The territory of supply response and traditional supply response assessments is familiar. But with AI, teachers must move into new territory. It is a region that, as always, requires domain knowledge. But now the educator needs to learn new ways with which to magnify their knowledge and

distribute their presence in and through AI. It is necessary now to apply a specific educational terminology for tools, by means of which the teacher can re-mediate generative AI.

- Knowledge artifacts: In the new AI pedagogy, knowledge representation tasks can still be drawn from the legacy repertoire of teachers: read about this; listen to this; investigate this; discuss this; consider this; solve this; reflect on this. However, to render their understanding accessible to AI, this cognitive, social and representational work needs to be directed to the creation of digital knowledge artifacts. This, too, is a familiar pedagogical move, not only in the legacy artifacts of writing and image-making (essays, projects, worked examples, etc.), but the broader scope of now readily available multimodal digital media, including videos, data visualizations and software code (Kalantzis and Cope, 2025*a*).
- Epistemic meta-reflection: Beyond the focus on creating knowledge artifacts, students need to externalize their thinking, for instance, not just to write out the solution to a mathematical equation, computer code or a history essay, but also to think aloud about their work, explicitly articulating their underlying knowledge processes. Cognition must be supplemented by a complementary discourse of metacognition. AI can only engage effectively with cognition when cognition is externalized. The AI will be able to tell when this is happening and prompt students when it is not.
- Rubric agents: So far, this is only an updated version of what excellent (although perhaps overworked!) teachers have always done. From here, the teacher needs to move into a new AI space, and for this they will only need pedagogical, not technical skills. The teacher will need to establish a framework of what we term 'rubric agents'. The word 'rubric' will be immediately understood by educators. However, rubric agents spell out in some different ways the kinds of knowledge activity and learning outcome expected of the learner at specified levels of performance. Each criterion in a rubric should be framed as an agent in the discursive sense that they speak like kinds of people from their distinctive epistemic perspectives – say, experiential, conceptual, analytical or applied (Zapata et al., 2024). And while legacy rubrics tersely pack a lot into each criterion and bury their full import in implicit assumptions, the orientation of rubric agents must be spelled out explicitly. The richer that spelling out, the more helpful the agent will be. Technically, this is prompt engineering.
- Knowledge bases: Historically, educators have chosen textbooks and prescribed readings. Usually this was at least enough and, perhaps optimally, a little too much. Now the teacher can create a knowledge base much bigger than what is necessary or feasible for students to absorb in full – say, half a dozen textbooks covering the same subject, articles, links and all kinds of perhaps messy artifacts they have created themselves, such as lesson notes or slide decks. All they need to do is upload or link to this material, and this becomes their own, pedagogical large language model (LLM). Technically, this material is vectorized on upload for 'retrieval augmented generation' or 'cache augmented generation'. This curated corpus constitutes trusted knowledge, validated by the teacher. The knowledge base is then prioritized by the AI over the foundation model. The teacher will, of course, have to be critically aware of cultural, gender and epistemic biases

in their sources, balancing these to the extent this might be needed in the framing of their rubric agents.

- AI thresholds: The teacher can and will need to determine the range of minimum and maximum use of AI. No use of AI is a fail; just as in the past, no reading before writing an essay would be a fail. Total use of AI is a fail; just as in the past plagiarism was always a fail. Now, the teacher can set the target parameters of the learner/AI mix: 80/20, 50/50, 20/80, whatever, depending perhaps on how new the domain is for the student. Contextual AI (represented to the learner and the teacher in data visualizations drawn from analysis of keystrokes, timestamps, clickstream and other logfile data) will make teacher and student constantly aware of their proximity to the target threshold.
- Human/AI moderation: Never AI without human moderation! AI systems can orchestrate this – self reflections, peer reviews and teacher commentary. The humans are the bosses here, and the AI learns the only way it can, by 'reinforcement learning with human feedback'.
- AI assessment: No more old-style grading, because when re-mediated by the educator in the ways we have just described, the AI can give qualitatively richer feedback and quantitatively more reliable ratings than any human teachers – and in seconds, and 24/7, and all the time, and any time. Via learning analytics visualizations based on a massive number of datapoints, the teacher can see which students need their human help and in what specific areas.

We have reached a point now where this can hardly be called AI. The very idea of 'artificial intelligence' is fraught, as if somehow these binary notation machines could replicate human intelligence (Cope and Kalantzis, 2024). When we use this phrase or the acronym, we do so with great reluctance. Human and artificial intelligence are so very different. When the AI talks with the learner, when it gives them support and feedback in the learning environment we have just described, it is really the mediated pedagogical intelligence of the teacher. It is the teacher who has created the rubric agents and defined the domain of study and validated reliable sources for the knowledge base. Behind this, it is the collective human knowledge of the ages in the foundation LLM. Then it is the AI learning about the learner, becoming optimally helpful as it learns what they can and cannot do. Without the humans in the loop, the AI has zero intelligence. This is why we suggest the phrase 'cybersocial learning' to characterize the phenomenon, emphasizing the productive feedback relationship between humans and the machine, rather than to imply that the machine is in any sense replicating human intelligence. Of course, the acronym AI is so pervasive that it is impossible to avoid. We only want to indicate that we mean something at least slightly different.

Our main conclusion here: teachers now have a new job (Kalantzis and Cope, 2025*b*). Their pedagogical expertise will be re-mediated through a cybersocial learning relationship. When the AI speaks, it is the teacher speaking. Now they can focus on the human-to-human things that really need attention. Their intelligence has been used to best effect and magnified. From 1:n, they now have a classroom of simultaneous 1:1. And best of all, a good deal of the tedium in teaching has been removed.

Education equality for all

We have begun each section of this piece with assessment. This is because too often, this end of the pedagogical process determines all that comes before. What,

then, of its legacy purposes? Assessments of all kinds – select or supply response – have until now almost invariably insisted on inequality of outcomes. Sometimes this is not in so vicious and discriminatory a manner as the bell curve, but by design inequality is nevertheless the result. For the few to do very well, the many must be mediocre, and some should fail. This is one of the main functions of education as we have known it until now, to inculcate and establish a moral basis for inequality. Institutionalized in education systems, assessment becomes a key fulcrum for this social project.

However, as every teacher knows, those who didn't do so well in the test could have done just as well as the best if they had received adequate feedback or were given more time before the test to meet its expectations. But relentlessly the bell tolls.

Benjamin Bloom famously articulated the possibility of shifting the curve in his notion of 'mastery learning' (Bloom, 1968). Under what conditions, he asked, can everyone achieve mastery? After all, if a student has been allowed entry to the class, it is on the presupposition that they can achieve the teacher's and the curriculum's expectations.

A responsive AI can determine and then offer feedback on the basis of what we call, revising Lev Vygotsky (1962), a zone of proximal knowledge (ZPK). We don't use the word 'development' as in zone of proximal development (ZPD) because we want this range to be domain- and epistemologically based, rather than a notion rooted in developmental child psychology. The starting point for learning should be within the zone, but with AI we can track the microdynamics of learning and incrementally expand the upper threshold of that zone. Rather than 'mastery', with cyber-social learning we push towards the kinds of epistemic proficiency articulated by the rubric agents.

Cybersocial learning can also calibrate the interactions to engage learner lifeworld differences – students' particular interests, epistemic frames, affinities, emotional nuances and sensibilities. Such education systems should be tuned to learner differences, changing the register of their interactions with learners according to their responses. Over time, the AI will develop a finely calibrated understanding of the learner's needs and potential learning trajectories. This is to move AI from an external support to an intimate extension of personal and collective epistemic life. But it is never to leave the learner in the space where they are to be found; rather to push them into new spaces. In this way, the AI becomes finely tuned to the learner, while embedded within a broader architecture of creative divergence, surfacing unfamiliar perspectives, introducing challenging ideas and fostering dialogical openness.

Finally, this is cheap. Some foundation models are free, and others are less expensive than an individual subscription when purchased by the token for AI education apps. The layers of educational re-mediation we have described above are no more difficult than making a web app. With an inexpensive web-connected device, the best education in the world can be made available to every learner in the world.

Here, by way of conclusion are some key R&D objectives exemplifying what can be built into dedicated AI education workspaces. Over the last several years, we have been building one such educational web app, CyberScholar (Saini et al., 2024; Tzirides et al., 2025). The research accompanying this project has demonstrated ways in which it is possible to create cybersocial learning environments for students and instructors that are safe, productive and ethical. The table that follows compares cybersocial platform pedagogy with the unmediated use of publicly accessible generative AIs.

Table 2: Comparison of cybersocial platform pedagogy with the unmediated use of publicly accessible generative AIs

	Cybersocial platform pedagogy (for instance, CyberScholar)	LLMs 'in the wild'
1. Space for learner activity	Multimodal workspace with full AI support. Thinking with AI.	Question > Answer. The machine thinks for you, 'cognitive offloading' (Gerlich, 2025).
2. Cognitive engagement	Embedded meta-prompts push thinking but will not do the work for you (brainstorm, research, plan, etc. – cache augmented generation).	Will do your work for you. Even when you don't ask for this, it will make an offer that is perhaps too tempting to resist.
3. Provenance of student work	Contextual AI tracks the mix and kinds of generative AI support in relation to the learner's own thinking (keystrokes, time stamps, clickstreams, etc.).	There is no reliable way of knowing the extent or nature of student use of the AI.
4. Learning designs and measures	Rubric agents (prompt engineering), setting the parameters of the learning outcomes, created by the teacher, instructional designer or drawn from a library of shared, educator-created rubrics.	Teacher has no control over prompts, nor an audit trail of dialogue with the LLM.
5. Learning content	Knowledge base (retrieval augmented generation) providing validated subject or disciplinary content, created by the teacher, instructional designer or drawing from a library of educator-curated knowledge bases. Minimizes 'hallucinations' by prioritizing vetted content.	LLM contains general knowledge, not the focused knowledge of a topic or discipline. Factually unreliable.
6. Human-AI relations	The AI can always and should always be moderated by humans (peers, teachers and self), explicitly calling out the question, what are the differences and complementarities between human and machine intelligence?	A private user <-> AI relationship, unmoderated.
7. Learner differences	Calibrated within the learner's changing zone of proximal knowledge (ZPK), while pushing every student to expand their ZPK to reach criterion.	At the mercy of the student-created prompt. Limited prompting skills produces limited outcomes, which are not necessarily aligned with learning objectives.
8. Assessment	Rubic-aligned (quantitative) with detailed rationale (qualitative) formative as-you-go assessment, pushing every student to reach criterion. Updated on the fly, fully transparent to learners. Summative assessment is a retrospective view of progress.	Absent.
9. LLM dependency	Choice of any foundation model you wish.	Tied by account and subscription to one LLM at a time.
10. Privacy	Secure, private, personal LLM, designed to push students with in their ZPK and to expand its range over time.	Hand your identity over to the AI company.

Recapping the three main points of this piece, here is an agenda for dedicated AI in education:

1. Discard assessment of superficial learning and replace with dialogical spaces of deep learning, in which formative feedback is everywhere, embedded and integral to instruction.
2. Now, teaching and learning become sites of cybersocial interaction, transforming the profession of teaching and the schooling lives of learners.
3. With these tools, we can again and with renewed conviction strive for epistemic justice, where every learner has access to education of comparable quality.

The dangers that accompany AI are, of course, manifold. This piece has, however, chosen to focus on the opportunities. It is imperative now to harness the affordances of cybersocial learning in support of educational values that have until now been diminished by heritage assessment practices. By reimagining assessment through generative AI, we are offered a chance not merely to automate old methods, but to reframe what counts as meaningful evidence of learning, growth and understanding. The end of conventional testing may be near, but the future of assessment remains open, a space we must now shape with transformative purpose and social imagination.

References

Biesta, G. J. J. 2016. *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. London, Routledge.

Bloom, B. S. 1968. Learning for mastery. *Evaluation Comment*, Vol. 1, No. 2. Los Angeles, University of California.

Cope, B. and Kalantzis, M. 2024. On cyber-social learning: A critique of artificial intelligence in education. *Trust and Inclusion in AI-Mediated Education: Where Human Learning Meets Learning Machines*. Champaign, Springer, pp. 3–34.

Cronbach, L. J. and Meehl, P. E. 1955. Construct validity in psychological tests. *Psychological Bulletin*, Vol. 52, No. 4. Washington, DC, American Psychological Association, pp. 281–302.

Dixon-Román, E. 2024. AI and psychometrics: epistemology, process, and politics. *Journal of Educational and Behavioral Statistics*, Vol. 49, No. 5. Washington, DC, American Educational Research Association, pp. 709–14.

Gergen, K. J. and Dixon-Román, E. J. 2013. *Epistemology in Measurement: Paradigms and Practices*. Princeton, The Gordon Commission.

Gerlich, M. 2025. AI tools in society: impacts on cognitive offloading and the future of critical thinking. *Societies*, Vol. 15, No. 6. Basel, MDPI, pp. 1–28. https://doi.org/10.3390/soc15010006 (Accessed 5 August 2025.)

Greenberg, K. 1992. Validity and reliability: issues in the direct assessment of writing. *Writing Program Administration,* Vol. 16. West Lafayette, Council of Writing Program Administrators, pp. 7–22.

Hao, J., Von Davier, A. A., Yaneva, V., Lottridge, S., Von Davier, M. and Harris, D. J. 2024. Transforming assessment: the impacts and implications of large language models and generative AI. *Educational Measurement: Issues and Practice*, Vol. 43, No. 2. Bloomington, National Council on Measurement in Education, pp. 16–29.

Kalantzis, M. and Cope, B. 2025*a*. Multiliteracies since social media and artificial intelligence. *Harvard Educational Review*, Vol. 95, No. 1. Cambridge, Harvard Education Publishing Group, pp. 135–51.

Kalantzis, M. and Cope, B. 2025*b*. Literacy in the time of artificial intelligence. *Reading Research Quarterly*, Vol. 60. Newark, International Literacy Association, pp. 1–34. https://doi.org/10.1002/rrq.591 (Accessed 5 August 2025.)

Lord, F. M. 1980. *Applications of Item Response Theory to Practical Testing Problems*. New York, Routledge.

Lord, F. M., Novick, M. R. and Birnbaum, A. 1968. *Statistical Theories of Mental Test Scores*. Reading, Addison-Wesley.

Peters, M. A., Jackson, L., Papastephanou, M., Jandrić, P., Lazaroiu, G., Evers, C. W., Cope, B., Kalantzis, M., Araya, D., Tesar, M., Mika, C., Chen, L., Wang, C., Sturm, S., Rider, S. and Fuller, S. 2023. AI and the future of humanity: ChatGPT-4, philosophy and education – critical responses. *Educational Philosophy and Theory*, Vol. 56, No. 9. London, Taylor & Francis, pp. 840–44.

Saini, A. K., Cope, B., Kalantzis, M. and Zapata, G. C. 2024. *The Future of Feedback: Integrating Peer and Generative AI Reviews to Support Student Work.* Washington, DC, Center for Open Science. https://doi.org/10.35542/osf.io/x3dct (Accessed 5 August 2025.)

Shepard, L. A. 2000. The role of assessment in a learning culture. *Educational Researcher*, Vol. 29, No. 7. Washington, DC, American Educational Research Association, pp. 4–14.

Tzirides, A., Galla, M., Cope, B. and Kalantzis, M. 2025. *Thinking Through AI: Advancing Cognitive and Collaborative Research for AI in Education*. Washington, DC, Center for Open Science. https://doi.org/10.35542/osf.io/s8hqe (Accessed 5 August 2025.)

Uher, J. 2020. Psychometrics is not measurement: unraveling a fundamental misconception in quantitative psychology and the complex network of its underlying fallacies. *Journal of Theoretical and Philosophical Psychology*, Vol. 44, No. 1. Washington, DC, American Psychological Association, pp. 58–84.

Vygotsky, L. S. 1962. *Thought and Language*. Cambridge, MIT Press.

Zapata, G. C., Saini, A. K., Tzirides, A.-O., Cope, B. and Kalantzis, M. 2024. The role of AI feedback in university students' learning experiences: an exploration grounded in activity theory. *Ubiquitous Learning: An International Journal*, Vol. 18, No. 2. Champaign, Common Ground Research Networks, pp. 1–30. https://doi.org/10.18848/1835-9795/CGP/v18i02/1-30 (Accessed 5 August 2025.)

5. Revaluing and recentring human teachers

Keeping the primary goals of education in the AI era: What do educators need to consider?

Ching Sing Chai, Jiun-Yu Wu and Thomas K.F. Chiu

While AI offers ever stronger support for professionals to achieve higher productivity, it is a disruptive technology that challenges all professionals to redefine their roles and articulate acceptable human-AI collaborative models. For teachers and teacher educators, such relational models need to be centred on human development. Emerging empirical findings point towards cautious acceptance, with negative concerns (for example, overreliance, academic integrity, impediments to higher-order thinking) about the use of AI for human development (Hu et al., 2025; Sun et al., 2025). Many argue that AI has the potential to empower students' learning through personalization and open enquiry, as well as providing timely and high-quality feedback (Wu, Jong and Kwok, 2025). Recent studies highlight the benefit of generative AI in addressing teachers' lack of knowledge in interdisciplinary education (Hong et al., 2025), which is important when teachers do not have ready access to professional development. As digital products, well-designed, AI-based pedagogical agents (see Lan and Chen, 2024), online AI platforms (for example, Khan Academy) and intelligent tutoring system (ITS) can be widely distributed to the underprivileged. This has led to the recommendation of international collaboration to leverage leapfrogging with AI to bypass traditional developmental stages (Khan et al., 2024). Nonetheless, AI can also potentially harm students' cognitive and emotional development and encourage cheating and intellectual laziness (Kassenkhan et al., 2025; Zou and Huang, 2023). AI can amplify biases and be manipulated to swing public opinions (Coeckelbergh, 2023; Zhang et al., 2025). Hence, AI is viewed as a double-edged sword that challenges current conceptions of human development, as well as soft skills such as leadership, critical thinking and creative thinking (Tian et al., 2025; Wu, Jong and Kwok, 2025). To address these emerging issues, it is important for educators to re-examine the bases for human development and hence how AI should be deployed. This think piece offers five broad and interrelated theoretical lenses to understand the Artificial Intelligence in Education (AIED) phenomenon: the relational challenges that AI presents, the teleological aims of education, the epistemic limitations of generative AI, the psychological needs for human development and, consequently, the pedagogical design considerations that need attention.

Relational challenges between humans and AI

The core of human existence is fundamentally relational. A valuable form of teaching proposed by Martin Buber is characterized as the formation of an I-Thou relationship through genuine teacher-student encounter (Buber, 1932). Relational teaching emphasizes relation-building before information transaction. It occurs when students and teachers recognize each other as unique beings and are mutually engaged in growing. It is through the I-Thou encounters that educators affirm learners as unique individuals with intrinsic values, worthy of trust, respect and autonomous development. Such encounters propel teacher and learner

into the world of mutual knowing and knowing the world, with the teacher being the institutionally licensed expert to guide learners without compromising learners' freedom. Buber's vision of education is intrinsically human-centric. He paints the teachers as authoritative professionals with knowledge and responsibilities to cultivate learners who could enter I-Thou encounters. Through meeting learners as a Thou in the educational moment, the educator awakens learners' capacities for dialogue, autonomy and reciprocal presence in the world.

Inferring from this person-to-person position, intelligent machines are positioned as I-It relationally, with the machines supporting both teachers and learners. Studies on AIED for English language learning reveal that intelligent machines such as automatic writing evaluation systems and generative AIs can be used to enhance students' writing performances, with differentiated effects on the learners' self-image, but there is a concern about over-reliance on generative AI (Shi et al., 2025). Over-use of machines will affect human development, and educators (including parents) need to regulate appropriately. Zhihui Zhang et al. (2025) report positive outcomes and highlight that generative AI with human feedback is powerful for writing enhancement, especially for competencies associated with higher-order thinking. These studies support person-to-person teaching without denying human-machine usefulness.

Intelligent machines are human-like when prompted to provide answers, and they do not express impatience or ridicule naive questions. Many AIs are outperforming human teachers in providing summarized, high-quality information. They will continue to improve in helping humans with many epistemic tasks as they accumulate backpropagations that better index correct responses. Nevertheless, human-like intelligent machines do not encounter a Thou, nor do teachers or students encounter a Thou when they interact with machines. Machines do not know, care, value, discipline, relate and struggle with the students in the same way as genuine I-Thou encounters. Engineers may create human-like intelligent machines that assume a pseudo Thou, but is that an advisable course of action? Technically, the advancement of natural language processing has rendered the human-machine interactions as indistinguishable from human-human interactions. And this could undermine genuine encounters. Hence, if we accept that education is more than equipping students with knowledge and skills and it needs to entail the facilitation of I-Thou encounters, the conceptualization of education that advocates machines as teachers needs careful examination to determine the nuanced differences. When teachers and AIs are equally capable of providing information and socio-emotional support verbally for students, should the intrinsic differences between humans and machines be treated as irrelevant for the discussion of education? We suggest designing pedagogical I-Thou-It collaborative models that complement each other, with Thou as the pinnacle of the triadic relationship as the way forward. Replacing any of the foundational poles in this historical time is likely harmful. It is necessary to test the evolving relationships within the possible contexts (schools, workplaces, communities, etc.) through the relational perspective. Some emerging research (Hong et al., 2025; Tian et al., 2025; Zou and Huang, 2023) addresses these relationships, but not necessarily through Buber's lens.

Teleological purposes of education and AIED

Gert Biesta (2009, p. 39) explicates that education is essentially 'someone educating someone', with three possible purposes: qualification, socialization and subjectification. His perspective can be viewed as contextualizing Buber's existential positions in the current education arena, and it reinforces the need to understand AIED from a critical standpoint. To counter currently accountability-dominated schooling and measurement-driven discourse, undergirded by treating students as objects or data points to be changed in education, Biesta advocates that all three purposes need to be fulfilled simultaneously in a balanced and interdependent manner. We have not yet read about the design of AI systems based on Biesta's teleological vision of education. Biesta's purposes are directed towards developing Thou (that is, learners) and achieved through constant balancing competing aims. Current AI systems are developed for more specific purposes. AI has been tested to facilitate achievement of qualification, especially with domain-specific intelligent tutoring systems (Kulik and Fletcher, 2016). Its potential in promoting socialization through suggesting useful collaborative knowledge-building activities in the online/blended learning context has also received support (see, for example, Zheng et al., 2024). With regard to the role of AI in supporting critical thinking, creative thinking and other forms of higher-order thinking that are essential for subjectification, emerging research indicates that students prefer and trust human teachers or hybrid systems with a human presence, while teachers are concerned about AI dependence (Fakour and Imani, 2025; Kassenkhan et al., 2025; Zhang et al., 2025). Overall, emerging research indicates that when we dwell more on achieving the socialization and subjectification of Thou, the more we encounter the objectification nature of current AI sciences and technologies, and hence the need to pedagogically position AI as epistemic tools for sense-making on open-ended topics (Wu, Lee et al., 2025).

Good socialization is characterized by the flow of communicative acts that stem from real intentions to interact in genuine encounters. It is an I-Thou (individual or collective) relationship. The nuances of relational knowing, as well as how teachers' care translates into students' socialization, need to be foregrounded to counter current technical rationality prevalent in AIED. Unlike AI outputs that can be ignored, human teachers are responsible and ethically obligated to raise good citizens, and they should not disregard signs of growing poor socialization and subjectification. Subjectification arises from the I-Thou encounters, and it emphasizes individuals as the subject of their own being, who actively construct their unique self, rather than being the objects of others, passively being shaped by others. Socialization and subjectification require human modelling to exhibit the courage and wisdom of being a unique individual. Biesta's perspective further extends the need to adopt a human-centric perspective for AIED as advocated by UNESCO (Miao et al., 2024). Technologists and teachers should design and review progress in AIED not just from the qualification perspective but also incorporate socialization and subjectification perspectives. We foresee that for the latter, qualitative studies that document how learners' identities evolve in pedagogical environments grounded in social constructivist learning theories would be helpful to clarify the way forward. Ting Tian et al. (2025) document how university students' epistemic agency evolved when they were empowered to use generative AIs to pursue inquiry in a knowledge-building community. While these students from a research university are able negotiate a new

identity with AI, we are less optimistic about others who are less advantaged. There is a need for longitudinal studies (three to five to ten years) of AI effects on the socialization and subjectification of diverse learners.

Epistemic concerns: Human vs. machine ways of knowing

AI lacks true epistemic agency – it processes inputs and runs algorithms on its databases to generate the most probable outputs. It does not 'know' or seek truth as humans do (Hicks et al., 2024). Humans interact with physical and sociocultural environments to form beliefs and ideas about the seen and unseen worlds. To seek truths, some of these ideas are subjected to ways of knowing through which the knowing agents articulate the claims they aim to establish, the reliable processes they are engaging, and the epistemic standards they abide by for public scrutiny. Clark Chinn et al. (2011) characterize such processes as epistemic cognition. Epistemic cognition involves physical, cognitive and sociocultural processes that differ based on disciplines to establish socially accepted knowledge. It is essential to recognize that human ways of knowing are different from machine ways of predicting.

The increasing focus on the epistemic issues associated with AI highlights the need to pay specific attention to this dimension in AIED. Given that generative AIs cannot be trusted as epistemic authorities because of their underlying processes (Hicks et al., 2024), AI in narrow domains such as ITS are designed narrowly for qualification (that is, to guide users towards a single, correct answer). In addition, given that current machine-generated outputs are outperforming many human teachers, AIs are viewed as epistemic tools and resources that could spur epistemic cognition. Recent studies report that students and teachers adapt to epistemic characteristics of generative AIs by assuming epistemic agency to critically review their output (Hong et al., 2025; Tian et al., 2025). Emerging studies are examining the roles of AI in mediating sensemaking (Sivola et al., 2025). To educate involves enabling learners to seek truth with adaptable epistemic posture, valid goals, and trustworthy justification and verification methods (Wu, Lee et al., 2025). It seems that the use of generative AIs naturally demands users to act as active epistemic agents who review, evaluate and synthesize generated outputs for knowledge inquiry. However, generative or other forms of AI are insufficient to support fruitful knowledge quests. For example, natural sciences require learners to interact with the natural world (Tang and Cooper, 2024); social sciences and humanities require learners to understand the first-hand knowledge of human beings. It seems obvious that as AI progresses, deepening learners' responsibility to assume epistemic agency, emphasizing the primacy of first-hand data, and building learners' capacities to engage in valid and reliable knowledge work, are accentuated. If users understand the nature of AIs, they can utilize AIs as a powerful epistemic partner to facilitate epistemic works. Huang-Yao Hong et al. (2025) research how teachers make use of generative AIs to navigate their lack of knowledge for interdisciplinary STEM education, with full awareness of the machines' limitation as a text generator. Anni Sivola et al. (2025) examined students' sensemaking with AI outputs, which led the students to realize that the outputs are lacking in creativity and insights. The shortcomings of AI can be a context that frames users' epistemic cognition for productive outcomes. Future human development efforts can leverage the epistemic strengths and weaknesses of humans and machines to formulate productive collaboration models in the context of open-ended inquiries. This

direction could facilitate many situated, meaningful encounters for subjectification. Gao et al. (n.d.) examine how AI workflows facilitate secondary students' epistemic cognition for engineering product design targeted at promoting the social good.

Psychological risks of AI in learning

Psychologically, the current concerns of over-reliance on AI (Kassenkhan et al., 2025; Zou and Huang, 2023) may undermine the three basic intrinsic human needs for autonomy, competence and relatedness (Ryan and Deci, 2017). These needs are congruent with Buber's I-Thou position and Biesta's perspective on education. The concerns of over-reliance on AI necessitate careful pedagogical design to ensure AI supports rather than hinders holistic human development.

While AI offers efficiency and personalized learning, excessive dependence risks diminishing students' abilities to think independently and critically, solve problems creatively and engage meaningfully with others (Stonjanov et al., 2024). Autonomy – the need for self-directed action – may weaken if AI makes too many decisions for learners. Reliance on shortcuts could erode critical cognitive thinking. Students can use AI to complete challenging tasks without cognitive learning; for example, they can use AI to create videos for their presentations without deep thinking. Relatedness – the desire for human connection – might suffer if AI-mediated interactions replace authentic collaboration. Students could ask AI for help instead of humans; they could trust AI more than humans and care less about building relationships. Given that these systems are often trained on data containing gender biases, uncritical use may reinforce stereotypes and undermine progress toward gender equality (UNESCO, 2024). To prevent these pitfalls, educators must intentionally design learning environments where AI augments rather than replaces human agency (Chiu and Rospigliosi, 2025). This involves fostering metacognitive skills, encouraging AI-assisted reflection rather than passive consumption, and prioritizing activities that require emotional intelligence and teamwork (Zhou et al., 2025). Current interventions that build on self-determination theory (Ryan and Deci, 2017) rely on the teachers to fulfil the psychological needs. Educators must guide learners to focus on developing their competencies in higher-order thinking, ethical reasoning and interpersonal growth. Education must harness AI's benefits, while preserving the irreplaceable human elements of curiosity, empathy and well-being.

Pedagogical responsibility: Educators must lead AI integration

Brayan Diaz and Miguel Nussbaum (2024) emphasize the incorporation of pedagogical intelligence for AIED. While AI exhibits strong affordances for education, it is an epistemic technology that needs to be pedagogically tamed for various subject matter or interdisciplinary cognition. In other words, teachers are required to design new technological pedagogical content knowledge for AIED (Celik, 2023; An et al., 2023), and this is a difficult endeavour that takes time and wisdom. Existing studies support that AI serves as an educational tool – not a replacement – for meaningful human interaction and intellectual growth. Students and teachers are beginning to accept AI use, but with an emerging understanding of its limitations (Hong et al., 2025; Ravšelj et al., 2025; Tian et al., 2025). Resources such as UNESCO's *Red Teaming AI for Social Good* (2025) provide practical guidance for educators and students to collaboratively test generative AI models for flaws and vulnerabilities that may uncover harmful behaviour. Educators should design learning experiences that leverage AI's

strengths – such as instant feedback and data analysis – along with consideration of the issues from the above perspectives. By framing the pedagogical design with these references, teachers may be better able to cultivate well-rounded learners who thrive in an AI-driven world without losing their humanity.

Conclusion: Balancing AI's benefits with human-centric education

As advances in AI find ways to be more human-like, we are likely to face a period of ambiguity. Educators are currently designing and testing new pedagogical practices and investigating naturally occurring human adaptations to technological advancements. These efforts are moving us closer to the explication of theoretically supported and practically helpful human-AI collaboration models. Meanwhile, we must exercise caution to prevent technological impacts from overwhelming the needs of human development. Educators are obliged to ask not only how to integrate AI into education but also how the integrated instructional practices can facilitate human subjectification and how learning experiences can be shaped to foster human capacities for better living.

References

An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C. and Chen, M. 2023. Modeling English teachers' behavioral intention to utilize artificial intelligence in middle schools. *Education and Information Technologies*, Vol. 28, No. 1. Berlin, Springer Nature, pp. 5187–5208. https://doi.org/10.1007/s10639-022-11286-z (Accessed 8 August 2025.)

Biesta, G. 2009. Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, Vol. 21. Berlin, Springer Nature, pp. 33–46.

Buber, M. 1932. *I and Thou* (R. G. Smith ,Trans.). Edinburgh, T. & T. Clark.

Celik, I. 2023. Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, Vol. 138. Amsterdam, Elsevier. https://doi.org/10.1016/j.chb.2022.107468 (Accessed 8 August 2025.)

Chinn, C.A., Buckland, L.A. and Samarapungavan, A. 2011. Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, Vol. 46, No. 3. London, Taylor & Francis, pp. 141–167.

Chiu, T. K. F. and Rospigliosi, P. A. 2025. Encouraging human-AI collaboration in interactive learning environments. *Interactive Learning Environments*, Vol. 33, No. 2. London, Taylor & Francis, pp. 921–924. https://www.tandfonline.com/doi/full/10.1080/10494820.2025.2471199 (Accessed 8 August 2025.)

Coeckelbergh, M. 2023. Narrative responsibility and artificial intelligence: How AI challenges human responsibility and sense-making. *AI & SOCIETY*, Vol. 38, No. 6. Berlin, Springer Nature, pp. 2437–2450. https://doi.org/10.1007/s00146-021-01375-x (Accessed 7 August 2025.)

Díaz, B. and Nussbaum, M. 2024. Artificial intelligence for teaching and learning in schools: The need for pedagogical intelligence. *Computers & Education*. Amsterdam, Elsevier, Vol. 217. https://doi.org/10.1016/j.compedu.2024.105071 (Accessed 8 August 2025.)

Fakour, H. and Imani, M. 2025. Socratic wisdom in the age of AI: A comparative study of ChatGPT and human tutors in enhancing critical thinking skills. *Frontiers in Education*, Vol. 10. Lausanne, Frontiers Media. https://doi.org/10.3389/feduc.2025.1528603 (Accessed 8 August 2025.)

Gao, L., Jong, M. and Chai, C. S. n.d. Understanding secondary students' epistemic cognition in the design process with the support of a personalized multi-agent system. Unpublished (Forthcoming).

Hicks, M. T., Humphries, J. and Slater, J. 2024. ChatGPT is bullshit. *Ethics and Information Technology*, Vol. 26, No. 2. Berlin, Springer Nature, pp. 1–10. https://doi.org/10.1007/s10676-024-09775-5 (Accessed 8 August 2025.)

Hong, H. Y., Chen, M. J., Chang, C. H., Tseng, L. T. and Chai, C. S. 2025. AI-supported idea-developing discourse to foster professional agency within teacher communities for STEAM lesson design in knowledge building environment. *Computers & Education*, Vol. 229. Amsterdam, Elsevier.

Hu, X., Gong, W. and Cortazzi, M. 2025. Researching China's pre-service teachers' perceptions of AI education for K-12: An elicited metaphor analysis. *European Journal of Education*, Vol. 60, No. 2. Hoboken, John Wiley & Sons.

Kassenkhan, A. M., Moldagulova, A. N. and Serbin, V. V. 2025. Gamification and artificial intelligence in education: A review of innovative approaches to fostering critical thinking. *IEEE Access*, Vol. 13. New York, IEEE, pp. 98699–98728. https://doi.org/10.1109/ACCESS.2025.3576147 (Accessed 7 August 2025.)

Khan, M. S., Umer, H. and Faruqe, F. 2024. Artificial intelligence for low income countries. *Humanities and Social Sciences Communications*, Vol. 11, No. 1. Berlin, Springer Nature, pp. 1–13. https://doi.org/10.1057/s41599-024-03947-w (Accessed 7 August 2025.)

Khan, S. 2024. *Brave New Words: How AI Will Revolutionize Education (and Why That's a Good Thing)*. London, Penguin.

Kulik, J. A. and Fletcher, J. D. 2016. Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, Vol. 86, No. 1. Washington DC, American Education Research Association, pp. 42–78.

Lan, Y.-J. and Chen, N.-S. 2024. Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents. *Educational Technology & Society*, Vol. 27, No. 1. Taipei, National Taiwan Normal University, pp. 1-18.

Miao, F., Shiohira, K. and Lao, N. 2024. *AI competency framework for students*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391105 (Accessed 8 August 2025.)

Ravšelj, D., Keržič, D., Tomaževič, N., Umek, L., Brezovar, N., et al. 2025. Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE*, Vol. 20, No. 2. San Francisco, PLOS. https://doi.org/10.1371/journal.pone.0315011 (Accessed 8 August 2025.)

Ryan, R. M. and Deci, E. L. 2017. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York, Guilford Press.

Shi, H., Chai, C. S., Zhou, S. and Aubrey, S. 2025. Comparing the effects of ChatGPT and automated writing evaluation on students' writing and ideal L2 writing self. *Computer Assisted Language Learning*. London, Taylor & Francis, pp. 1–28. https://doi.org/10.1080/09588221.2025.2454541 (Accessed 8 August 2025.)

Silvola, A., Kajamaa, A., Merikko, J. and Muukkonen, H. 2025. AI-mediated sensemaking in higher education students' learning processes: Tensions, sensemaking practices, and A-assigned purposes. *British Journal of Educational Technology*. Hoboken, John Wiley & Sons. https://doi.org/10.1111/bjet.13606 (Accessed 8 August 2025.)

Stojanov, A., Liu, Q. and Koh, J. H. L. 2024. University students' self-reported reliance on ChatGPT for learning: A latent profile analysis. *Computers and Education: Artificial Intelligence*, Vol. 6, No. 4. Amsterdam, Elsevier. http://dx.doi.org/10.1016/j.caeai.2024.100243 (Accessed 8 August 2025.)

Sun, Y., Yang, H., Yu, H. K. and Suen, R. 2025. Boon or bane? Evaluating AI-driven learning assistance in higher education professional coursework. *Education and Information Technologies*. Berlin, Springer Nature, pp. 1–34.

Tang, K. S. and Cooper, G. 2025. The role of materiality in an era of generative artificial intelligence. *Science & Education*, Vol. 34. Berlin, Springer Nature, pp. 731-746. https://doi.org/10.1007/s11191-024-00508-0 (Accessed 7 August 2025.)

Tian, T., Chai, C. S., Chen, M.-H. and Liang, J.-C. 2025. University students' perception of learning with generative artificial intelligence. *Educational Technology & Society*, Vol. 28, No. 3. Taipei, National Taiwan Normal University, pp. 151–165.

UNESCO, 2024. *"Your opinion doesn't matter, anyway": exposing technology-facilitated gender-based violence in an era of generative AI*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000387483 (Accessed 8 August 2025.)

—. 2025. *Red Teaming artificial intelligence for social good - The PLAYBOOK*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000394338 (Accessed 8 August 2025.)

Wu, J.-Y., Jong, M. S.-Y. and Kwok, O.-M. 2025. Guest editorial: Application and research of generative AI in education. *Educational Technology & Society*, Vol. 28, No. 3. Taipei, International Forum of Educational Technology and Society, pp. 1–3.

Wu, J.-Y., Lee, Y.-H., Chai, C. S. and Tsai, C.-C. 2025. Strengthening human epistemic agency in the symbiotic learning partnership with generative artificial intelligence. *Educational Researcher*. Washington DC, American Educational Research Association. https://doi.org/10.3102/0013189X251333628 (Accessed 8 August 2025.)

Wu, Y., Zhang, W. and Lin, C. 2025. Generative artificial intelligence in university education. *IT Professional* Vol. 27, No. 2, pp. 69–74.

Zhang, Z., Aubrey, S., Huang, X. and Chiu, T. K. 2025. The role of generative AI and hybrid feedback in improving L2 writing skills: A comparative study. *Innovation in Language Learning and Teaching*. London, Taylor & Francis, pp. 1–19. https://doi.org/10.1080/17501229.2025.2503890 (Accessed 7 August 2025.)

Zheng, L., Fan, Y., Chen, B., Huang, Z., LeiGao and Long, M. 2024. An AI-enabled feedback-feedforward approach to promoting online collaborative learning. *Education and Information Technologies*, Vol. 29, No. 9. Berlin, Springer Nature, pp. 11385–11406.

Zhou, X., Li, Y., Chai, C. S. and Chiu, T. K. 2025. Defining, enhancing, and assessing artificial intelligence literacy and competency in K-12 education from a systematic review. *Interactive Learning Environments*. London, Taylor & Francis, pp. 1–23. https://doi.org/10.1080/10494820.2025.2487538 (Accessed 8 August 2025.).

Zou, M. and Huang, L. 2023. The impact of ChatGPT on L2 writing and expected responses: Voice from doctoral students. *Education and Information Technologies*, Vol. 29, No. 11. Berlin, Springer Nature, pp. 13201–13219. https://doi.org/10.1007/s10639-023-12397-x (Accessed 7 August 2025.)

Compassion by design: Building AI with and for caring educators

Arafeh Karimi

Introduction

> 'In one rural classroom, nine-year-old Aisha grew silent halfway through an AI-powered math activity, her shoulders slumping as the system moved on without her. Her teacher, Ms Devi, later explained, "The AI wouldn't pause for her half-correct answer; she just shut down."'[1]

When AI overlooks empathy and relational nuance, it risks eroding trust, learner agency and inclusive, social-emotional development. Too many so-called human-centred EdTech tools arrive as closed-box, efficiency-driven products; automating tasks while neglecting educator empowerment, learner belonging and nurturing pedagogy. Prioritizing technical optimization over equity and care, risks hollowing out the human core of teaching (UNESCO, 2025).

This think piece proposes a 'compassion by design' approach: seven actionable policy shifts that ensure AI listens, defers and co-learns with educators. Five foundational shifts protect agency, trust and belonging.

Two transformational shifts embed ethical transparency, adaptive wayfaring[2] and sustainable data stewardship. Foundational shifts set minimum policy guardrails to protect educator agency, trust and inclusion; transformational shifts build longer-term systemic levers that reimagine pedagogy and governance. Together, they form a values-aligned blueprint, grounded in UNESCO's vision for inclusive, human-centred digital learning (UNESCO, 2024). This piece primarily supports national ministries and policy-makers, with secondary guidance for EdTech vendors and local education leaders.

Compassion by design

As UNESCO 2025 affirms, integrating AI must not replace relational pedagogy with automated efficiency. We must shift from automation, where AI dictates pace and content, to attunement, where systems respond to educator judgement, learner emotion and local context. A 'compassion by design' approach ensures that AI systems listen, defer and co-learn with educators and students, preserving the deeply relational core of education.

Figure 1: Compassion by design: Seven shifts to align AI with human-centred education

Foundational shifts (Establish minimum guardrails)	Transformational shifts (Build longer-term systemic levers)
Shift 1. Participatory co-design Shift 2. Equity explainability Shift 3. Trust and well-being audits Shift 4. Nurturing pause workflows Shift 5. Belonging and reflection assessment	Shift 6. Adaptive wayfaring and narrative continuity Shift 7. Ethical data stewardship

1. Author's field notes, Malaysia, 2022-2025 (Product Leader, FrogAsia Company)

2. Wayfaring, coined by Tim Ingold (2007), is where progress is lived navigation through terrain, not waypoint hopping; AI should notice spirals, back eddies and side trails.

Conceptual foundations

This framework builds on two research grounds:

- Human-centred AI, which informs Shifts 1–3, focuses on participatory design, explainability and emotional attunement. Shift 1 draws on co-design (Alfredo et al., 2024; Simonsen et al., 2020). Shift 2 builds on the concept of epistemic justice (Fricker, 2007). Shift 3 centres relational trust and pedagogical care (Bryk and Schneider, 2003; Noddings, 2005). Shift 4 integrates feminist ethics of care, addressing the emotional labour of educators (Haraway, 2016). Complementing these, the Positive Artificial Intelligence in Education (P-AIED) framework (Bittencourt et al., 2023) shows how Artificial Intelligence in Education (AIED) can be purpose-built to foster student well-being, resilience and meaning, rather than narrow performance gains.
- Critical AI and society, which informs Shifts 5–7, interrogates power, equity and data ethics. Shift 5 embraces entangled pedagogy and learner reflection.[3] Shift 6 invokes Tim Ingold's 'wayfaring' metaphor: AI as a compass for culturally responsive journeys (Ingold, 2007). Shift 7 draws from Abeba Birhane's relational ethics (2021) to promote bias disclosure and educator-led data stewardship.

These conceptual anchors echo UNESCO's call to embed empathy, participation and educator agency into AI policy. The seven shifts that follow translate this vision into practical steps for building attuned, relational and just AI in education.

Foundational shifts: From automation to attunement

What happens when AI doesn't pause, and how do we rebuild trust from that silence?

Goal: Reset the baseline so no AI deployment compromises teacher agency, learner belonging or social-emotional well-being.

These shifts move from rigid automation towards attuned, human-centred education. They embed empathy, agency and inclusion into the design and use of AI tools in learning.

Shift 1: Participatory co-design

How do we ensure the next tool pauses when a child like Aisha does?

Action: Require all AI procurements to demonstrate ongoing, participatory co-design with educators, including joint planning workshops, iterative pilots and documented feedback loops.

Rationale: When educators help shape AI from the start, technology becomes a responsive partner rather than a black-box automation tool. Co-design protects teacher agency, fosters pedagogical trust and ensures AI serves real classroom needs. Co-design centres the voices of rural, Indigenous and other under-represented educators, preventing 'digital colonialism' and ensuring AI reflects diverse classroom realities.

Implementation: Ministries ask two key procurement questions:

- How were educators involved in shaping this tool?
- What documented evidence shows teachers' input informed the design?

3. In 'entangled pedagogy', learning emerges from inseparable knots of humans, tools and contexts; design starts from relationships, not functions (Fawns, 2022).

Proposals must include clear co-design artefacts, such as workshop notes, teacher prototypes or pilot feedback. Bids lacking this are returned for revision, with professional development funds paused until compliance. A short annex offers vendors sample artefacts and mock-ups, clarifying what meaningful co-design looks like.

Practice snapshot: In Brazil, over 60 educators co-created the *Guia IA para Educadores* with CIEB and the Ministry of Education (UNESCO and CIEB, 2024). They set priorities for how AI should serve writing instruction. As a result, pilot tools now directly reflect teachers' goals (CIEB, 2024). This process is influencing Brazil's national EdTech procurement policy, embedding teacher voice into future AI deployments.

Shift 2: Equity-driven explainability

Who defines what counts as knowledge and how do we make this visible and contestable?

Action: Mandate that all educational AI include equity-driven explainability features that expose whose knowledge, cultural frames and epistemic assumptions shaped the model and create clear pathways to challenge or adapt these. This moves beyond technical transparency to protect diverse ways of knowing, especially critical in under-resourced or culturally plural settings.

Rationale: Traditional explainability only reveals how a model works. Equity-driven explainability exposes who it serves and whose knowledge it overlooks, making it a tool of epistemic justice that protects learner identity, trust and belonging. By surfacing whose knowledge is absent, equity-driven explainability protects minoritized learners from epistemic erasure and equips them to contest biased outputs.

Implementation: Ministries include two clear requirements in tenders:

- Whose knowledge shaped this model, and how is that documented?
- What mechanisms allow teachers and students to question or contest outputs?

Vendors must provide culturally explicit model cards and show simple, multi-language pathways for contesting AI outputs. Bids without these are returned for revision. In low-resource settings, ministries can mandate SMS or paper-based systems to keep contestability accessible.

Practice snapshot: This builds on a wave of new regulation. The EU *AI Act* requires vendors to publish plain-language summaries detailing data sources, gaps and bias mitigation (European Commission, 2024). Brazil's Draft AI Bill contains parallel transparency and *explicabilidade* requirements, aimed at protecting cultural rights (Federal Senate, Brazil, 2024). In education, this means ministries can mandate disclosures that show not just how an AI works, whose epistemologies it encodes, and equips educators to critically question or adapt it.

This is supported by Riordan Alfredo et al. (2024), who found co-design paired with explainability features increases educator trust, adoption and classroom relevance.

Shift 3: Trust and well-being audits

When trust breaks down, how do we rebuild it with students, teachers and communities alike?

Action: Embed short, rights-based 'trust and well-being audits' into all AI pilots. Co-authored by educators, students and communities, audits assess relational trust, inclusion, learner agency and emotional safety.

Rationale: Accuracy alone does not ensure dignity. Trust audits reveal early harm signals and ensure AI supports equitable, inclusive learning. By spotlighting how the system treats girls, rural students and other marginalized groups, the audit becomes a guardrail against widening the AI divide.

Implementation: Ministries fund mixed audit teams made up of psychologists, counsellors and teachers. Teams run a pre-pilot baseline, a mid-term check and a final review. In lower-resource settings, audits can use simple rubrics developed with teacher leaders and run through peer networks.

Practice snapshot: In 2023, New South Wales, Australia, launched a mandatory AI Assurance Framework (NSW Department of Education, 2023). Every AI tool must pass this audit to be approved in schools. Reviewers assess whether tools support well-being, agency and inclusion. A career guidance tool was rejected in 2024 after the audit found gender bias and weak rural representation. The vendor revised its model and was later approved. This framework shows that well-being audits can be scaled across public systems and hold vendors accountable to human-centred values, not just technical accuracy.

This model reflects a growing global norm, as seen in the Government of Uruguay's (2024) pilot, adapting UNICEF's (2021) *Policy Guidance on AI for Children*. It also aligns with Olga Viberg et al. (2024), who found emotional safety and learner autonomy significantly predict trust in EdTech.

Shift 4: Nurturing pause workflows

What would it take for Ms Devi to pause the AI and be supported in doing so?

Action: AI tools should be required to include configurable 'pause windows' so teachers can reflect, connect or adjust lessons. Monthly peer support groups should be funded to address ethical and emotional challenges.

Rationale: Fast-paced AI can suppress empathy and erode key human interactions. Pause windows and reflection groups restore emotional safety, relational care and teacher voice. Pause windows especially benefit students who are often overlooked by fast-moving algorithms, such as language-minority, neurodivergent or trauma-affected learners, restoring space for care.

Implementation: Ministries revise AI standards, requiring pause functions and time for teacher reflection. Where funding is limited, protected time is scheduled into the school calendar for 'AI Reflection Groups'.

Practice snapshot: Carnegie Learning's MATHia LiveLAB Dashboard alerts teachers when students are struggling, prompting them to pause AI delivery and step in with support (Sloan, 2024). These insights are discussed in weekly Professional Learning Communities, where teachers share approaches and adapt strategies. Independent reviews show lower student anxiety and increased teacher confidence when AI includes pause and reflection features.

Finland's School Day dashboard and Spain's ODITE '*pausas reflexivas*' show similar benefits (Ministry of Education and Culture, Finland, 2024; ODITE, 2025). When teachers are empowered to interrupt automation and reflect together, AI becomes a tool for care, not just efficiency.

Low-cost options: Pause-and-reflect workflows do not require expensive tools. Schools can use open-source dashboards (for example, Kolibri Coach), simple surveys and peer-led groups. In Rwanda, VVOB's learning circles support emotional check-ins with time, not tech (VVOB, 2021). A designated 'well-being lead' can coordinate this effort.

Shift 5: Belonging and reflection assessment

How do we know when a student feels unseen, and how can AI assessments help us listen better?

Action: Require all AI-assisted assessments to include a brief 'Inclusive belonging index' (for example, adapted from OECD and India's Thriving index) and a short student reflection

on how the AI feedback resonated with their lived experience. Ministries work with educators to localize the design.

Rationale: Scripted answers alone miss relational and emotional dimensions. Including prompts like 'I felt heard' or 'I could ask for help' centres the learner experience. Reflections surface bias, affirm identity and promote metacognition. Belonging data reveals hidden disparities across gender, caste, disability and culture, allowing timely support before small exclusions become structural gaps.

Implementation: Ministries integrate the Belonging index into existing dashboards (for example, EMIS) and publish anonymized school-level results alongside attainment data. Examination boards may allocate up to 10 per cent of marks in AI-assisted assignments for student reflections, encouraging active engagement with feedback rather than passive acceptance.

Practice snapshot: In India (Ministry of Education, India, 2020), Dream a Dream's Thriving Index, used with 100,000-plus students, measures life skills like agency and empathy. When it flagged a drop in girls' agency during early secondary years, schools responded with targeted mentoring and social and emotional learning. Recent impact data show 93.9 per cent of 7,434 learners improved on the Life-Skills Assessment Scale, with 75 per cent gaining at least one standard deviation; girls out-paced boys (76 per cent vs 73 per cent) (Dream a Dream, 2023). An AI system trained on these rubrics could prompt reflective questions post-assignment and offer teachers real-time insights into students' well-being and sense of belonging.

OECD identifies student voice and reflection as key predictors of equity, with schools that enable regular learner reflection scoring higher on belonging and emotional safety (OECD, 2023; 2025). Embedding reflection and belonging metrics into AI assessments ensures they listen, not just score, and support more equitable pedagogical redesign.

Transformational shifts: Care, ethics and wayfaring

How do we design an AI ecosystem that honours learning as a shared journey and data as a responsibility we carry together?

The five foundational shifts establish a necessary baseline of relational trust and educator agency. Building on this stable ground, the two Transformational shifts then propose to fundamentally reshape pedagogical practice and data governance, moving from incremental improvement to systemic evolution.

Goal: Once trust and educator agency are secured, these advanced shifts embed sustainable, values-centred practices, ensuring AI remains a caring, future-ready companion.

'Care, ethics and wayfaring' represents a significant evolution from prescriptive curricula towards dynamic, responsive educational journeys. Drawing from Ingold's (2007) metaphor of wayfaring, education shifts from predefined paths to emergent navigation, cultivating learners who actively orient, adapt and choose wisely within unfolding landscapes. Ethical transparency and educator-led data stewardship complement this approach by reinforcing care, relational ethics and cultural responsiveness, thereby anchoring learning within enduring values of integrity and inclusivity.

Shift 6: Adaptive wayfaring and narrative continuity

What if learning isn't a path to follow, but a story we co-author, step by step, with care?

Action: Launch annual wayfaring micro-grants for schools to pilot flexible, learner-centred AI modules offering culturally responsive pathways (for example, project-based learning, scaffolded inquiry, multimedia). Teachers and students co-navigate and adapt these in real-time. Each pilot builds a Learning narrative archive; a record of AI reflections, feedback loops and student journals, to sustain narrative continuity.

Rational: Traditional curricula provide rigid maps that often fail to reflect learners' dynamic, lived experiences. Wayfaring honours agency, resilience and discovery, aligning learning with real-world contexts. Archiving learning narratives fosters trust, continuity and cultural relevance, helping educators build on past insights and support each learner's evolving journey. Wayfaring grants let learners pursue culturally sustaining pathways, countering one-size-fits-all curricula that privilege dominant narratives and disenfranchise local knowledge.

Implementation: Ministries co-sponsor annual grants for schools to propose AI-supported, adaptable learning pathways shaped by students. Awardees share concise 'Adaptation Stories' showcasing real-time, local innovations. A Data Stewardship Fund supports training for district officers in ethical archiving to ensure secure, accessible narrative continuity. Ministries can also promote open-source AI tools to enable flexible pedagogy without costly licences.

Practice snapshot: The *Redes de Tutoría* model (Rincón-Gallardo, 2023), used in thousands of Mexican public schools, flips traditional instruction: students choose a topic they care about, master it and teach it to a peer. Learning unfolds through curiosity, dialogue and shared exploration, building agency, motivation and metacognitive depth. A wayfaring micro-grant could support the next step: an AI-enhanced platform like a hypothetical ConexionesAI, co-designed with educators, would scaffold and archive these journeys. Instead of prescribing content, the AI could suggest branching paths (for example, interviews, maps, artefacts), prompt peer reflection, and help learners build personalized *Rutas de Aprendizaje*, rich portfolios of how they explored and adapted their paths. This illustrates how AI can extend learner-led pedagogy and deepen culturally grounded education, particularly in the Global South.

Shift 7: Ethical data stewardship

If data is the story we tell about students, who holds the power to shape it, and who protects what truly matters?

Action: Mandate all AI vendors to publish plain-language values and bias disclosures, co-developed with educators and communities. These must outline data sources, known biases, risks and mitigation steps. Simultaneously, establish an educator-led data stewardship fund to train school-based stewards in equity audits, ethical governance and privacy safeguards.

Rational: Opaque AI erodes trust, inclusion and educator agency. Transparent disclosures empower teachers to question AI outputs, while stewardship training equips them to protect privacy, cultural continuity and student rights, anchoring AI in care and ethical practice. School-based data stewards defend communities with weak legal protections, ensuring privacy, cultural continuity and fair use of data in low-resource settings.

Implementation: Bias disclosures and data stewardship plans are required as conditions for AI procurement. Ministries fund annual training for school-based stewards and host workshops to share best practices in inclusive

auditing, secure archiving and participatory review. This embeds ethical governance within school systems.

Practice snapshot: Under the EU's 2024 *AI Act* (European Commission, 2024), educational AI systems are classified as 'high risk', requiring vendors to publish plain-language summaries of training data, representational gaps and mitigation strategies before school deployment, setting a new legal standard for transparency.

Building on the EU *AI Act*, in Portugal, the Ministry of Education is piloting educator-led stewardship in 30 secondary schools (OECD, 2024). Teachers and ICT leads are trained to review AI tools for equity, pedagogical fit and cultural relevance. They are empowered to question vendor disclosures and shape procurement decisions.

These examples show how legal mandates and school-based oversight can work together, ensuring AI tools align with inclusive, context-sensitive education goals.

Translating these shifts into practice demands system-level scaffolding. Education ministries need clear political backing (cabinet-level mandates), stable multi-year funding lines and data-sharing agreements that link curriculum, teacher-training, child-protection and ICT portfolios. Inter-ministerial task forces – including in education, digital affairs, finance and child-rights – should co-issue standards, pool expertise and align budgets so that compassionate, equity-centred AI becomes a whole-of-government priority rather than a siloed pilot.

Equity lens

The AI divide is ultimately a question of whose knowledge, pace and privacy count. By bringing marginalized educators into design discussions through co-design, requiring equity-driven explainability that lays bare a model's cultural lens, running trust and well-being audits before scale-up, and embedding pause windows that let teachers shield vulnerable learners from one-speed automation, systems start to level power. An 'Inclusive belonging index' amplifies unheard student voices, wayfaring micro-grants finance culturally grounded, learner-shaped pathways, and educator-led data stewardship keeps privacy and context close to communities, so that equity is woven through every technical layer and governance choice. Together, these guardrails keep 'Compassion by design' from becoming another form of digital privilege.

Conclusion

We began by identifying how automation-first AI in education can erode educator agency, learner belonging and emotional well-being. The five foundational shifts – participatory co-design; explainability; rights-based audits; slow-paced nurturing workflows; and holistic assessment with learner voice – reset the baseline for relational trust, equity and inclusive pedagogy.

Building on this, the two transformational shifts – adaptive wayfaring with narrative continuity; and ethical transparency with educator-led data stewardship – embed values of care, cultural responsiveness and sustainable governance. Together, these shifts reflect global priorities of empathy, dignity, equity and the co-creation of just futures.

This is not a debate between tradition and technology. It is an opportunity to reimagine AI as an attuned educational partner who listens, defers and co-evolves with educators and learners.

Policy-makers can embed 'Compassion by design' into procurement, pilot funding and professional learning. EdTech developers must co-design from the outset, prioritizing explainability, safety and adaptability.

Ministries and researchers can trial these shifts in diverse school settings and capture equity-centred data to guide policy. Educational communities, including teachers, students and families, must be recognized as co-stewards who shape local ecosystems with cultural wisdom and care.

By choosing 'Compassion by design', we do not just mitigate harm. We uphold the moral imperative to honour human dignity and the pedagogical essence of teaching.

The path forward is not only technically feasible, it is also ethically urgent and within our collective grasp.

'Somewhere, a child hesitates halfway through a thought, not seeking correction, but simply asking that her learning be honoured.

What happens next will not be decided by algorithms alone, but by the policies we craft and the ecosystems we choose to build around them.

Imagine classrooms where pause marks the beginning of understanding, where learner stories are held with care, and where paths are not just followed, but made, together.'[4]

4. Author's field notes, Malaysia, 2022-2025 (Product Leader, FrogAsia Company)

References

Alfredo, R., Echeverría, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D. and Martínez-Maldonado, R. 2024. Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, Vol. 6. Amsterdam, Elsevier. https://doi.org/10.1016/j.caeai.2024.100215 (Accessed 11 August 2025.)

Birhane, A. 2021. Algorithmic injustice: A relational ethics approach. *Patterns*, Vol. 2, No. 2. Cambridge, Cell Press. https://doi.org/10.1016/j.patter.2021.100205 (Accessed 11 August 2025.)

Bittencourt, I. I., Chalco, G., Santos, J., Fernandes, S., Silva, J., Batista, N., Hutz, C. and Isotani, S. 2023. Positive Artificial Intelligence in Education (P-AIED): A roadmap. *International Journal of Artificial Intelligence in Education*, Vol. 34. Leeds, International Artificial Intelligence in Education Society (IAIED). https://doi.org/10.1007/s40593-023-00357-y (Accessed 11 August 2025.)

Bryk, A. S. and Schneider, B. 2003. Trust in schools: A core resource for school reform. *Educational Leadership*, Vol. 60, No. 6. Alexandria, ASCD. https://www.ascd.org/el/articles/trust-in-schools-a-core-resource-for-school-reform (Accessed 11 August 2025.)

CIEB. 2024. *Inteligência Artificial na Educação Básica: Novas Aplicações e Tendências para o Futuro* [Artificial Intelligence in Basic Education: New Applications and Trends for the Future]. São Paulo, Centro de Inovação para a Educação Brasileira (CIEB). (In Portuguese.) https://cieb.net.br/wp-content/uploads/2024/06/Inteligencia-Artificial-na-Educacao-Basica_2024.pdf (Accessed 11 August 2025.)

Dream a Dream. 2023. *Direct-Intervention Impact Report 2022–23*. Bangalore, Dream a Dream, pp. 18–19. https://dreamadream.org/about/impact-2 (Accessed 11 August 2025.)

European Commission. 2024. *Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. Brussels, European Commission. https://artificialintelligenceact.eu (Accessed 11 August 2025.)

Fawns, T. 2022. An entangled pedagogy: Looking beyond the pedagogy–technology dichotomy. *Postdigital Science and Education*, Vol. 4. Berlin, Springer Nature, pp. 711–728. https://doi.org/10.1007/s42438-022-00302-7 (Accessed 11 August 2025.)

Federal Senate, Brazil. 2024. *Projeto de Lei n° 2338, de 2023: Dispõe sobre o uso da Inteligência Artificial* [Bill No. 2338, of 2023: Provides for the Use of Artificial Intelligence]. Brasilia, Federal Senate. (In Portuguese.) https://www25.senado.leg.br/web/atividade/materias/-/materia/157233 (Accessed 11 August 2025.)

Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford, Oxford University Press.

Government of Uruguay. 2024. *Estrategia Nacional de Inteligencia Artificial del Uruguay 2024–2030* [National Artificial Intelligence Strategy 2024–2030]. Montevideo, Government of Uruguay. (In Spanish.) https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/publicaciones/estrategia-nacional-inteligencia-artificial-del-uruguay-2024-2030 (Accessed 11 August 2025.)

Haraway, D. J. 2016. *Staying with the Trouble: Making Kin in the Chthulucene*. Durham, Duke University Press.

Ingold, T. 2007. *Lines: A Brief History*. London, Routledge.

Ministry of Education, India. 2020. *National Education Policy 2020*. New Delhi, Ministry of Education, India. https://www.education.gov.in/en/nep/about-nep (Accessed 11 August 2025.)

Ministry of Education and Culture, Finland. 2024. *Finnish Model for Leisure Activities Increases Schoolchildren's Participation Rate and Reaches Vulnerable Groups*. Helsinki, Ministry of Education and Culture, Finland. https://okm.fi/en/-/finnish-model-for-leisure-activities-increases-schoolchildren-s-participation-rate-and-reaches-vulnerable-groups (Accessed 11 August 2025.)

Noddings, N. 2005. Identifying and responding to needs in education. *Cambridge Journal of Education*, Vol. 35, No. 2. Cambridge, University of Cambridge, pp. 147–159.

NSW Department of Education. 2023. *Artificial Intelligence Assurance Framework*. Sydney, NSW Government. https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assessment-framework (Accessed 11 August 2025.)

ODITE. 2025. *Inteligencias Conectadas: Cómo la IA está Redefiniendo el Aprendizaje Personalizado* [Connected Intelligences: How AI is Redefining Personalized Learning]. Barcelona, Observatorio

de Innovación Educativa y Cultura Digital (ODITE). https://ciberespiral.org/informe-odite-2025 (Accessed 11 August 2025.)

OECD. 2023. *Equity and Inclusion in Education: Finding Strength through Diversity*. Paris, OECD Publishing. https://doi.org/10.1787/e9072e21-en (Accessed 11 August 2025.)

—. 2024. *Digital Education Outlook 2024: Shaping AI Use in Schools*. Paris, OECD Publishing. Unpublished (Forthcoming).

—. 2025. *A Snapshot of Equity and Inclusion in OECD Education Systems: Findings from the Strength through Diversity Policy Survey*. Paris, OECD Publishing. https://doi.org/10.1787/801dd29b-en (Accessed 11 August 2025.)

Rincón-Gallardo, S. 2023. *México: Redes de Tutorías* [Mexico: Tutoring Networks]. Washington, DC, Inter-American Development Bank (IADB). (In Spanish.) http://dx.doi.org/10.18235/0005222 (Accessed 11 August 2025.)

Simonsen, J., Karasti, H. and Hertzum, M. 2020. Infrastructuring and participatory design: Exploring infrastructural inversion as analytic, empirical and generative. *Computer Supported Cooperative Work*, Vol. 29. Berlin, Springer Nature, pp. 115–151.

Sloan, K. 2024. *AI in MATHia: It's More Than Right and Wrong Answers*. Pittsburgh, Carnegie Learning. https://www.carnegielearning.com/blog/mathia-ai (Accessed 11 August 2025.)

UNESCO. 2024. *AI Competency Framework for Teachers*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391104 (Accessed 11 August 2025.)

—. 2025. *Digital Learning Week 2025: Steering Technology for Education*. Paris, UNESCO. https://www.unesco.org/sites/default/files/medias/fichiers/2025/05/dlw-2025-cn-en.pdf (Accessed 11 August 2025.)

UNESCO and CIEB. 2024. *Guia IA para Educadores: Relatório de Co-Criação* [AI Guide for Educators: Co-creation Report]. Paris, UNESCO and São Paulo, CIEB. (In Portuguese.) Unpublished.

UNICEF. 2021. *Policy Guidance on AI for Children*. New York, UNICEF. https://www.unicef.org/innocenti/reports/policy-guidance-ai-children (Accessed 11 August 2025.)

Viberg, O., Cukurova, M., Feldman-Maggor, Y., Alexandron, G., Shirai, S., Kanemune, S., Wasson, B., Tømte, C., Spikol, D., Milrad, M., Coelho, R. and Kizilcec, R. F. 2024. What explains teachers' trust in AI in education across six countries? *International Journal of Artificial Intelligence in Education*, Vol. 34. Leeds, International Artificial Intelligence in Education Society (IAIED), pp. 1–29. https://doi.org/10.1007/s40593-024-00433-x (Accessed 11 August 2025.)

VVOB. 2021. *Teacher Motivation for Online Learning: Piloting a Microlearning Support System*. Kigali, VVOB. https://rwanda.vvob.org/news/teacher-motivation-online-learning-piloting-microlearning-support-system (Accessed 18 July 2025).

6. Ethical and governance imperatives for AI futures in education

Towards an ethics of care by design in AI in education

Kaśka Porayska-Pomsta and Isak Nti Asare

Introduction

The field of Artificial Intelligence in Education (AIED) has a relatively long history, which is grounded in the interplay between learning sciences and AI engineering (Woolf, 2008; McCalla, 2023). Its mission has traditionally been shaped by core principles such as responsible AI, learner-centred support and scalable, adaptive learning (Porayska-Pomsta, 2024). However, as AIED becomes more embedded in mainstream educational practice, it is increasingly being reframed through a commercial lens, often by actors detached from its research foundations. In this context, and echoing Paulo Freire's (1970) critique of the 'banking model' of education, education is cast as a market, learners as consumers, and educators as content deliverers. As AIED is increasingly shaped by commercial imperatives and generic AI models, there is a growing risk of displacing the foundational ideals of education as a public good and an instrument of freedom (Freire, 1970), since commercial interests tend to prioritize efficiency and scale, often at the expense of context, human connection and care.

A key part of the problem lies in the nature of the AI technologies that underpin most widely available AIED systems. These systems typically adopt a standard model of AI that assumes fixed, known objectives that are pursued through rational optimization, regardless of their (mis-)alignment with human intent (Russell, 2019). This can lead to bypassing the relational, dialogic and subjective dimensions of learning, long understood as essential for human development, resulting in interventions that may be ineffective at best, or harmful at worst (Treviranus, 2023). The combined pressures of marketization and the limitations of the standard model of AI highlight the acute need to reconsider how the role of education as a public good may be supported and reinforced.

In this think-piece we explore these questions through the prism of the ethics of care framework (Tronto, 1993), which foregrounds attentiveness, relational responsibility and responsiveness as a process of distributing care and responsibility among all stakeholders. This approach stands in contrast to entrusting the design of educational processes and tools only to selected actors who happen to be privileged by an economic or political system. The framework offers an important perspective on AIED as a process (rather than merely a product) that is embedded within shared and situated ethical educational practices.

Beyond critique, we explore what an emancipatory, care-centred AIED might look like in practice. We identify concrete approaches from AIED research through which AIED tools can be reimagined as co-created infrastructures that are responsive to diverse lived experiences, locally grounded, and shaped by learners and educators. While some care-centred approaches may appear aspirational in today's commercial and policy landscape,

they are far from speculative. Long-standing practices, such as open learner models (OLMs), participatory co-design and learning analytics, have over thirty years of research behind them, with demonstrated efficacy across diverse AIED contexts (Conati et al., 2018; Ramaswami et al., 2023). By foregrounding these practices as exemplars of designs that promote pedagogical quality and learner agency, while subverting the dominant model of AI and its assumptions, we aim to expand AIED's ethical and design horizons beyond solely personalization-driven tools, which are necessary but insufficient for care, to systems that are both adaptive to and modifiable by users in real time and in their immediate contexts.

What is ethics of care?

Ethics of care is a philosophical framework that recognizes human life as fundamentally shaped by dependency on others, environments and systems. Responsible action, therefore, must attend to the situated, concrete needs of individuals, especially those made invisible by dominant structures of power and knowledge (Tronto, 1993).

Joan Tronto's formulation identifies four moral and functional dimensions of care, each linked to a distinct phase of the process of care. Attentiveness requires recognizing the needs of others; responsibility involves accepting the obligation to respond; competence ensures that care is delivered effectively; and responsiveness centres the perspective of the cared-for and their reaction to the care received. These dimensions form a dynamic, recursive cycle that calls for ongoing reflection, adjustment and shared moral engagement.

The ethics of care presents a compelling alternative to abstract, tool-centric AI ethics. Rather than reducing moral questions to optimization or compliance, it foregrounds the lived realities of learners and educators, allowing them to assess AIED systems based on their relational sensitivity and ethical impact across the entire life cycle of their use. In what follows, we explore how this approach may help reframe the ethics of AIED.

Specifically, AI in education (as a field of research and as a type of technology) has the capacity both to reinforce and to disrupt Freire's 'banking model' of education. When designed around predetermined outcomes, constrained definitions of what is to be learned, rigid feedback structures (for example, correct/incorrect assessment) and performance-oriented learner surveillance, AIED systems can deepen passivity and disempowerment (Fanning, 2024). However, when approached through critical, learner-centered, constructionist pedagogies (Papert, 1980), such as those supported by OLMs, AI can be used to foster emancipatory learning and shared agency (between learners, systems and educators) through co-constructed personalization, made possible by systems that are both adaptive and adaptable (Du Boulay et al., 2023).

Thus, the liberatory or controlling role of AI in education is not fixed. It depends on the pedagogical design choices that precede and shape the technological models we adopt. Ethics of care shepherds us back to foundational best practices in education; it surfaces the deep tensions between the standard model of AI's ambition for full autonomy and the essential conditions for human learning and well-being.

In this context, the ethics of care offers more than abstract commitments to fairness or transparency. It provides a path towards grounded, value-driven design, deployment and use of AI. In the sections that follow, we draw on Tronto's four phases of care to explore the tensions between what AI technology, especially the standard model of AI, 'wants' (Kelly, 2010) and what the human mind needs to develop and thrive.

Principles of ethics of care by design in AIED

Caring about: Attentiveness

The first phase in Tronto's ethics of care corresponds to the moral quality of attentiveness, which prompts essential questions: Who is being seen? What needs are recognized? and Whose experiences are made visible?

These questions are particularly urgent in AIED, where systems are often developed around narrow, mainstream learner profiles that typically reflect Western, well-resourced contexts (Baker and Hawn, 2022). Consequently, the local and global needs and realities of underrepresented or marginalized learners are often overlooked or poorly served (Holmes and Porayska-Pomsta, 2023; Porayska-Pomsta, 2024).

This issue echoes a broader educational trend towards standardization in pedagogy, assessment and outcomes, which tends to favour the so-called 'norm', while de facto marginalizing those who do not conform. Those with alternative ways of knowing, for example, neurodivergent learners, are often cast in deficit terms and are seen as needing remediation rather than being recognized for their strengths (Valencia, 1997).

In contrast, attentiveness demands a shift from outcome-driven objectives to designs grounded in a context-sensitive understanding of learners: their identities, histories, environments and individual capacities. This shift highlights the need for continuity of informed insight and adaptive practices throughout people's lived experiences. When transposed onto AIED artefacts as vehicles of care, attentiveness calls for alignment and relevant knowledge across the entire AIED development life cycle, from initial conception to deployment and use, to ensure systems remain attuned to real-world needs and are effective at the front line. By the same token, this approach emphasizes the need for a functioning and sustainable ecosystem capable of supporting such continuity and relevance of insight throughout the AIED life cycle.

While current AI policies and ethical frameworks identify risks and offer general safeguards against them, they often lack the specificity and responsiveness required for true attentiveness. Caring involves more than rule-following and legal frameworks. It requires judgement, context awareness and moral imagination. In the context of AIED, 'caring about' means extending the kind of AI that is based solely on similarity patterns and rigid objective functions with approaches which are based on 'designing for difference'.

It is important to acknowledge that from their inception, research-driven AIED systems have been aimed to cater for attentiveness (McCalla, 2023). A core component of every classic Intelligent Learning Environment (ILE)[1] is their ability to detect and reason about learners' pertinent behaviours – the core of the adaptive

1. The term Intelligent Learning Environment is typically used to describe AI-driven systems that support human learning. The term has been increasingly adopted since the late 1990s to extend the original Intelligent Tutoring Systems approach by pedagogical approaches other than those focused on tutoring and drill-and-practice, for example, by exploratory learning environments, teachable agents or environments allowing students to self-direct in their help-seeking (Du Boulay et al., 2023).

capacities of AIED systems – enabling them to tailor real-time feedback and interaction accordingly, that is, to personalize. Similar detection mechanisms are also used in learning management systems and learning analytics-based dashboards to alert educators to critical learning incidents requiring their attention. However, the 'typical' patterns on which these systems are based carry a real risk of bias against any divergence, potentially marginalizing learners who do not conform. Moreover, teachers' routine reliance on these systems may lead to confirmation bias and/or skill atrophy, including a diminished ability to intuit the root causes of unusual behavioural patterns or to improvise their support (Lin et al., 2005; Baker and Hawn, 2022).

Used within a care-centred pedagogical framework that prioritizes attentiveness, such systems could be reframed more explicitly as supports that augment educators' curiosity and pedagogical responsiveness. However, to achieve this would require reframing learner data from serving performance diagnostics to supporting a better understanding of the learners' contexts, differences and needs. It would also likely involve a shift in the intended use of these tools from real-time decision augmentation towards supporting teacher training off-line and away from the front-line interventions, for instance through simulations of diverse learners and learning situations.

Caring for: Responsibility

The second phase of Tronto's care framework centres on responsibility, that is, the moral obligation to anticipate needs and ask: *Who will ensure care is delivered? Who is accountable when a need is known?* In the field of AIED, ethical concerns are increasingly acknowledged (Weber, 2020; Holmes et al., 2019), yet this recognition rarely translates into concrete moral or institutional action, particularly within commercial settings. A core barrier is that many actors involved in developing or deploying AIED are not structurally positioned or incentivized to assume ethical responsibility (McLennan et al., 2020), leading to a persistent accountability gap.

This gap is compounded by the misalignment between commercial incentives and the principles of care (Strümke et al., 2022). Commercial priorities, such as profit, speed and market share, often override commitments to pedagogical integrity or ethical depth. As a result, systems may be optimized for market logic rather than for the actual needs of learners and educators, posing a real risk of what Tronto describes as a 'breakdown in democratic care'.

This is not to dismiss the vital role of commercial actors in AIED innovation. Instead, this highlights the need to restructure how responsibility is defined, shared and enacted across stakeholders. To support real educational needs, responsibility must be distributed equitably across institutions and sectors, with input from relevant expertise at every stage of the AIED life cycle. Educators and learners must be positioned as central agents, not passive users, in the design, evaluation and use of these systems. Achieving this requires investment in collaborative, care-driven innovation, supported by incentive structures that prioritize public good over narrow commercial objectives.

At the technical level, personalization is often framed as AIED's signature feature, that is, a means of delivering care by tailoring responses to real-time assessments of learner performance. Yet, ethical responsibility for these personalized decisions remains ambiguously allocated among developers, institutions and users. Moreover, the accuracy of such assessments is limited by the systems' capacity to access and interpret relevant learner behaviours.

Even when systems generate rational inferences, these are shaped by the data and theoretical assumptions encoded in their models.

In contrast, when personalization is co-designed with educators and learners, and when it includes mechanisms such as OLMs that let users override, reshape or negotiate AI decisions (Bull and Kay, 2016), moral responsibility can be more equitably and transparently shared. OLMs exemplify one of several forms of collaborative approach developed and tested in AIED research. They have been shown to promote more accountable design and to foster users' deep engagement with both learning processes and pedagogical practice (Bull and Kay, 2016; Conati et al., 2018).

Care-giving: Direct action, competence

Competence refers to the ability to deliver care effectively and appropriately. Good care requires skill, contextual awareness and domain-specific knowledge – without these, even well-intentioned actions can cause harm or fail to meet actual needs.

In AIED, this raises an often-overlooked question: Do all stakeholders in AIED possess the skills, resources, and institutional support necessary to participate meaningfully in its ethical governance? Distributing responsibility without investment in stakeholders' capacity to act effectively is insufficient. Technical literacy, contextual understanding and decision-making powers are unevenly distributed in the AIED ecosystem (Miao et al., 2021). Yet, most existing ethical AI frameworks set aside the question of whether all actors can actually engage on an equal footing and what specific pre-conditions may enable meaningful participation in this context.

Research in AIED (as distinct from commercial implementations) offers promising evidence and prototypes. Notably, it highlights the pedagogical and ethical value of alternative models to the standard AI paradigm, such as OLMs already mentioned (Bull and Kay, 2016). These systems allow users to engage with and refine AI models in real time, increasing relevance, accuracy and transparency thereof (Conati et al., 2018).

This same research highlights that co-creation and collaboration are not just ethical imperatives, but that they are also competencies that can be learned given flexible system designs. When users can scrutinize, negotiate or edit AI models during use, they build AI literacy and metacognitive competencies through exerting their agency and self-reflection, thus also developing an intimate relationship with their learning, thinking and practices (Kay et al., 2023).

In this light, care-giving may need to be reframed as creating environments and opportunities that enable stakeholders to direct both the care they receive (via AIED design and policy structures) and their own self-care through meaningful ownership of their educational experiences.

Competence in AIED systems is increasingly demonstrated through features that give users control over their learning, via appropriate help-seeking (Aleven et al., 2016) or encouraging exploratory behaviour (Mavrikis et al., 2022). While these features reflect a form of technical competence, their ethical competence depends on how well they are embedded within culturally responsive pedagogy. A genuinely care-centred system goes beyond functionality; it ensures alignment with learners' diverse contexts through co-design, and continuous evaluation of the relevance and effectiveness of the interventions delivered.

Care receiving: Responsiveness

Responsiveness focuses on how those receiving care experience it, asking whether their needs have truly been met. It distinguishes ethical care from merely well-intentioned action by requiring ongoing attentiveness, feedback and adaptation.

In AIED, responsiveness is perhaps the most underdeveloped dimension of ethics. While existing frameworks often emphasize fairness, transparency or accountability during design and deployment (Weber, 2020; Holmes et al., 2019), they rarely consider whether systems remain ethical in practice, or how they evolve in response to users over time. Where feedback mechanisms exist, they are often superficial, one-off or entirely absent, particularly regarding ethical performance.

One barrier here is the reliance on abstract, universal AI ethics principles, such as fairness or trust, whose meaning changes across educational, cultural and social contexts. From a care ethics perspective, this highlights the need for ethical reflection to be situated in the lived realities of learners and educators, rather than anchored in generalized ideals.

A further issue is the tendency to treat AIED systems as fixed artefacts that are deployed, evaluated and left unchanged, rather than as 'living infrastructures' that should grow in dialogue with those they serve. Genuine responsiveness requires sustainable, embedded feedback mechanisms that empower users to surface concerns, provide insights and help shape system updates.

OLMs yet again offer a promising model of embedded responsiveness. By allowing learners to view, interpret and adjust how the system represents them, they support continuous, contextual feedback loops. Unlike top-down or isolated evaluations, interactions with OLMs are situated within everyday learning, allowing for real-time negotiation of meaning and intent. Such dynamic, user-centred mechanisms mark a shift from uni-directional interventions (and their evaluations), where AI acts upon the learner, to a more reciprocal relationship, where systems learn from and can be modified by those they aim to support.

Although responsiveness remains the least developed of Tronto's phases within AIED, emerging practices like OLMs point towards a future in which learners and educators are not just recipients of automated care, but active participants in shaping how care is delivered and sustained.

Discussion and conclusions

This think piece has explored the ethics of care in AIED, using Tronto's framework to surface often-overlooked dimensions of ethical practice. Three interrelated areas emerge as priorities for building and sustaining responsible AI in and for education, each requiring coordinated action across the polycentric AIED ecosystem.

First, policy must shift from abstract ethical principles to context-sensitive, actionable definitions. Current ethical AI frameworks, whether in strategy documents, guidelines or regulations, tend to generalize concepts such as 'fairness' or 'accountability' without specifying what they mean in diverse educational contexts. A care-informed approach urges policy-makers to move beyond static values and instead support ethics that are co-produced, situated and responsive. For example, fairness might be meaningfully enacted by requiring that AI systems be co-designed with teachers and students, equipped with mechanisms for local oversight, and evaluated for their impact on historically marginalized groups. Crucially, this shift also demands reconsideration of the technical design of AI systems. For instance, systems should enable forms of user control and

negotiation, such as those offered by OLMs, that embed ethics into everyday practice. While the importance of contextualized ethics has been noted (Holmes and Porayska-Pomsta, 2023), the ethics of care highlights ongoing gaps in how such principles are enacted. It exposes the limitations of generalized ethical constructs and highlights the need for ethical frameworks that reflect the relational and situated realities of those most impacted by AIED: learners and educators.

Second, funding models must prioritize sustainable, ecosystemic approaches across the full AIED life cycle. Rather than supporting isolated pilot projects, funders (public and private) should enable long-term relationships between developers, educators and communities. This includes investments in ongoing teacher professional development, iterative co-design grounded in real-world use, and accountability structures for identifying and responding to unintended harms over time. Although participatory design has a presence in educational technology, it is often tokenistic in AIED, ending at the deployment or prototype stage. Research on OLMs, seen through the lens of care ethics, offers an alternative: one where all stakeholders are viewed as integral to sustainable and responsive AIED. This model departs from dominant AI paradigms focused on autonomous prediction and prescription, and instead supports systems that collaborate with users, negotiating predictions and co-shaping interventions. Such models also offer a compelling way to rethink governance: while regulation remains essential, it is often too slow to keep pace with innovation. Open and cooperative models provide more agile and responsive infrastructures for distributing ethical responsibility, resonating with Elinor Ostrom's (1990) vision of commons-based governance.

Third, new competencies are required to sustain these care-based, participatory ecosystems. Frameworks like UNESCO's (Miao and Cukurova, 2024) emphasize collaborative partnerships and recognise that the skills needed for ethical, participatory engagement with AI are learnable. Care ethics positions these competencies not as optional but as foundational. As responsibility and decision-making become more distributed, and as AI systems become increasingly adaptive, educators, learners and other stakeholders must be equipped not only to collaborate with each other but also to interact meaningfully with AI systems themselves. This includes understanding how systems work, how to question or modify them, and how to participate in their evolution over time.

Ethical reflection in AIED must begin not with the technology itself, but with the people, relationships and contexts that education is meant to support. By virtue of being centred on attentiveness, responsibility, competence and responsiveness, Tronto's ethics of care offers a relational alternative to compliance-based models by emphasizing lived experience and shared accountability. Rather than retrofitting ethics after deployment, we advocate for care-informed design from the outset, using context-aware, participatory and technical approaches that embed ethical reflection into AI systems and that allow adjustment to individual and contextual needs. Open AI models, such as OLMs, exemplify how care and responsibility can be built into the fabric of AIED, enabling more adaptive, transparent and inclusive systems. Viewed through this lens, AIED can be reimagined as a 'commons' that are co-created, co-governed and sustained through collaborative care.

References

Aleven, V., Roll, I., McLaren, B. M., Koedinger, K. R. 2016. Help helps, but only so much: Research on help seeking with intelligent tutoring systems. *International Journal of Artificial Intelligence in Education*, Vol. 26. Leeds, International Artificial Intelligence in Education Society (IAIED), pp. 205–223.

Baker, R S. and Hawn, A. 2022. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, Vol. 32. Leeds, International Artificial Intelligence in Education Society (IAIED), pp. 1052–1092. https://doi.org/10.1007/s40593-021-00285-9 (Accessed 12 August 2025.)

Bull, S., and Kay, J. 2016. SMILI☺: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education*, Vol. 26, No. 1. Leeds, International Artificial Intelligence in Education Society (IAIED), pp. 293-331. https://doi.org/10.1007/s40593-015-0090-8 (Accessed 12 August 2025.)

Conati, C., Porayska-Pomsta, K. and Mavrikis, M. 2018. *AI in education needs interpretable machine learning: Lessons from Open Learner Modelling*. Ithaca, arXiv. https://doi.org/10.48550/arXiv.1807.00154 (Accessed 12 August 2025.)

Du Boulay, B., Mitrovic, A. and Yacef K. (eds). 2023. *Handbook of Artificial Intelligence in Education*. Cheltenham, Edward Elgar.

Fanning, S. 2024. Parallel minds: A thematic literature review of Artificial Intelligence in Education. MA dissertation, Department of Kinesiology and Educational Psychology, Washington State University. https://doi.org/10.7273/000007186 (Accessed 12 August 2025.)

Freire, P. 1970. *Pedagogy of the Oppressed*. New York, Seabury Press.

Holmes, W., Bektik, D., Woolf, B. and Luckin, R. 2019. Ethics in AIED: Who cares? *20th International Conference on Artificial Intelligence in Education*. Milton Keynes, the Open University. https://oro.open.ac.uk/60361 (Accessed 12 August 2025.)

Holmes, W. and Porayska-Pomsta, K. 2023. *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*. London, Routledge.

Kay, J., Kummerfeld, B., Conati, C., Porayska-Pomsta, K. and Holstein, K. 2023. Scrutable AIED. B. du Boulay, A. Mitrovic and K. Yacef (eds), *Handbook of Artificial Intelligence in Education*. Cheltenham, Edward Elgar.

Kelly, K. 2010. *What Technology Wants*. New York, Penguin Books.

Lin, X., Schwartz, D. L. and Hatano, G. 2005. Toward teachers' adaptive metacognition. *Educational Psychologist*, Vol. 40, No. 4. Philadelphia, Taylor & Francis, pp. 245–255.

Mavrikis, M., Rummel, N., Wiedmann, M., Loibl, K. and Holmes, W. 2022. Combining exploratory learning with structured practice educational technologies to foster both conceptual and procedural fractions knowledge. *Educational Technology Research and Development*, Vol. 70. Indianapolis, Association for Educational Communications and Technology, pp. 691–712. https://doi.org/10.1007/s11423-022-10104-0 (Accessed 12 August 2025.)

McCalla, G. 2023. The history of artificial intelligence in education: The first quarter century. B. du Boulay, A. Mitrovic and K. Yacef (eds), *Handbook of Artificial Intelligence in Education*. Cheltenham, Edward Elgar.

McLennan, S., Fiske, A., Celi, L. A., Müller, R., Harder, J., Ritt, K., Haddadin, S. and Buyx, A. 2020. An embedded ethics approach for AI development. *Nature Machine Intelligence*, Vol. 2, No. 9. Berlin, Springer Nature, pp. 488–490. https://doi.org/10.1038/s42256-020-0214-1 (Accessed 12 August 2025.)

Miao, F. and Cukurova, M. 2024. *AI competency framework for teachers*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391104 (Accessed 12 August 2025.)

Miao, F., Holmes, W., Huang, R. and Zhang, H. 2021. *AI in education: Guidance for policy-makers*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709 (Accessed 12 August 2025.)

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, Cambridge University Press.

Papert, S. 1980. *Mindstorm: Children, Computers, and Powerful Ideas*. New York, Basic Books.

Pinkwart, N. 2016. Another 25 years of AIED? Challenges and opportunities for intelligent educational technologies of the future. *International Journal of Artificial Intelligence in*

Education, Vol. 26. Leeds, International Artificial Intelligence in Education Society (IAIED), pp. 771–783. https://doi.org/10.1007/s40593-016-0099-7 (Accessed 12 August 2025.)

Porayska-Pomsta, K. 2024. A manifesto for a pro-actively responsible AI in education. *International Journal of AI in Education*, Vol. 34. Leeds, International Artificial Intelligence in Education Society (IAIED), pp. 73–83. https://doi.org/10.1007/s40593-023-00346-1 (Accessed 12 August 2025.)

Porayska-Pomsta, K., Holmes, W. and Nemorin, S. 2023. The ethics of AI in education. B. du Boulay, A. Mitrovic and K. Yacef (eds), *Handbook of Artificial Intelligence in Education*. Cheltenham, Edward Elgar.

Ramaswami, G., Susnjak, T. and Mathrani, A. 2023. Effectiveness of learning analytics dashboards for increasing student engagement levels. *Journal of Learning Analytics*, Vol. 10, No. 30. Beaumont, Journal of Learning Analytics. https://doi.org/10.18608/jla.2023.7935 (Accessed 12 August 2025.)

Russell, S. J. 2019. *Human Compatible: Artificial Intelligence and the Problem of Control*. New York, Viking.

Strümke, I., Slavkovik, M. and Madai, V. I. 2022. The social dilemma in artificial intelligence development and why we have to solve it. *AI and Ethics*, Vol. 2. Berlin, Springer Nature, pp. 655–665. https://doi.org/10.1007/s43681-021-00120-w (Accessed 12 August 2025.)

Treviranus, J. 2023. Learning to learn differently. W. Holmes and K. Porayska-Pomsta (eds), *The Ethics of Artificial Intelligence in Education*. New York, Routledge.

Tronto, J. 1993. *Moral Boundaries: A Political Argument for an Ethic of Care*. New York, Routledge.

Valencia, R. R. (ed.). 1997. *The Evolution of Deficit Thinking: Educational Thought and Practice*, The Stanford Series on Education and Public Policy. London, Routledge.

Weber, A. 2020. Ethics concerns in artificial intelligence use in education. *INTED2020 Proceedings*. Valencia, International Academy of Technology, Education and Development (IATED), pp. 4539–4544.

Woolf, B. 2008. *Building Intelligent Tutoring Systems: Student-centred strategies for Revolutionizing E-learning*. San Francisco, Morgan Kauffman.

Artificial intelligence and governing education: Rethinking democratic action, resistance and participation

Kalervo N. Gulson and Sam Sellar

> 'We have been assimilated, all too willingly, and there is probably no going back.'
>
> Paul Edwards (2018)

In 2022, we published a book, *Algorithms of Education: How Datafication and Artificial Intelligence Shape Policy* (Gulson et al., 2022), in which we speculated on how AI would shape the future governance of education. Now, a short three years later, we find that many of our speculations have come to pass. We are getting a clear sense of what kinds of educational worlds, and what kinds of governing truths, are being created through AI. Alongside these new worlds and truths, substantial new challenges to democratic action and resistance are also being posed.

Our book, written prior to the release of ChatGPT and the ubiquity of foundational models and generative AI (GenAI), outlined what we thought would become an emerging synthesis of humans and machines in education policy and governance. We termed this emerging phenomenon 'synthetic governance' to describe how a technology like AI enters a mutually constitutive relationship with social actions and cultural values (Mackenzie, 2022). We proposed that synthetic governance is not human or machine governance, but human and machine governance, arising from 'conjunctive syntheses' (Deleuze and Guattari, 1983) that bring together and integrate data-driven human and computational rationalities. Our contention was that synthetic governance is an amalgamation of human classifications, rationalities, values and calculative practices with new non-human political rationalities embedded in algorithms, data infrastructures and AI, and it is changing how we think and govern.

At the time of publication, our interest was in how AI was being used in some early-stage EdTech products, as part of digital infrastructures, and in the work of data scientists employed by education systems. However, AI was not widely used at all in education governance, and our empirical examples of synthetic governance were thin on empirical detail. We speculated that we would see further intensification of this form of governance. However, we certainly did not predict just how quickly some things would change and we now have a much better sense of the challenges to democratic action and resistance that AI brings to education governance.

First, the relationship between AI and expertise is changing. While we had studied how AI was being used in policy and governance, we found that this required specialist expertise and intra-organizational cultural shifts in the types of expertise that were required and valued (for example, data science). We observed government education departments employing data scientists, and hence importing knowledge about using AI, adding an additional dimension to historically embedded statistical management of education and the professional judgement of educators (Sellar and Gulson, 2021). We anticipated that this expertise was going to be key to understanding how AI shaped future practices of governing. What we have seen with the advent of foundation models and GenAI, like ChatGPT, is a dispersal of expertise about the use of AI in governance through the extensive integration of AI into all areas from teaching and learning to administration. For

example, GenAI is being incorporated into management platforms, which in education includes bundles of products and services relating to student information systems and learning management systems. When GenAI is used here it includes activities such as drafting emails, reports and public relations materials. Another still understudied area is how GenAI is being used as a recommender system, such as providing advice for school leadership teams or education systems. Importantly, the accessibility of GenAI (that is, not requiring users to have any special expertise in data science or programming) allows a combination of evidence and natural language prompts to be used in decision-making. This includes instances where student data can be used to prompt ChatGPT to assist students with selecting elective courses (Chiu, 2024). This development forces us to rethink our earlier notion that new kinds of expertise are needed in education to implement AI effectively.

Second, AI now exhibits some interesting similarities to cognitive dynamics associated with human thinking. For example, GenAI produces text and images that have no reference point in human knowledge, vernacularly described as 'hallucinations', with evidence emerging of the implications of these processes for the public sector (Cantens, 2025), including creating false references that can undermine trust in decision-making in education. We saw such an instance in Alaska, when an official from the state education department used GenAI to create a policy document on banning the use of mobile phones. It transpired that many of the citations in the document either did not exist or were not relevant to the policy context (Fisher Phillips, 2024). In this case, while GenAI created new text in the form of citations, it did not create legitimate new knowledge. The hallucinations produced by the system are offered as valid outputs given the parameters of the system. Here, the use of GenAI must be understood from the perspective that there 'are new digitized power/knowledges, new truth regimes constructed by the impeccable logic of algorithms, a logic that has become abstractized and riven from human realities' (Henman, 2020, p. 25). This logic introduces a 'radical otherness' that is created when 'machine-built systems use machine logic, not human logic' (Edwards, 2018, p.23).

Third, and connected to the point above, is the role technology companies play to enable integration of AI across public services like education. Access to GenAI is provided in various ways from subscription models to freely available versions, and this includes systems in which a commercial AI model is procured, and the company modifies it on behalf of an educational user (Jones, 2023). Over the past decade, especially after the Covid pandemic, there have been growing concerns about the role of technology companies in education (Williamson and Hogan, 2020). These concerns are exacerbated by new forms of AI and we must consider technology companies as no longer being just product and service providers; these companies are now part of the infrastructure of governing education itself through inter-organizational relationships that fundamentally shape the problems and solutions that drive the design and use of EdTech products and services. What is becoming evident is that AI use in public service platforms, such as Amazon Web Services, exacerbates existing dependencies on single proprietary sources of infrastructure and information (Plantin et al., 2016).

We suggest that all three of these dynamics point to the ways the introduction of AI, and the changes to governance and education, demand that we rethink democratic action, resistance and participation in education. In what follows, we outline some possible policy responses.

First, better regulation of AI is a necessary yet insufficient response, and future regulatory developments must consider the specificities

of education. Current regulation largely presuppose that AI can be used in education towards socially progressive ends by utilizing existing policy instruments and governance mechanisms. Yet, aside from some examples such as the Council of Europe's work on regulating AI, there is an absence of work focusing on the values of education. This is a space where we need to think more broadly about regulation as producing new values and not just managing risk. That is, new policy and regulation must move beyond setting boundaries on how existing technologies can be used and instead create conditions for collaborative and creative engagement with new technological developments. This will involve expanding the network of those included in deciding what kinds of AI and EdTech are desirable and possible. As Arathi Sriprakash and colleagues (2025) propose, it is 'time for engaged, collaborative research in partnership with activists, educators and academics who are already challenging extractive imaginaries of sociodigital futures in education and opening up alternatives' (p. 563).

Second, democratic resistance is being impeded by the embeddedness of technology companies in governance infrastructures. Decisions are made by technology companies with little to no oversight from user organisations like education departments and schools. For example, in June 2025 Google embedded thirty new AI tools into Google Classroom without explanation of purpose, or asking if these tools were desired by schools and educators (Dougherty, 2025). The deeper technology companies reach into the management of everyday life, public and private, the more challenging it becomes to shape technology for democratic purposes. Indeed, we may require policy action that, somewhat paradoxically, takes the form of what Blair Attard-Frost (2023) calls 'counter-governance'. That is, a strategy that refuses and reorients AI governance through a focus on power and organization rather than merely more democratic and ethical designs and uses of AI. Counter-governance of AI involves movements that 'primarily direct organized, collective, community-led opposition against the underlying political and economic systems that are constitutive of "AI," rather than intervening in the design of the technology itself' (Attard-Frost, 2023). Examples of counter-governance in practice include community-led audits of AI acts in Canada, including public dissemination in the form of reports, roundtables, media work and freedom of information (Attard-Frost, 2023).

Third, we need policy-making to be informed by new theories of democracy, or at the least we need to rethink our current participatory and deliberative democratic forms, to deal with the new dynamics created by machine-human syntheses. This requires thinking about democracy in terms of not only competing values but also radical otherness and incommensurate logics that are meshed in human-focused applications. This will involve experimenting with how to engage with non-human actors like AI in governing education. We do have conceptual precedents here, such as work in the field of law that has considered the role of non-human actors such as rivers. One political option is to fully embrace possibilities for co-creating socio-technical simulations of more democratic approaches to governing. For some, these opportunities will be uncomfortable because they unsettle human thinking and values. But the premise of these approaches would be that the capabilities of AI can combine with human analytical skills to create new kinds of foresight for governing AI. However, as is already becoming evident in the use of machine learning to create synthetic datasets, using AI always carries with it the risk of further entrenching inequities while not generating any useful cognitive or administrative advantages (Lee et al., 2025). As such, any new forms of democratic thinking

will need to forge new methods of engaging with AI to produce different possibilities for technological change.

References

Attard-Frost, B. 2023. AI countergovernance. *Midnight Sun*. https://www.midnightsunmag.ca/ai-countergovernance (Accessed 13 August 2025.)

Cantens, T. 2025. How will the state think with ChatGPT? The challenges of generative artificial intelligence for public administrations. *AI & Society*, Vol. 40, No. 1. Berlin, Springer Nature, pp. 133–144.

Chiu, T. K. F. 2024. The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, Vol. 32, No. 10. London, Taylor & Francis, pp. 6187–6203. https://doi.org/10.1080/10494820.2023.2253861 (Accessed 13 August 2025.)

Deleuze, G. and Guattari, F. 1983. *Anti-Oedipus*. Minneapolis, University of Minnesota Press.

Dougherty, J. 2025. *Default on; quality off: Google Classroom's new AI tools*. Oakland, Beta Classroom. https://betaclassroom.wordpress.com/2025/07/05/default-on-quality-off-google-classrooms-new-ai-tools (Accessed 13 August 2025.)

Edwards, P. N. 2018. We have been assimilated: Some principles for thinking about algorithmic systems. U. Schultze, M. Aanestad, M. Mähring, C. Østerlund and K. Riemer (eds), *Living with Monsters? Social Implications of Algorithmic Phenomena, Hybrid Agency, and the Performativity of Technology. IS&O 2018. IFIP Advances in Information and Communication Technology*, Vol. 543. Cham, Springer. https://dx.doi.org/10.1007/978-3-030-04091-8_3 (Accessed 13 August 2025.)

Fisher Phillips. 2024. *Education officials learn dangers of AI after citing false studies in official document: 5 steps for school administrators to avoid similar fate*. Fisher Phillips. https://www.fisherphillips.com/en/news-insights/education-officials-learn-dangers-of-ai.html (Accessed 13 August 2025.)

Gulson, K.N., Sellar, S. and Webb, P.T. 2022. *Algorithms of Education: How Datafication and Artificial Intelligence Shape Policy*. Minneapolis, University of Minnesota Press.

Henman, P. 2020. Governing by algorithms and algorithmic governmentality: Towards machinic judgement. M. Schiulenberg and R. Peeters (eds), *The Algorithmic Society*. London, Routledge, pp. 19–34.

Jones, E. 2023. *Foundation Models in the Public Sector*. London, Ada Lovelace Institute.

Lee, F., Hajisharif, S. and Johnson, E. 2025. The ontological politics of synthetic data: Normalities, outliers, and intersectional hallucinations. *Big Data & Society*, Vol. 12, No. 2. Thousand Oaks, SAGE Publications. https://doi.org/10.1177/20539517251318289 (Accessed 13 August 2025.)

Mackenzie, A. 2002. *Transductions: Bodies and Machines at Speed*. London, Continuum.

Plantin, J. C., Lagoze, C., Edwards, P. N. and Sandvig, C. 2016. Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, Vol. 20, No. 1. Thousand Oaks, SAGE Publications, pp. 293–310.

Sellar, S. and Gulson, K. N. 2021. Becoming information centric: The emergence of new cognitive infrastructures in education policy. *Journal of Education Policy,* Vol. 36, No. 3. London, Taylor & Francis, pp. 309–326. https://doi.org/10.1080/02680939.2019.1678766 (Accessed 13 August 2025.)

Sriprakash, A., Williamson, B., Facer, K., Pykett, J. and Valladares Celis, C. 2025. Sociodigital futures of education: Reparations, sovereignty, care, and democratisation. *Oxford Review of Education*, Vol. 51, No. 4. London, Taylor & Francis, pp. 561–578. https://doi.org/10.1080/03054985.2024.2348459 (Accessed 13 August 2025.)

Williamson, B. and Hogan, A. 2020. *Commercialisation and Privatisation in/of Education in the Context of Covid-19*. Brussels, Education International.

7. Confronting coded inequalities in education

Ensuring inclusive, contextualized AI in education: Considerations towards a roadmap

Vukosi Marivate, Nombuyiselo Caroline Zondi and Baphumelele Masikisiki

Introduction: The shifting demands on teachers

Oh, what we ask of our teachers today. They are expected to nurture the very young, inspire restless teenagers, and serve both as subject experts and guides through a complex world. In societies still grappling with historical and ongoing inequalities, they must confront systemic failures while helping students imagine better futures. And now, layered onto these demands, is the quiet but growing presence of AI in the classroom. Smartboards sit largely unused (Van der Walt, 2024); AI tools arrive faster than teachers or learners can fully grasp their implications. Students explore these tools with curiosity, while teachers often navigating unfamiliar digital terrain must make sense of AI's promises, pitfalls and unknowns. How should learners responsibly engage with AI? How should teachers balance opportunity with caution? This think piece explores these dilemmas, reflecting on the challenges and possibilities within South Africa's education system and drawing connections to broader global concerns. We propose key considerations rather than a definitive roadmap for equitable and context-sensitive AI integration.

Understanding AI in education

AI refers to computer systems that can perform tasks that usually require human intelligence, such as recognizing patterns, making decisions or solving problems. A specific type of AI, called large language models (LLMs), are trained on massive volumes of text to process and generate human-like language. They can assist with essay grading, personalized tutoring or administrative tasks. These tools offer new ways to support learning but also present significant challenges (Abumusab, 2024; Wang et al., 2024). Many African communities have been left out of how these technologies are designed and developed. This exclusion leads to cultural gaps, gender gaps and ethical concerns, which are aggravated by practical barriers, including poor infrastructure and limited internet access. We believe AI can contribute to more inclusive education systems but only if local knowledge is integrated, cross-sector expertise is mobilized, and enabling conditions like electricity and connectivity are established.

AI in assessment: Promise and precautions

One of the most immediate uses of AI in education is in aiding assessment. Essay grading remains labour-intensive, inconsistent and often subjective. Research at a South African university shows how AI-driven scoring models can help (Twabu and Nakene-Mginqui, 2024). In a comparative study with physiology instructors, AI-generated scores and accompanying textual justifications led nearly half of the instructors to revise their original grades (Masikisiki et al., 2024). This suggests that, if well-designed, AI can improve grading consistency. It can also accelerate feedback loops, allowing students to revise and improve their work more quickly.

However, AI cannot solve assessment challenges alone. When trained on superficial features like essay length, models may reward verbosity over quality. To prevent this, researchers develop rubrics with educators and train AI systems to assess coherence, reasoning and content depth. 'Penetration testing' further checks model reliability by feeding low-quality essays to confirm that longer responses are not unfairly rewarded. Still, human-in-the-loop systems are vital. Teachers regularly review AI-marked essays to ensure alignment with educational values and disciplinary expectations.

Beyond grading: Supporting learning in marginalized contexts

AI tutoring tools can support personalized learning by identifying misconceptions, suggesting resources or adjusting lesson difficulty. This can be valuable in under-resourced schools where human support is limited. But if introduced without concurrent investments in digital literacy, devices and connectivity, such tools may deepen educational inequities. AI implementation must include teacher training and equitable technology access. This is especially urgent in schools that experience frequent power outages or lack adequate hardware.

Multilingual classrooms and the limits of text-based AI

In South African classrooms, learning is multilingual, cultural and interactive. Teachers draw on songs, play and oral traditions to nurture curiosity. Text-centric AI tools risk ignoring this richness. A Zulu-speaking learner may convey meaning through poetry or gesture forms that standard AI systems cannot interpret. Ignoring these expressions marginalizes non-Western ways of knowing and reduces learning to a narrow, standardized metric.

To avoid this, AI must evolve beyond mere translation. Models should be designed to understand multiple modes of communication, especially in indigenous contexts. This requires co-design with teachers, language experts and community members who understand local education practices. Participatory workshops where learners, storytellers and educators share oral histories and cultural practices can inform training data for AI tools rooted in community knowledge. Further, we need to understand deeply embedded biases in LLMs. Recent work focused on a context in India highlights gender biases that result from different stages of building LLMs (Aneja et al., 2025).

Language, vocabulary and epistemic exclusion

AI systems trained primarily on English fail to support the many African learners who use indigenous languages daily. Vocabulary development is foundational to literacy, yet African languages often lack age-appropriate word lists, corpora or digital tools. This is not merely a technical gap, it is an epistemic crisis. When languages are undocumented, they are excluded from AI training. This reinforces inequality: well-documented languages become 'AI-ready', while others remain invisible.

Moreover, the structure of African languages is often agglutinative and tonal, and this confuses English-trained tokenizers. A single word in isiXhosa may carry as much meaning as an entire sentence in English. Without linguistically appropriate tools, AI systems misinterpret or invalidate learners' responses, marking correct expressions as wrong and reinforcing deficit views of learners.

To address this, communities must build indigenous language corpora, design linguistically appropriate tokenizers, and ensure collaboration between AI developers, teachers and linguists. Projects such as

Masakhane[1] are already taking these steps by developing open datasets and models for African languages.

Grounded innovation: Local projects leading change

In South Africa, the Data Science for Social Impact (DSFSI) Lab at the University of Pretoria offers an inspiring example of community-rooted innovation. DSFSI collaborates with linguists, educators and technologists to build localised, culturally relevant content and reading tools. An example of one of the projects is that of automated grading for medical students, developed in tandem with professors and which takes into account local contexts and needs.

Transparency, ethics and governance

Many AI systems function as black boxes, making decisions without transparency. In education, this is unacceptable. Teachers must understand how AI arrives at conclusions, particularly in assessment. Interpretability layers that explain which phrases influenced a grade allow educators to review and override decisions. Human oversight is essential not just for trust, but to prevent harm.

We must also establish ethical frameworks for data use. Historically, education data collection in Africa has been extractive. Community data trusts with locally governed repositories can offer a more equitable model. Universities and legal clinics can support these efforts by developing fair-use policies and ensuring data is not repurposed without informed consent.

Global relevance through local insight

While rooted in the South African experience, the lessons shared here resonate globally. Indigenous, multilingual and underserved communities worldwide face similar exclusion from AI development. By emphasizing linguistic diversity, cultural expression and equitable access, we offer principles that can inform AI integration in any context striving for justice.

Key considerations for equitable AI in education

1. Centre local voices: Empower educators, learners and communities to co-create AI tools that reflect their contexts.
2. Maintain human oversight: Design interpretable systems and retain teacher judgement in AI-assisted decisions.
3. Invest in infrastructure and collaboration: Support offline-compatible tools, transdisciplinary partnerships and grassroots-led innovation.

Conclusion: Optimism with accountability

We remain cautiously optimistic about AI's potential in education but only if its development centres equity, humility and local knowledge. When teachers shape the tools, when learners see their languages reflected, and when communities govern their own data, AI becomes a tool for empowerment rather than erasure. The goal is not to build perfect systems, but to build participatory ones. The future of AI in education must echo the plurality of its learners, their tongues, their stories and their right to be heard.

Acknowledgements

We are thankful for the editing assistance of Angel Maduke.

1. See https://www.masakhane.io

References

Abumusab, S. 2024. *LLMs in education: Forged from years of educational challenges*. Newark, American Philosophical Association (APA). https://blog.apaonline.org/2024/09/26/llms-in-education-forged-from-years-of-educational-challenges (Accessed 13 August 2025.)

Aneja, U., Gupta, A. and Vashistha, A. 2025. Beyond semantics: Examining gender bias in LLMs deployed within low-resource contexts in India. *FAccT '25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. New York, Association for Computing Machinery. https://doi.org/10.1145/3715275.3732180 (Accessed 13 August 2025.)

Masikisiki, B., Marivate, V. and Hlophe, Y. 2024. Investigating the efficacy of large language models in reflective assessment methods through chain of thought prompting. *AfriCHI '23: Proceedings of the 4th African Human Computer Interaction Conference*. New York, Association for Computing Machinery. https://doi.org/10.1145/3628096.3628747 (Accessed 13 August 2025.)

Twabu, K. and Nakene-Mginqu, M. 2024. Developing a design thinking artificial intelligence driven auto-marking/grading system for assessments to reduce the workload of lecturers at a higher learning institution in South Africa. *Frontiers in Education*, Vol. 9. Lausanne, Frontiers Media. https://doi.org/10.3389/feduc.2024.1512569 (Accessed 13 August 2025.)

Van der Walt, I-M. 2024. Smart board integration: Mixed results from N$9 million investment. *Windhoek Express*. https://www.we.com.na/focus-tel-and-tech/smart-board-integration-mixed-results-from-n9-million-investment2024-10-29135469 (Accessed 13 August 2025.)

Wang, S., Xu, T., Li, H., Zhang, C., Liang, J., Tang, J., Yu, P.S. and Wen, Q. *Large language models for education: A survey and outlook*. Ithaca, arXiv. https://doi.org/10.48550/arXiv.2403.18105 (Accessed 13 August 2025.)

From compliance to creativity: Reimagining AI in young women's learning

Kiran Bhatia and Payal Arora

This think piece advocates for a paradigmatic shift from paternalistic, risk-centric approaches to AI in education to feminist, culturally grounded frameworks that prioritize creativity, autonomy and joy in the learning experiences of young women. Grounded in qualitative research and participatory workshops in India and Brazil, this article elucidates the ways in which young women tactically deploy AI technologies to enhance their everyday and informal learning. While giving fair weight to the risks and harms associated with AI – including gender bias, surveillance, violence and exclusion – this piece contends that current discourses often obscure women's agency and innovation, cultivating a fear-based approach to digital literacies, policies and design practices when targeting young women. By placing positive and joyful digital experiences at the centre, we argue for a reclaiming of digital educational spaces by focusing on motivational approaches for women and girls. This demands building and sustaining digital freedoms by placing the responsibility on instead of curtailing bad actors. An inclusive, intersectional and gender-responsive approach to AI integration in education thereby recognizes the need for playful, personalized and emotionally attuned engagement with technology.

AI and learning risks in the Global South

'Does AI cause more harm than good in schools,' asks a technology learning Malaysian professor, Nurkhamimi Zainuddin, as he suggests for 'halting AI adoption until comprehensive legislation safeguards against risks' (Zainuddin, 2024). He is not alone in the global education space, with calls from policy-makers to educationists increasing to contain, curb and control AI in the classroom. Caution is the sentiment of today. As AI systems become increasingly integrated into educational infrastructures, a dominant discourse of moral panic has emerged (Sidorkin, 2025), which is characterized by fear of algorithmic harm, misinformation and declining academic integrity. AI in education is increasingly perceived as a threat to learning and assessment, with growing concerns that its use undermines academic integrity (Currie 2023), devalues critical thinking (Scribelli and Stevens, 2024) and facilitates plagiarism (Hutson, 2024) – concerns that have led to bans and restrictive policies across educational systems (Johnson, 2023; Bernstein, 2023). These anxieties are magnified around fears of misinformation and disinformation, especially on the 'moral' content around school curricula (Melo-Pfeifer and Dedecek Gertz, 2023).

Across policy contexts, this discourse is shaping restrictive interventions, including within the Global South, where the stakes of digital engagement are already heightened by systemic inequalities. Well-intentioned efforts to protect youth often result in technocratic solutions that prioritize deterrence, restriction and control, disproportionately impacting on women (Arora, 2024). This is especially acute in regions where patriarchal norms and infrastructural inequities already curtail access to digital tools. For marginalized groups, particularly young women, AI's integration into education is viewed as deepening existing digital and gender divides. Muayyad Ahmad et al. (2024) conducted a large-scale survey among Asian and African participants from eleven

countries, revealing that women are more likely than men to worry about data protection and the moral effects of AI use, which leads to less use of AI technologies by young women in education. While these concerns are understandable, recent scholarship also reveals that if these tools are used constructively and thoughtfully, they can bring more equitable and personalized learning in our everyday learning, and those most marginalized may be at the forefront of these benefits (Peters and Tukdeo, 2024).

Beyond risk-aversion in AI and education

We need to shift our emphasis from restrictive to more enabling approaches by encouraging the co-design of new learning systems that better integrate AI tools in ways that work for young people, especially young women and other systemically disadvantaged groups. It is essential to move the conversation from compliance to creativity, from paternalism to partnership. Drawing on empirical data and grounded narratives, we highlight how young women are not passive recipients of technological harms but active agents who creatively negotiate AI systems to serve their educational, emotional and aspirational needs (Sebastian et al., 2024). We advocate for educational frameworks that do not merely mitigate risk but actively enable them to thrive in digital ecosystems.

Reframing the narrative: From harm mitigation to enabling agency

The prevailing binary in AI discourse – safety versus freedom – creates a false dichotomy. For many women and girls, particularly those in under-resourced contexts (Bhatia, 2024), digital safety is not something imposed by institutional actors; it is practised daily through tactical strategies and emotional labour. These actors employ creative, low-tech interventions, such as pseudonymous profiles, shared device use, time-shifted engagement and encryption, to protect themselves while pursuing learning goals (Bhatia et al. 2021). In this context, digital safety and digital agency are co-constitutive.

A UNHCR project in Brazil with young refugee women reveals that their dominant motivations for getting and staying online and using AI-enabled platforms was for playful purposes – creating and sharing humorous memes, building eyebrow tutorials on YouTube, and playing with their visual profiles when seeking social connection (UNHCR, 2023). Many women wanted to educate others about their plight, their passions and their struggles, and saw these tools as essential to gain an audience for their storytelling. These engagements extend beyond formal curricula, challenging the notion that standardized, linear education systems can adequately capture the diverse driving factors for such learners. These demographics are constructing plural, playful and polymodal learning experiences, including with AI (Zheng and Han, 2024) – not despite systemic constraints, but often in defiance of them.

Case study: Tarishri and the role of AI in everyday learning

Our fieldwork in India (Bhatia et al., 2021) demonstrates that young women are reimagining new digital tools as allies in personal growth, circumventing limitations imposed by both technology infrastructure and social norms. In early 2024, we encountered Tarishri, an 18-year-old student living in a low-income settlement in Mumbai. Despite economic precarity, she had access to a smartphone and used ChatGPT extensively as a personalized learning companion. Her case is emblematic of how young women are

deploying generative AI tools to transcend the limitations of their educational environments (Herman and Payal, 2023).

Tarishri used ChatGPT not only to supplement her understanding of chemistry and biology but also to create customized tests, visualize complex processes and receive emotionally intelligent feedback. Her learning approach was hybrid, incorporating affect, play and critical thinking. She reported using different learning modes depending on her mood and cognitive capacity. For instance, on days of low concentration, she preferred visual diagrams; on high-focus days, she engaged in self-assessment quizzes. ChatGPT's adaptability allowed her to move fluidly between these modes. As she explained,

> 'When I think about how I use AI in learning, I get excited because it is not boring or fixed. I can change the style or approach any day based on my mood and even the vibe of the day. Some days I am competitive, and I want to see how much I know about study topics – take a test and get feedback. And some days, I just want to draw charts and diagrams and color them, like the different phases of the cell cycle. And I can do all of it simultaneously and based on my requirements with ChatGPT. And because it is so personal, some days when I am struggling to learn a topic, ChatGPT reminds me that maybe I am struggling because I fought with my friend in the morning, and so my mind is somewhere else.'

This example disrupts dominant narratives that associate AI with academic dishonesty or dependency. For Tarishri, the AI system did not replace learning; it enhanced it. More importantly, it respected her emotional tempo, cultural rhythms and personal aspirations. This illustrates the need for a broader definition of educational value – one that includes affective learning, cultural relevance and personalization (Bhatia et al., 2023).

Systemic barriers and the gendered digital divide

Many countries frame AI through the deficit lenses. In India, Nigeria, Kenya and beyond, educators and parents express heightened concern about girls misusing technology. These fears translate into punitive restrictions, moralistic gatekeeping (Johnson, 2023) and overemphasis on compliance. Often, boys are encouraged to use AI for upskilling and innovation, while girls are subjected to control and suspicion.

Statistical data from organizations such as GSMA and UNESCO corroborate these trends. As of 2025, 90 per cent of adolescent girls in low-income countries remain offline (Global Education Monitoring Report Team, 2024), while their male peers are nearly twice as likely to access the internet. Women comprise only 35 per cent of STEM students globally, and a mere 26 per cent of the workforce in data and artificial intelligence (UNESCO, 2025). Such data points reflect not only access gaps but a systemic devaluation of women's technological fluency.

This regulatory asymmetry reinforces structural inequality. It sends a clear message: that girls are at risk, while boys are innovators; that surveillance is safety; and that control is care. A feminist ethics of AI demands the opposite. It insists on trust, co-agency and design justice.

Privacy, play and personalization in AI-enabled education

Our fieldwork and participatory workshops across the Global South underscore the ways in which women and girls actively engage in tactical negotiations of digital

privacy to sustain their educational aspirations. Within contexts of gendered surveillance and infrastructural limitations, these learners cultivate what can be described as vernacular digital literacies (Pathak-Shelat and Bhatia, 2024) – switching SIM cards, borrowing siblings' devices, employing pseudonyms and leveraging encrypted platforms such as WhatsApp to engage in learning. These are not simply acts of circumvention; they are demonstrations of adaptive expertise, where privacy is not passively granted but tactically constructed.

Such practices are essential not only for maintaining access but also for fostering emotionally resonant and playful engagements with AI-enabled tools. When institutional constraints restrict formal learning opportunities, AI systems – particularly generative models like ChatGPT – become vital companions that facilitate polymodal, asynchronous and affect-driven educational experiences. Young women use these tools to simulate classroom-like feedback, create visual and interactive learning materials, and explore taboo or sensitive topics without fear of moral judgement or disciplinary repercussions.

Importantly, these engagements often exist outside sanctioned pedagogical structures. They allow learners to rehearse new identities, interrogate social norms and pursue curiosity-driven knowledge acquisition. AI is not merely a repository of information; it becomes a responsive interlocutor – attuned to the user's emotions, time constraints and changing ability to focus or learn.

From a design and policy perspective, this necessitates a reorientation of how AI learning environments are conceptualized. Standardized, risk-centric protocols – centred on surveillance and compliance – fail to account for the nuanced ways that women and girls configure safety, autonomy and aspiration in digital spaces. To foster inclusive learning ecosystems, AI tools must embed customizable privacy controls, adaptive content flows, and localized, culturally relevant prompts. This includes features such as low-bandwidth access modes, narrative simulations, visual annotation tools and feedback mechanisms that are empathetic and non-punitive.

Ultimately, personalization and play are not peripheral to educational success; they are integral to it. For young women navigating AI in the Global South, these dimensions represent not indulgences but necessary affordances for meaningful, self-directed learning.

Policy and design implications: Co-creation as mandate

To create inclusive AI learning environments, systemic interventions must move from consultative tokenism to genuine co-design. This entails:

- Engaging young women in iterative, participatory design cycles that value lived experience as expertise.
- Ensuring platforms are optimized for intermittent access, low-bandwidth environments and shared device contexts.
- Integrating localized content that reflects regional languages, cultural norms and hybrid educational models.
- Moving beyond risk-centric frameworks to prioritize emotional safety, consent and contextual trust.

Such systems should not assume linear learning trajectories, individualized device ownership or homogeneous user needs. Instead, they must reflect the

complex intersectionality of gender, class, caste, geography and digital access. In doing so, they affirm that the futures of AI and education are not technological inevitabilities, but social contracts that must be co-authored.

Designing for desire and play

Reimagining AI in education for young women requires a shift from protectionist paradigms to enabling ecosystems. While acknowledging the importance of addressing algorithmic bias, AI-enabled gender-based violence and structural exclusion, we emphasize that risk mitigation must not come at the cost of agency, joy and innovation.

By embracing design philosophies rooted in empathy, cultural sensitivity and play, we can build digital environments where women and girls learn not under constraint, but in community; not through fear, but through freedom. In doing so, we centre the feminist promise of AI: not as a threat to be managed, but as a tool to reimagine education itself.

References

Ahmed, M., Subih, M., Fawaz, M., Alnuqaidan, H., Abuejheishesh, A., Naqshbandi, V. and Alhalaiqa, F. 2024. Awareness, benefits, threats, attitudes, and satisfaction with AI tools among Asian and African higher education staff and students. *Journal of Applied Learning & Teaching*, Vol. 7, No. 1. Sydney, Kaplan Business School Australia. https://doi.org/10.37074/jalt.2024.7.1.10 (Accessed 13 August 2025.)

Arora, P. 2024. *From Pessimism to Promise: Lessons from the Global South on Designing Inclusive Tech*. Cambridge, MIT Press.

Bernstein, G. 2023. ChatGPT is the wake-up call schools need to limit tech in classrooms. *Time*. https://time.com/6266311/chatgpt-tech-schools (Accessed 13 August 2025.)

Bhatia, K. V. 2024. *Children's Digital Experiences in Indian Slums: Technologies, Identities, and Jugaad*. Amsterdam, Amsterdam University Press.

Bhatia, K. V., Arora, P. and Pathak-Shelat, M. 2021. Good girls don't go online: Unpacking the quotidian playful resilience influencing girls' social and digital engagements. *International Journal of Communication*, Vol. 15. Long Angeles, USC Annenberg Press, p. 19. https://ijoc.org/index.php/ijoc/article/view/17552 (Accessed 13 August 2025.)

Bhatia, K. V., Pathak-Shelat, M. and Arora, P. 2023. DIY education in the digital era: Youth-driven learning strategies and curricula for the future of work opportunities. *Education Information Technologies*, Vol. 29. Mödling, IFIP, pp. 1407–1426. https://doi.org/10.1007/s10639-023-11750-4 (Accessed 13 August 2025.)

Currie, G. M. 2023. Academic integrity and artificial intelligence: Is ChatGPT hype, hero or heresy? *Seminars in Nuclear Medicine*, Vol. 54, No. 5. Amsterdam, Elsevier, pp. 719–730. https://doi.org/10.1053/j.semnuclmed.2023.04.008 (Accessed 13 August 2025.)

Global Education Monitoring Report Team. 2024. *Global education monitoring report 2024, gender report: technology on her terms*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000389406 (Accessed 13 August 2025.)

Herman, L. M. and Payal, A. 2023. Decolonizing creativity in the digital era. *IASDR 2023: Life-Changing Design*. https://doi.org/10.21606/iasdr.2023.307 (Accessed 13 August 2025.)

Hutson, J. 2024. Rethinking plagiarism in the era of generative AI. *Journal of Intelligent Communication*, Vol. 3, No. 2. Birmingham, Scientific Publishing Limited. https://doi.org/10.54963/jic.v4i1.220 (Accessed 13 August 2025.)

Johnson, A. 2023. ChatGPT in schools: Here's where it's banned – and how it could potentially help students. *Forbes*. https://www.forbes.com/sites/ariannajohnson/2023/01/18/chatgpt-in-schools-heres-where-its-banned-and-how-it-could-potentially-help-students (Accessed 13 August 2025.)

Melo-Pfeifer, S. and Dedecek Gertz, H. 2023. Learning about disinformation through situated and responsive pedagogy: Bridging the gap between students' digital and school lives. L. Parker (ed.), *Education in the Age of Misinformation*. Cham, Palgrave Macmillan.

Pathak-Shelat, M. and Bhatia, K. V. 2024. Reconceptualizing ICTD: Prioritizing place-based learning experiences, socio-economic realities, and individual aspirations of young students in India. *Social Sciences*, Vol. 13, No. 7. Basel, MDPI. https://doi.org/10.3390/socsci13070379 (Accessed 13 August 2025.)

Peters, M. A. and Tukdeo, S. 2024. Beyond the utopic/dyspotic frames: Towards a research agenda for AI in education and development (AI4ED) in the Global South. *Contemporary Education Dialogue*, Vol. 22, No. 1. Thousand Oaks, SAGE Publications, pp. 176–184.

Scribelli, D. and Stevens, B. 2024. Is AI flattening the curve of critical thinkers leaving behind a cognitive gap for learners? *Issues in Information Systems*, Vol. 25, No. 3. International Association for Computer Information Systems (IACIS), pp. 133–147. https://doi.org/10.48009/3_iis_2024_111 (Accessed 13 August 2025.)

Sebastian, P., Sharma, S., Iivari, N., Kinnula, M., Monga, C., Verma, D. and Shahroz Abbas, M. 2024. Emerging technologies in Global South classrooms: Teachers imagining future of education. *PDC '24*, Vol. 1. New York, ACM, pp. 234–247. https://doi.org/10.1145/3666094.3666109 (Accessed 13 August 2025.)

Sidorkin, A. M. 2025. AI in education in the media. *AI-EDU Arxiv*. Sacramento, AI-EDU Arxiv. https://doi.org/10.36851/ai-edu.vi.5460 (Accessed 13 August 2025.)

UNESCO. 2025. *Closing the digital divide for women and girls in Africa through education*. Paris, UNESCO. https://www.unesco.org/en/gender-equality/education/digital-divide (Accessed 13 August 2025.)

UNHCR. 2023. *The digital leisure divide and the forcibly displaced*. Geneva, UNHCR. https://www.unhcr.org/innovation/wp-content/uploads/2023/04/The-Digital-Leisure-Divide-Field-Research.pdf (Accessed 13 August 2025.)

Zainuddin, N. 2024. Does artificial intelligence cause more harm than good in schools? *International Journal of Language Education and Applied Linguistics*, Vol. 14, No. 1. Pekan, Universiti Malaysia Pahang Al-Sultan Abdullah Publishing. https://doi.org/10.15282/ijleal.v14i1.10432 (Accessed 13 August 2025.)

Zheng, S. and Han, M. 2024. The impact of AI enablement on students' personalized learning and countermeasures: A dialectical approach to thinking. *Journal of Infrastructure, Policy and Development*, Vol. 8, No. 14. El Monte, EnPress Publisher. https://doi.org/10.24294/jipd10274 (Accessed 13 August 2025.)

Conceptual clarity: The missing link in the implementation of AI technologies for inclusive education

Yuchen Wang

In 2024, we celebrated the 30th anniversary of the Salamanca statement (UNESCO, 1994), a historical agreement that marked a global endeavour to ensure education is inclusive for all, regardless of learners' backgrounds and differences. Despite the evident shift in the educational landscape due to progressive policies relating to equality, diversity and inclusion, the question of how to realize inclusive education in classrooms remains a pressing challenge. With the increasing interest in AI integration in education, often coupled with a promise to solve longstanding educational challenges, we must consider whether AI may and to what extent promote inclusivity in learning and teaching. Although there are some discussions about the relations between AI and inclusive education (Pagliara et al., 2024), my observation is that insufficient understanding of what inclusive education means can lead to mistaken assumptions about the usefulness of AI technologies. I will draw on two examples of popular AI applications to illustrate my concern.

Assistive technology

The implementation of assistive technology is often presented as a key demonstrative example of AI's role in supporting inclusive education (Samson and Pothong, 2025). Indeed, recent developments in AI have been advancing the performance of assistive technology impressively, and it is not my intention to dispute how the right kind of assistive technology can enhance lived experiences for many. However, what tends to be ignored is that the real-life context of learning and teaching, such as in an inclusive classroom, is complex, involving social, cultural and political factors that shape our decision-making over who learns what and how. The differences the deployment of any AI technology can make still largely depend on how the barriers in the learning environment are addressed, beyond supporting an individual learner's specific functions. For instance, a student may benefit from having a new AI-powered wheelchair, which gives them better mobility control. Nevertheless, at school, they might still be expected to always stay at the front of a classroom near the entrance, while their peers have the choice to move between learning tasks and join friends. The introduction of the new technology has therefore not fully addressed the experiences of division and marginalization.

The main issue to note here is the startling similarity between a traditional special educational needs approach to supporting diversity and the current practice of technology deployment in the name of 'inclusion'. Both approaches identify educational 'difficulties' within individuals, overlooking how social, cultural and political contexts give rise to the perceived difficulties. Both approaches overly focus on providing something special, extra or additional for the concerned individuals to mediate or fix their 'needs', so they could fit into an intact and normative process of learning and teaching. New technologies like AI are frequently presented in such terms.

However, the special educational needs approach has long been critiqued for its limited impact on achieving inclusion (Florian, 2014). Research shows that to enact inclusive pedagogy, competent practitioners should consider extending what is generally available to all without marginalizing some, what values and beliefs

are at play, the dynamic interactions in a learning community and, importantly, the learners' experiences (Florian and Beaton, 2017; Florian and Black-Hawkins, 2011). While the need to integrate assistive technology appropriately into classroom practices was noted previously (Chambers, 2020; Magana, 2019), I argue that it is particularly helpful for practitioners to reflect on its use through the lens of inclusive pedagogy. Assistive technology should not be perceived as a stand-alone solution beyond one type of specialist support, as inclusive learning and teaching will require much more significant pedagogical transformation.

Personalized tutoring systems

The other example of AI technology frequently brought up in discussions of inclusive education is the personalized tutoring system. It is not uncommon to hear remarks such as, 'We're at the cusp of using AI for probably the biggest positive transformation that education has ever seen … And the way we are going to do that is by giving every student on the planet an artificially intelligent but amazing personal tutor' (Khan, 2023). Personalized tutoring systems are often uncritically assumed to be able to provide technical solutions to complex educational challenges, and the existing research, mostly focusing on the impact on academic performance, tends to support such an opportunistic view (Villegas-Ch et al., 2025). However, the pedagogical beliefs underpinning this type of technology and its purposes – the intended and the served – have rarely been examined.

First, the assumption that personalized teaching is what we ultimately need to help diverse learners flourish cannot be isolated from the influence of individualism, particularly in Anglo-Eurocentric societies. Every individual, by being taught in a way tailored to their learning differences, may then achieve their full potential, and the role of personalized tutoring systems is often to track the learning trajectories of individuals to offer more differentiated prompts, tasks and guidance. While a personalized tutoring system may bring some benefits to learners in terms of academic progress, its overemphasis on the idea that 'individuals learn by themselves and for themselves' is a rather problematic interpretation of inclusive education. Research in the field of inclusive education, including my study on inclusive schooling in China (Wang, 2021), confirms that the hugely important impact of inclusive education resides in relationships, collaboration and togetherness in a learning community. By participating in inclusive education, we are also experiencing something rather profound about us as humans – how to learn with and from each other, how to respect and celebrate differences, and how to live together in peace and harmony. The current implementation of personalized tutoring systems, which often literally place learners on isolated computer workstations, appears to add very little to the values of inclusive education.

Second, personalized tutoring systems tend to maintain the status quo of educational provision. For example, as pointed out in UNESCO's (2021) *AI and education: guidance for policy-makers*, these tutoring systems are often developed with a pre-determined curriculum and sets of fixed objectives. Such design reinforces a restricted view of learning outcomes in contrast to the diverse range of interests, strengths and achievements we should recognize in our learners, and it potentially fuels a curriculum of limited relevance to their lives and concerns (such as an interest in decolonization). Furthermore, these tutoring systems work hand in hand with measuring and testing based on a bell-curve model, maintaining competition and selection in education that will inevitably legitimize marginalization and the failure of some learners (Slee, 2011). Personalized tutoring systems appear to exist in their

apolitical vacuum, oblivious to opportunities for democratic changes in response to diversity.

Inclusion is an invitation

The global movement of inclusive education is a means to its very own end (Knight, 2000). Human diversity, the differences between us, should make a difference to the progression of education. Inclusive education is an invitation to change, and it is a deeply ethical project (Slee and Tait, 2022). There is no magical formula or quick fix – getting it right for our learners requires principled pedagogical actions and relentless negotiations. On the contrary, the ongoing AI integration in education often promotes the technology as a catch-all solution for all complex challenges, diverting us from developing a critical understanding of its wider impacts and implications (Knox et al., 2019; Shi and Palenski, 2024). Unpacking the above examples of assistive technology and personalization sheds light on the limitations of AI technologies in initiating the substantial changes required to make inclusion a reality. The key problem is that the pedagogical beliefs underpinning the implementation of AI technologies are not aligned with the principles of inclusive education. Without tackling this gap promptly, we risk repeating rather than disrupting unhelpful approaches. As discussed above, the missing link is conceptual clarity.

It is beyond the capacity of this short paper to offer comprehensive guidance; nonetheless, the following action points may be considered to advance policy, practice and research:

- Examine AI and inclusive education policies to review how key concepts are defined and to what extent a solutionist view of technology is reinforced. For instance, Scotland's National Framework for Inclusion (Scottish Universities Inclusion Group, 2022) is a useful example for guiding reflections about technology and exclusion.
- Develop new research informed by the accumulated wealth of research insights in the field of inclusive education, such as when evaluating AI's impact in classrooms and engaging AI-based pedagogical practice. Future research needs to foreground the views of learners, truly respecting their rights to be heard (Wang, 2023).
- Ensure teachers' meaningful participation in deciding what AI technologies are used and how, as we must value their indispensable role – the commitment, agency and creativity – in overcoming challenges to realize inclusive education. We should also consider how teacher education curricula may evolve to support student teachers and teachers' critical use of AI technologies for inclusion.
- Initiate collaborative and transdisciplinary working across traditional divides of disciplines and non-academic stakeholders to enable knowledge sharing about the nuances of the intersections between AI and inclusive education, and to co-create context-sensitive roadmaps for transformative changes.[1]

1. See a similar process led by the AIED Unplugged project in Brazil (Isotani et al., 2023).

References

Chambers, D. 2020. Assistive technology supporting inclusive education: Existing and emerging trends. D. Chambers (ed.), *Assistive Technology to Support Inclusive Education*. Leeds, Emerald, pp. 1–16.

Florian, L. 2014. Reimagining special education: Why new approaches are needed. L. Florian (ed.), *The SAGE Handbook of Special Education: Two Volume Set*. London, Sage Publishing, pp. 9–22.

Florian, L. and Beaton, M. 2017. Inclusive pedagogy in action: Getting it right for every child. *International Journal of Inclusive Education*, Vol. 22, No. 8. London, Taylor & Francis, pp. 870–884.

Florian, L. and Black-Hawkins, K. 2011. Exploring inclusive pedagogy. *British Educational Research Journal*, Vol. 37. No. 5. Hoboken, John Wiley & Sons, pp. 813–828. https://www.jstor.org/stable/23077052 (Accessed 12 August 2025.)

Isotani, S., Bittencourt, I. I., Challco, G.C., Dermeval, D. and Mello, R. F. 2023. AIED Unplugged: Leapfrogging the digital divide to reach the underserved. *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky*. Cham, Springer, pp. 772–779.

Khan, S. 2023. *How AI could save (not destroy) education*. New York City, TED Conferences. https://www.ted.com/talks/sal_khan_how_ai_could_save_not_destroy_education/transcript (Accessed 12 August 2025.)

Knight, T. 2000. Inclusive education and educational theory: Inclusive for what? *Melbourne Studies in Education*, Vol. 41. London, Informa UK, pp. 17–43.

Knox, J., Wang, Y. and Gallagher, M. 2019. Introduction: AI, inclusion, and 'everyone learning everything'. J. Knox, Y. Wang and M. Gallagher (eds), *Artificial Intelligence and Inclusive Education: Speculative Futures and Emerging Practices*. Singapore, Springer Nature, pp. 1–13.

Magana, A. J. 2019. Disruptive classroom technologies. *Oxford Research Encyclopedia of Education*. Oxford, Oxford University Press.

Pagliara, S. M., Bonavolontà, G., Pia, M., Falchi, S., Zurru, A. L., Fenu, G. and Mura, A. 2024. The integration of artificial intelligence in inclusive education: A scoping review. *Information*, Vol. 15, No. 12. Basel, MDPI, pp. 774. https://doi.org/10.3390/info15120774 (Accessed 12 August 2025.)

Samson, R. and Pothong, K. 2025. *A learning curve? A landscape review of AI and education in the UK*. London, Ada Lovelace Institute. https://www.adalovelaceinstitute.org/wp-content/uploads/2025/01/Ada-Lovelace-Institute-A-learning-curve.pdf (Accessed 12 August 2025.)

Scottish Universities Inclusion Group. 2022. *National Framework for Inclusion (3rd edition)*. Edinburgh, General Teaching Council for Scotland. https://www.gtcs.org.uk/documents/national-framework-3rd-edition-2022 (Accessed 12 August 2025.)

Shi, L. P. and Palenski, T. 2024. *From optimism to caution: An analysis of the seemingly balanced policy discourse on artificial intelligence in education*. London, BERA. https://www.bera.ac.uk/publication/from-optimism-to-caution (Accessed 12 August 2025.)

Slee, R. 2011. *The Irregular School: Exclusion, Schooling and Inclusive Education*. Abingdon, Routledge.

Slee, R. and Tait, G. 2022. *Ethics and Inclusive Education: Disability, Schooling and Justice*. Berlin, Springer Nature.

UNESCO. 1994. *The Salamanca statement and framework for action on special needs*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000098427 (Accessed 12 August 2025.)

—. 2021. *AI and education: Guidance for policy-makers*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709 (Accessed 12 August 2025.)

Villegas-Ch, W., Buenano-Fernandez, D., Navarro, A. M. and Mera-Navarrete, A. 2025. Adaptive intelligent tutoring systems for STEM education: Analysis of the learning impact and effectiveness of personalized feedback. *Smart Learning Environments*, Vol. 12, no. 41. Berlin, Springer Nature, pp. 41. https://doi.org/10.1186/s40561-025-00389-y (Accessed 12 August 2025.)

Wang, Y. 2021. 'Teachers did not let me do it': Disabled children's experiences of marginalisation in regular primary schools in China. *Disability and the Global South*, Vol. 8, No. 2. Attard, The Critical Institute, pp. 2053–2070. https://dgsjournal.org/wp-content/uploads/2021/02/dgs_08_02_03.pdf (Accessed 12 August 2025.)

Wang, Y. 2023. 'It is the easiest thing to do': University students' perspectives on the role of lecture recording in promoting inclusive education in the UK. *Teaching in Higher Education*, Vol. 29, No. 8. London, Informa UK, pp. 1974–1991. https://doi.org/10.1080/13562517.2022.2162814 (Accessed 12 August 2025.)

Inclusion or illusion? Rethinking AI for learners who are deaf or hard of hearing in under resourced settings (the Global South)

Marloes Williams van Elswijk

We are living in an era where AI is increasingly positioned as the great equalizer of education. The potential of Artificial Intelligence in Education (AIED) for inclusive education is vast, especially for learners living in marginalized communities. Yet, as access to the internet and AI expands its footprints in classrooms and educational policy (African Union, 2024; UNESCO, 2023; Marwala et al., 2023), a sobering question arises: can AIED truly deliver access to long-overdue equitable education for primary school learners who are Deaf or Hard of Hearing (DHH) in under resources settings? Or will it perpetuate and deepen exclusion under the guise of innovation?

In many cases, AIED tools are introduced faster than policy and guidelines can regulate (Giannini, 2023). Without addressing the long-standing gaps in teacher training, DHH inclusive curriculum development or linguistic access, AIED risks overlooking the unique and diverse needs of DHH learners. Community-led design and attention to intersectional barriers needs to be more than policy – it needs to be enacted.

Beyond one-size-fits-all: Diversity meets intersectionality

The term DHH represents a wide spectrum of identities, communication modes, preferences and learning needs. Some children identify culturally as Deaf and use national sign language(s). Others rely on spoken language and/or assistive technologies, such as hearing aids.

Adding to the diverse needs and preferences is the sometimes compounded issue of language deprivation, particularly for the 90 per cent to 95 per cent of DHH children who are born to hearing parents and who have lost their ability to hear at birth or prelingual age (Mitchell and Karchmer, 2004). A DHH child who has had very little language input needs different pedagogical and technological support than a DHH child exposed to a (sign) language-rich environment early (Skotara et al., 2012; Hall et al., 2016). Simply put, the first needs to learn a language before literacy, while the latter requires instruction in a language of preference. Diverse language acquisition also impacts on executive functions of DHH children (Caselli et al., 2020). An AIED tool would need to be able to assess and adapt to these various needs in language acquisition. The diversity of needs is topped by the complexity of intersectionality, where factors such as socio-economic status, data poverty and gender further shape the experiences of DHH children and their access to education and digital tools (Male and Wodon, 2017). For example, girls in low-income countries are restricted or discouraged from using digital technology due to sociocultural values, roles and interests and are less likely to attend or finish primary education (UNICEF, 2021; Webb et al., 2020; World Bank, 2025). Hence, a DHH girl faces intersectional challenges to equitable education.

DHH learners cannot be served by a one-size-fits-all solution. Inclusive AIED must begin by asking what works for whom, under which conditions and in what languages or modalities. Human-centred

AIED can bridge the educational divides, if shaped with these diverse layers and communities in mind.

Global push, local challenges

Misconceptions about hearing loss and education are deeply embedded within the layers of educational systems. For example, professionals still deem DHH learners depending on sign language in an inclusive setting as linguistically and socially isolated (De Meulder, 2025). Universal Design for Learning (UDL), a multimodal approach often used to include all learners, does not always include the structures for dual-language exposure or culturally grounded teaching approaches that DHH learners need (Hall et al., 2019; Humphries et al., 2012). Captioning, as well as the possibility of augmented reality (AR) glasses to deliver supplementary information to learners (Zhang et al., 2024), has shown to improve vocabulary and incidental learning (Teng, 2019). The challenge there is accuracy: background noise, fast speaking or dialects that could alter words, resulting in misinterpretation or confusion. For example, the letter 'T' could be spelled as 'Tea' in captions. Many promising innovations are being put forward to improve access to education for marginalized learners; however, the educational needs of DHH learners are often still underestimated.

Community-led development, nevertheless, does show positive progress. A user study with sixteen DHH participants showed that they perceived conversations with AI persona who had DHH education experiences to be more human-like and trustworthy (Cheng et al., 2024). They could relate more to the AI persona with cultural knowledge of DHH communities. This shows that AIED tools designed with the DHH community rather than for them, are more promising. Inclusion should not be a retrofitted feature. It must be an original design principle, embedded in how funds are allocated, and how AIED systems are conceived, trained, tested and governed (Desai et al., 2024).

Yet, inclusion is not reached through AIED alone. AIED cannot replace the need for trained emphatic teachers. It cannot fix under-resourced schools. It cannot on its own remedy the lack of culturally and linguistically appropriate curricula for DHH learners. When treated as a shortcut, AIED risks masking these systematic gaps.

Techno solutionism

In the rush to deploy AIED, policy-makers may equate the presence of technology with progress, mistaking accessibility for full inclusion. The belief that complex social challenges, like educational inequality, can be solved through technical fixes alone is one-sided. Without complementary and appropriate classroom support, AIED can become performative, serving more to check accessibility boxes than genuinely support learning. It may reduce the urgency for actual system change. To ensure true inclusion and to maximize its potential, AIED must be integrated as an add-on, prioritizing the needs of those it serves.

A call to action

Community-led design mirroring the diverse layers of hearing loss

- Fund AIED projects that include DHH professionals in the design, pre-deployment and evaluation process.

Multimodal, multilingual

- Prioritize visual-spatial AIED that integrates indigenous (sign) language.

Layering AIED with human support

- Use as an add-on, not a replacement.
- Blend with DHH pedagogical methods and epistemology.

- Invest in system change, alongside innovation.

Conclusion

AI is not inherently inclusive; it mirrors the values and priorities of those who fund it, invest in it, build it and train it. Let us make AIED in DDH education truly human-centred, genuinely equitable and meaningfully accessible by taking deliberate, inclusive action at every stage – from funding to implementation, and from design to evaluation.

References

African Union. 2024. *Continental Artificial Intelligence Strategy*. Addis Ababa, African Union. https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf (Accessed 12 August 2025.)

Caselli, N. K., Hall, W. C. and Henner, J. 2020. American sign language interpreters in public schools: An illusion of inclusion that perpetuates language deprivation. *Maternal and Child Health Journal*, Vol. 24. London, Springer Nature, pp. 1323–1329.

Cheng, H., Chen, S., Perdriau, C. and Huang, Y. 2024. LLM-powered AI tutors with personas for d/Deaf and hard-of-hearing online learners. Ithaca, *arXiv*. https://doi.org/10.48550/arXiv.2411.09873 (Accessed 12 August 2025.)

De Meulder, M. 2025. Deaf in AI: AI language technologies and the erosion of linguistic rights. Ithaca, *arXiv*. https://doi.org/10.48550/arXiv.2505.02519 (Accessed 12 August 2025.)

Desai, A., De Meulder, M., Hochgesang, J. A., Kocab, A. and Lu A. X. 2024. Systemic biases in sign language AI research: A deaf-led call to reevaluate research agendas. Ithaca, *arXiv*. https://doi.org/10.48550/arXiv.2403.02563 (Accessed 12 August 2025.)

Giannini, S. 2023. *Generative AI and the future of education*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385877 (Accessed 12 August 2025.)

Hall, M. L., Eigsti I.-E., Bortfeld, H., Lillo-Martin, D. 2016. Auditory deprivation does not impair executive function, but language deprivation might: Evidence from a parent-report measure in deaf native signing children. *The Journal of Deaf Studies and Deaf Education*, Vol. 22, No. 1. Oxford, Oxford University Press, pp. 9–21.

Hall, M. L., Hall, W. C. and Caselli, N. K. 2019. Deaf children need language, not (just) speech. *First Language*, Vol. 39, No. 4. Thousand Oaks, SAGE Publications, pp. 367–395. https://doi.org/10.1177/0142723719834102 (Accessed 12 August 2025.)

Humphries, T., Kushalnagar, P., Mathur, G., Napoli, D. J., Padden, C., Rathmann, C. and Smith, S. 2012. Language acquisition for deaf children: Reducing the harms of zero tolerance to the use of alternative approaches. *Harm Reduction Journal*, Vol. 9, No. 1. London, Springer Nature. https://doi.org/10.1186/1477-7517-9-16 (Accessed 12 August 2025.)

Male, C. and Wodon, Q. 2017. *Disability Gaps in Educational Attainment and Literacy*. Washington, DC, World Bank. https://documents1.worldbank.org/curated/en/396291511988894028/txt/121762-replacement-PUBLIC-WorldBank-GapsInEdAttainmentLiteracy-Brief-v6.txt (Accessed 12 August 2025.)

Marwala, T., Fournier-Tombs, E. and Stinckwich, S. 2023. *The Use of Synthetic Data to Train AI Models: Opportunities and Risks for Sustainable Development*. Tokyo, United Nations University. https://collections.unu.edu/eserv/UNU:9216/UNU-TB_1-2023_The-Use-of-Synthetic-Data-to-Train-AI-Models.pdf (Accessed 12 August 2025.)

Mitchell, R. E. and Karchmer, M. A. 2004. Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, Vol. 4, No. 2. New York, JSTOR, pp. 138–163.

Skotara, N., Salden, U., Kügow, M., Hänel-Faulhaber, B. and Röder B. 2012. The influence of language deprivation in early childhood on L2 processing: An ERP comparison of deaf native signers and deaf signers with a delayed language acquisition. *BMC Neuroscience*, Vol. 13, No. 44. London, Springer Nature. https://doi.org/10.1186/1471-2202-13-44 (Accessed 12 August 2025.)

Teng, M. F. 2019. The effects of video caption types and advance organizers on incidental L2 collocation learning. *Computers and Education*, Vol. 142. Amsterdam, Elsevier. https://doi.org/10.1016/j.compedu.2019.103655 (Accessed 12 August 2025.)

UNESCO. 2023. *Guidance for generative AI in education and research*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693 (Accessed 12 August 2025.)

UNICEF. 2021. *Advancing girls' education and gender equality through digital learning*. New York, UNICEF. https://www.unicef.org/documents/advancing-girls-education-and-gender-equality-through-digital-learning (Accessed 12 August 2025.)

Webb, D., Barringer, K., Torrance, R. and Mitchell, J. 2020. *Girls' Education and EdTech: A Rapid Evidence Review*. Bolton, EdTech Hub. https://doi.org/10.5281/zenodo.4556939 (Accessed 12 August 2025.)

World Bank. 2025. *Girls' education*. New York, World Bank. https://www.worldbank.org/en/topic/girlseducation (Accessed 12 August 2025.)

Zhang, J., Zhu, R., Li, T. and Yi, C. 2024. Breaking barriers: Enhancing learning outcomes for deaf and hard of hearing students through augmented reality captioning. *ICIS 2024 Proceedings*. Atlanta, Association for Information Systems.

8. Reimagining AI in education policy: Evidence and geopolitical realities

Human and machine: Policy implications of emerging AI capabilities

George Siemens

Evidence is accumulating about the transformative capabilities of AI for humanity. AI competes in the domain of human cognition (Siemens et al., 2022), often exceeding human-level performance. The capabilities are not limited to only knowledge tasks. Katja Schlegel et al. (2025) find that AI can outperform humans on emotional intelligence tests. In medical diagnosis, AI is already outperforming human doctors and offers the prospect of personalized medicine (Nori et al., 2025). Michael Heinz et al. (2025) share early evidence that AI can offer human-level therapy support. The complexity and sophistication of today's neural networks goes beyond simply generating novel text, video, images and audio, and may lead to deeper understanding of human cognition and the human brain (Binz et al., 2025). While these represent a sampling of early research outputs and clarity will come over time as additional studies are conducted, a message is emerging: AI, with a range of ethical and humanistic concerns, has the pure cognitive capability to augment human cognition and, in some instances, to outperform humans.

AI is a technology that marks a departure from the ones that have defined and shaped human history. While there are various types of intelligence that are expressed in plant and animal kingdoms (Bridle, 2022), humanity has summited reasoning, planning and logic. The current trajectory of AI's progress challenges that cognitive supremacy. In some domains, AI pushes humanity to the side as an observer, rather than an active decider, in rapidly evolving and complex landscapes (Johnson et al., 2013).

AI and education

How AI impacts on traditional teaching and learning practices is still not clear. Some reports suggest that the change will be significant, enabling learners to vastly outperform the existing standards of academic performance (De Simone et al., 2025). Recent pre-prints suggest that AI may harm certain learning activities, such as essay writing (Kosmyna et al., 2025). More moderate conclusions indicate that students may learn to rely on, rather than learn from, AI (Darvishi et al., 2024). Additionally, knowledge processes, such as complex problem solving, may benefit significantly from AI (Joksimovic et al., 2023). The uncertain research evidence of impact presents educators and researchers with a range of practical and ethical questions regarding development, adoption and deployment of AI (Bond et al., 2024), including concerns of reinforcing existing stereotypes and gender bias.

Skills and mindsets are also still being determined in order to succeed in the future workforce (National Academies of Sciences, Engineering, and Medicine,

2025), as some will be replaced and others augmented by AI. This combination of skill and knowledge uncertainty, coupled with the rapid advancement of AI, makes it challenging for organizations to define the capabilities that are most critical for the future. At best, current thinking in this space will be provisional until the pace of change settles or until clarity is achieved around the skills needed to navigate perpetual complexity.

What is clear, however, is that the learning model that students experience today will no longer prepare them for an AI-centric world. Traditional tasks of assessing learner capabilities, such as writing essays, can now readily be generated by large language models like Gemini, Claude or ChatGPT. In response, some universities have advocated for a return to traditional paper-based assessment. While these might offer a short-term response to the cheating capabilities inherent with AI use, they offer no meaningful long-term path forward.

There are early examples of some schools that have taken drastic steps as a systemic restructuring of the calendar by focusing on two hours a day of academic work and the balance on human and relational skills.[1] Systems-level transformation at a state and national level, however, requires careful coordination of the needs of multiple stakeholders, including students, industry, parents, society and government.

In the immediate future, educational leaders must address questions about the requisite skills and literacies for effective AI utilization in learning, work and life. Numerous frameworks, often categorized as twenty-first-century or soft skills, have emerged over the past decade. However, most of these frameworks predate contemporary generative AI capabilities and necessitate revision to incorporate the now commonplace cognitive task-sharing and knowledge work (Shao et al., 2025). A primary policy objective is the identification and communication of future workforce requirements. These should be developed through collaborative efforts among peer institutions, regions and nations.

Given the growing influence and decision-making power of frontier AI laboratories, it is essential for educational systems to adopt a cohesive unified strategy that asserts their requirements. Without coordinated action, the voices of educators, students and communities risk being overshadowed by corporate agendas that may not fully align with educational values or long-term societal needs. A unified approach enables education leaders to articulate clear expectations, negotiate from a position of collaborative strength, and ensure that AI developments and deployments serve to enhance learning and equity.

Setting a strategy

Leaders cannot be complacent. There are increased responses available in establishing policies on AI adoption and use in education. Ajay Agrawal et al. (2022) detail the approaches available to general purpose technologies (those that have wide-ranging and cross-discipline impact), ranging from discrete adoption that has minimal systems impact to broadscale adoption that restructures entire industries. Paul David (1990) describes how electrification,

1. See https://alpha.school

over a century ago, started with small-scale changes and eventually resulted in system-level transformation that enabled new factory models, including the assembly line. All indications currently suggest that AI will be transformative, perhaps our last invention, according to Stephen Hawking (University of Cambridge, 2016).

To navigate the systemic transformations driven by AI, policy-makers will need capacities for anticipatory governance, integrating horizon scanning, scenario planning and adaptive regulatory frameworks that can evolve with technological shifts. This requires cross-sector coordination mechanisms linking education, labour, ethics and innovation policy. Governance must be proactive and iterative, and use experimental approaches that test interventions, learn from outcomes and adjust rapidly. Iterative and rapid testing of governance and policy is currently not the norm in education policy-making. However, this will need to change to match the pace of social, learning and employment changes that AI is expected to generate. This will require forward-looking strategies that prioritize equity-by-design, ensuring that AI's benefits are distributed inclusively, while embedding safeguards for transparency, accountability and human rights. In essence, the capacity to govern AI's transformation is not just about managing risks, but about shaping a shared, adaptive vision for societal futures, in which human agency remains central, and education systems serve this primary objective.

AI as statecraft

AI is statecraft. Its burgeoning capabilities and pervasive influence are not merely a technological advancement but a fundamental component of modern statecraft. AI's development, deployment and governance directly impacts on national security, economic competitiveness and social stability. The structure of society is shaped by how leaders develop policies to support learners and educators, notably around skills and mindsets required and the pedagogical processes whereby they attain them. This is contested space, and regions like the People's Republic of China and the United States are approaching AI with the same strategic foresight and comprehensive planning typically reserved for critical instruments of power, such as military capabilities, diplomatic relations and economic policy (Hook, 2025; The White House, 2025).

Treating AI as statecraft necessitates a multifaceted approach. This includes significant investment in domestic AI research and development to foster innovation and maintain a competitive edge, alongside robust cybersecurity measures to protect critical AI infrastructure from malicious actors. Much of the existing AI innovation in frontier models is occurring in the United States and the People's Republic of China. Smaller innovation centres and frontier models exist in Canada (Cohere) and the European Union (Mistral). International cooperation becomes paramount to establish norms, standards and ethical guidelines for AI development and use, but also to leverage shared resources to effectively compete with heavily funded frontier labs.

The task of education is to create a capable society and workforce that benefits from AI's capability and centres its impact on human well-being. For education ministers and governance bodies, the rapid advancements in AI necessitates not only anticipating the

impact of AI on curriculum design, teaching methodologies and assessment strategies, but also addressing ethical considerations, data privacy and the potential for digital divides. Policies must foster an environment that encourages innovation, while ensuring equitable access to AI-powered learning tools and resources for all students, regardless of socio-economic background or geographic location. This is the work of systems change: imagining and creating a new system of learning.

Governance also needs to consider the implications of AI on the future workforce and the skills required for success in an increasingly automated world. This includes promoting collaboration, creativity, problem-solving and adaptability. These skills are generally viewed as complementing AI, rather than being replaced by it. Policy frameworks should also facilitate ongoing professional development for educators, equipping them with the knowledge and competencies to effectively integrate AI into their pedagogy.

Ultimately, new approaches to policy-making for education in the age of AI must be collaborative, involving not just government officials, but also educators, technologists, industry leaders and the broader community. This collaborative effort will ensure that policies are comprehensive, responsive and ultimately serve to prepare students for a future where human and artificial cognition coexist and collaborate.

Preparing for human-centric AI

Longer-term vision and strategy should provide the desired future vision of AI and a projected pathway for its realization. UNESCO's *AI and education: guidance for policy-makers* report (Miao et al., 2025) states that 'applying AI in education should be to enhance learning, enabling every learner to develop their individual potential, which policies should reflect and support'. This framing, especially at an education minister's level, needs an additional lens on the system itself. What types of future education systems should we conceive of and what is the social contract that should underpin it (UNESCO, 2021)? How can schools and universities be restructured to account for the capabilities that AI makes available? How should assessment and evaluation be conducted? Does the existing heavily in-person education model still make social and economic sense?

Addressing questions of this nature will require visionary leadership. Simple literacies, such as teaching students what AI already does better than humans, is not a solid foundation for a new model of learning. Large-scale, multi-stakeholder conversations and engagements are needed to imagine a new classroom, a new focus on campus, and a new lens for determining competency and capability. Critical to this challenge is the need to meet the existential angst that fosters public discourse with a vision-oriented and hopeful future, where learners have agency in living their lives and society has the capacity to choose an intentional and ethical use of AI.

Thomas Kuhn (1962) argues that progress in science results in the collection of anomalies that do not fit the prominent frameworks of an era. Over time, those anomalies accrue and a dramatic change and restructuring of science emerges. Decades of technological change, while schooling retained its primary structure, has generated anomalies and affordances that now, coupled with AI, enable a window for a new, dramatic and extraordinary rethinking of the education experience. It is time for a new system of learning.

References

Agrawal, A., Gans, J. and Goldfarb, A. 2022. *Power and Prediction: The Disruptive Economics of Artificial Intelligence*. Boston, Harvard Business Review Press.

Binz, M., Akata, E., Bethge, M., Brändle, F., Callaway, F., Coda-Forno, J., Dayan, P., Demircan, C., Eckstein, M. K., Éltető, N., Griffiths, T. L., Haridi, S., Jagadish, A. K., Ji-An, L., Kipnis, A., Kumar, S., Ludwig, T., Mathony, M., Mattar, M., Modirshanechi, A., Nath, S. S., Peterson, J. C., Rmus, M., Russek, E. M., Saanum, T., Schubert, J. A., Schulze Buschoff, L. M., Singhi, N., Sui, X., Thalmann, M., Theis, F. J., Truong, V., Udandarao, V., Voudouris, K., Wilson, R., Witte, K., Wu, S., Wulff, D. U., Xiong, H. and Schulz, E. 2025. A foundation model to predict and capture human cognition. *Nature.* Berlin, Springer Nature, pp. 1–8. https://doi.org/10.1038/s41586-025-09215-4 (Accessed 17 August 2025.)

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W. and Siemens, G. 2024. A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, Vol. 21, No. 4. Berlin, Springer Nature. https://doi.org/10.1186/s41239-023-00436-z (Accessed 17 August 2025.)

Bridle, J. 2022. *Ways of Being: Animals, Plants, Machines: The Search for a Planetary Intelligence*. New York, Farrar, Straus and Giroux.

David, P. A. 1990. The dynamo and the computer: An historical perspective on the modern productivity paradox. *The American Economic Review*, Vol. 80, No. 2. New York, ITHAKA, pp. 355–361. https://www.jstor.org/stable/2006600 (Accessed 17 August 2025.)

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D. and Siemens, G. 2024. Impact of AI assistance on student agency. *Computers & Education*, Vol. 210. Amsterdam, Elsevier. https://doi.org/10.1016/j.compedu.2023.104967 (Accessed 17 August 2025.)

De Simone, M. E., Tiberti, F. H., Mosuro, W., Manolio, F. A., Barron, M. R. and Dikoru, E. J. 2025. From chalkboards to chatbots: Transforming learning in Nigeria, one prompt at a time. *World Bank Blogs*. Washington, DC, The World Bank. https://blogs.worldbank.org/en/education/From-chalkboards-to-chatbots-Transforming-learning-in-Nigeria (Accessed 17 August 2025.)

Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., Jewett, A. D., Salzhauer, A. J., Griffin, T. Z. and Jacobson, N. C. 2025. Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI*, Vol, 2, No. 4. Waltham, Massachusetts Medical Society.

Hook, D. 2025. Deepseek and the new geopolitics of AI. *Figshare*. https://figshare.com/articles/online_resource/Deepseek_and_the_New_Geopolitics_of_AI/29336588 (Accessed 17 August 2025.)

Johnson, N., Zhao, G., Hunsader, E., Qi, H., Johnson, N., Meng, J. and Tivnan, B. 2013. Abrupt rise of new machine ecology beyond human response time. *Scientific Reports*, Vol. 3, No. 1. Sci-Hub. https://doi.org/10.1038/srep02627 (Accessed 17 August 2025.)

Joksimovic, S., Ifenthaler, D., Marrone, R., De Laat, M. and Siemens, G. 2023. Opportunities of artificial intelligence for supporting complex problem-solving: Findings from a scoping review. *Computers and Education: Artificial Intelligence*, Vol. 4. Amsterdam, Elsevier. https://doi.org/10.1016/j.caeai.2023.100138 (Accessed 17 August 2025.)

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., Braunstein, I. and Maes, P. 2025. *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. Ithaca, arXiv. https://doi.org/10.48550/arXiv.2506.08872 (Accessed 17 August 2025.)

Kuhn, T. S. 1962. *The Structure of Scientific Revolutions*. Chicago, University of Chicago Press.

Maimann, K. 2025. Canadian universities grapple with evaluating students amid AI cheating fears. *CBC*. https://www.cbc.ca/news/canada/university-ai-exams-1.7551617 (Accessed 17 August 2025.)

Miao, F., Holmes, W., Huang, R. and Zhang, H. 2021. *AI and education: guidance for policy-makers*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709 (Accessed 17 August 2025.)

National Academies of Sciences, Engineering, and Medicine. 2025. *Artificial Intelligence and the Future of Work*. Washington, DC, The National Academies Press. https://doi.org/10.17226/27644 (Accessed 17 August 2025.)

Nori, H., Daswani, M., Kelly, C., Lundberg, S., Ribeiro, M. T., Wilson, M., Liu, X., Sounderajah, V., Carlson, J., Lungren, M. P., Gross, B., Hames, P., Suleyman, M., King, D. and Horvitz, E. 2025. Sequential diagnosis

with language models, Ithaca, arXiv. https://doi.org/10.48550/arXiv.2506.22405 (Accessed 17 August 2025.)

Schlegel, K., Sommer, N. R. and Mortillaro, M. 2025. Large language models are proficient in solving and creating emotional intelligence tests. *Communications Psychology*, Vol. 3, No. 80. Berlin, Springer Nature. https://doi.org/10.1038/s44271-025-00258-x (Accessed 17 August 2025.)

Shao, Y., Zope, H., Jiang, Y., Pei, J., Nguyen, D., Brynjolfsson, E. and Yang, D. 2025. *Future of work with AI agents: Auditing automation and augmentation potential across the U.S. workforce*. Ithaca, arXiv. https://doi.org/10.48550/arXiv.2506.06576 (Accessed 17 August 2025.)

Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S. and De Laat, M. 2022. Human and artificial cognition. *Computers and Education: Artificial Intelligence*, Vol. 3. Amsterdam, Elsevier. https://doi.org/10.1016/j.caeai.2022.100107 (Accessed 17 August 2025.)

UNESCO. 2021. *Reimagining our futures together: A new social contract for education*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000379707 (Accessed 17 August 2025.)

University of Cambridge. 2016. *The best or worst thing to happen to humanity – Stephen Hawking launches Centre for the Future of Intelligence*. Cambridge, University of Cambridge. https://www.cam.ac.uk/research/news/the-best-or-worst-thing-to-happen-to-humanity-stephen-hawking-launches-centre-for-the-future-of (Accessed 17 August 2025.)

The White House. 2025. *America's AI action plan*. Washington, DC, The White House. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf (Accessed 17 August 2025.)

Adding intelligence to AIED policy and practice

Ilkka Tuomi

In this think piece, I will ask what evidence-informed policy and practice can mean in Artificial Intelligence in Education (AIED). What counts as evidence, whose knowledge counts, and how research-based knowledge could inform policy and practice? I will argue that both the epistemology and ontology of AIED must be reconsidered; that we must focus on the development of learners' social agency instead of technical artefacts and their functionality; and that, instead of verified facts, new social sense-making and learning processes are needed to support the development of policy and practice.

AI is often declared to be an unavoidable technology that helps us learn knowledge and skills required in the imagined jobs of the AI-enabled future. Education, however, is more. Compulsory education, as it was understood in the Enlightenment tradition, generates members of society who can competently participate in social, cultural and economic life. Education, therefore, is inherently political and cultural. It creates epistemologies of ignorance and hierarchies of power. But it also creates undetermined social and personal development. In contrast to instrumental views that address the imagined needs of the industrial world under global economic competition, education can also be understood as the development of capabilities and capacity that expand our opportunities to live meaningful lives.

This more social and political understanding of education and its social functions means that we must view AIED more broadly than just a technology that increases the efficiency of learning. As large language models are now becoming experts in various domains of knowing, and knowledge is becoming a commodity, we therefore should ask why education is needed in the future. If it is not for producing useful skills and knowledge for the future labour market, what are the educational outcomes that matter? If the goal is something else than fast mastery of given curriculum content, what is the role of AI in education and how should its educational impact be measured?

In other words, we should ask what the impact is we are looking for and how can we know that the use of technology delivers that impact.

Evidence and impact

There have been repeated claims that technology will revolutionize education. But why should we think that future visions would be a reason to spend money and effort that could have other potentially useful uses. Why should we invest in dreams, when schools lack teachers, books and pens?

Evidence-based education emerged in the 1990s to address this problem (Hammersley, 2007). According to its proponents, the development of educational practice and policy was grounded on wishful thinking, intuition and hopes (Davies, 1999). To put education on the right track, educational interventions, including the use of educational technologies, should be based on robust evidence produced by research.

Evidence has since then become a key starting point in all policy-making. In particular, the European concept of 'better regulation' is based on the idea that the benefits of legislation must be justified by assessing their impact. It is now commonly accepted that policy should be informed

by research-based evidence, although policy cannot strictly speaking be based on science.

In education, evidence-based education led to the establishment of 'clearing houses' that collect and summarize research evidence on educational interventions. One of the leading initiatives in this field has been the What Works Clearinghouse (WWC),[1] set up in 2002 under the Institute of Educational Statistics of the US Department of Education. WWC has elaborated detailed guidelines on how educational research should be evaluated and to what extent its results should count as evidence in policy and practice. The highest-quality evidence, according to the WWC, can be gained through randomized controlled trials.

This 'gold standard' has been criticized from methodological and educational points of view (Cartwright, 2019; Parra and Edwards Jr, 2024; Tuomi, 2025). Experts in statistical methodology have pointed out that the validity of impact estimates cannot be guaranteed in practical experiments in educational contexts (Deaton and Cartwright, 2018; Heckman, 1991). Education researchers, in turn, have argued that experimental research designs miss the essence of teaching, learning and education (Biesta, 2010b; Simpson, 2019). Education is not a pill that cures a disease with a controlled dose. Both internal and external critiques on experimental research methods have therefore questioned the belief that randomized controlled studies can provide robust evidence that could guide educational practice and policy.

More realistic evidence

For the providers of educational technology, this eroding foundation of evidence has been a constant challenge. Recent EdTech evidence frameworks (for example, Kucirkova et al., 2025) move beyond learning impacts observed in laboratory settings, but randomized controlled studies are still often considered to provide the best possible evidence. Although some well-known meta-analyses of such research evidence have claimed clear benefits of using AI in education, critical reviews have shown that the presumed effects of AI decrease when known methodological biases are considered. A recent meta-analysis of 66 systematic meta-analyses of the impacts of AI in education notes that the results of existing studies are often based on low-quality primary research (Bond et al., 2024). A closer look at the supposedly high-quality meta-studies reveals huge conceptual variability and often dismal quality hidden behind statistical analyses (Tuomi, 2025).

In policy evaluation studies, the often-superficial methodological rigorousness about 'what works' has therefore increasingly been substituted with the question, 'what works, for whom, and in what context'. In contrast to the empiricist epistemology of experimental studies, realist evaluation (Pawson and Tilley, 1997; Tikly, 2015) builds on the epistemology of critical realism. Instead of trying to find a universal causal connection between an intervention and effect, it tests hypotheses about mechanisms that produce the outcome and contexts that activate these mechanisms. Rather than asking what the average effect of an intervention is, it starts from the assumption that the effect may be different for different intervention subjects. A high average impact can imply negative consequences for some and positive outcomes for others.

Zooming in to the local context where an intervention works is fatal for randomized controlled studies. Randomization was used

1. See https://ies.ed.gov/ncee/wwc

to get rid of context and its 'unobserved covariates' and to get population-level information about a well-defined intervention or chemical compound. When AIED systems personalize and adapt their functionalities according to the needs of the learner, randomized studies become pointless. The medical metaphor breaks down when the dosage, the content of the pill and the disease it is supposed to cure change from learner to learner.

Paradoxically, the belief that AI can improve learning by personalizing it is based on studies that claimed that individual human tutoring can improve learning outcomes by two standard deviations (Bloom, 1984). The developers of intelligent tutoring systems apparently never noted that standard deviations and randomized studies make little sense if the 'treatment' varies from one student to another. The 'gold standard' of evidence is incompatible with the idea that AI can personalize learning. More fundamentally, this contradiction shows that we need to rethink both the ontology and epistemology of AIED.

Towards relational epistemology and ontology

At the cognitive level, adaptive and personalized AI systems are intimately coupled with the individual's thinking and learning processes. This is a fundamental difference between simple automation and AI-based augmentation. When AI systems interact with human thought, the unit of analysis must therefore be the compound of AI-interacting-human. The proper ontology for AIED is relational, and it is not possible to extract causal impacts simply by studying a technical artifact.

Similarly, when teachers use AI systems in their teaching practice, the proper unit of analysis is technology-in-use. In general, the impact of information systems depends not only on technical artifacts and their functionality but also on how they are used and integrated in the ongoing practice (Tuomi, 2004). This means, for example, that it is not possible to study the impact of technology without addressing the capacity of a teacher to use the given system.

Knowledge about 'what works' in AIED is therefore deeply contextual and difficult to generalize. Quantitative knowledge about the effects of AIED that research can produce is structural and contextual. It can inform us about what worked, for whom and in what context.

This may be frustrating to technology providers, as they look for evidence that can convince administrators and policy-makers that their product has an impact that can be scaled and sold across different contexts. When knowledge is not just generalizable 'facts' or 'evidence', but structural and contextual, it cannot be simplified easily, and it does not travel easily in policy contexts (Cowen, 2019). Much tacit knowing is needed to make sense of it. Policy-makers and practitioners cannot just point to this knowledge, they have to learn to make sense of it (Tuomi, 2024). The creation of policy- and practice-relevant evidence shifts from academic laboratories towards the downstream, where knowledge claims are used to support arguments and deliberations about social choices. Evidence becomes part of an argument and discourse where it meets with counter-evidence and different interpretations of the situation.

Policy development as learning

For John Dewey (1991), when routine is disrupted and a problem needs to be solved, we generate hypotheses. Then we test them, first in thought, and then in practice. If the solution works, we have learned something. When a new problem emerges, we can use this learning to create more sophisticated hypotheses. Earlier experience and knowledge, therefore, become instruments

for thought. Knowledge, in this view, is a resource for intelligence, perception and judgement. In this sense, we can say that knowledge makes us more intelligent.

In evidence-informed policy and practice, this means that the process of developing policies and practices, itself, should be viewed as a learning process. Policy-making, as a process, can be made more intelligent by informing it with knowledge from learning theories and by accelerating the process of learning. Using what we know, including results from experimental, qualitative and ethnographic studies that can never be the whole picture, policy can integrate various points of view and different ways to frame the problem, and it must experiment. As all domain-specific epistemologies and knowledge infrastructures generate epistemologies of ignorance (Sullivan and Tuana, 2007), democratic processes are needed to integrate and negotiate knowledge.

Policy experimentation, therefore, does not prove eternal laws of nature or discover universal facts; it creates knowledge about the current situation. It leads to new concepts and new ways of thinking that build on experience. In a similar way as knowledge leads to human development, policy experimentation leads to policy development. It is a form of collective learning.

Such a developmental view on policy highlights a key challenge in current policies that regulate the use of AI in education. Regulation is mostly focused on restricting the uses of technology, protecting from harms and managing risk. In terms of ethics and political theory, regulation, therefore, focuses on 'negative freedoms' that limit the influence of others on individual's actions. Education, in contrast, inherently requires a focus on positive freedoms that make capability expansion and individual development possible. If we take learning and development as fundamental characteristics of what it is to be human, this also has implications on our understanding of democracy (Gould, 1988).

The commodification of knowing

As AI changes the infrastructures of knowing, the development of human agency is becoming increasingly important for education. Knowledge and expertise are becoming commodities. Human agency, understood as a socially and technically embedded capability to participate in the world and change it, replaces the old concept of competence that was defined as a combination of knowledge, skill and attitudes (Tuomi, 2022). This new definition of competence as the expansion of freedom, opportunity and capability is something that cannot be automated. In this sense, it is quite irrelevant to ask whether an AI system can pass a test. Education, as a system that makes competent adults, who can change individual and social realities, was never about passing tests.

AI in education, therefore, should not be viewed only as an instrument to speed up the route to mastery of given content. AI is a qualitatively new kind of technology that interacts with human thinking, communication and development in ways that no earlier technology has done. It is not just a medium or an instrument but an actor. AI systems do not have intentions of their own, but they have agentic powers, as they can make choices and transform material, cognitive and social realities.

Existing evidence frameworks typically start from the assumption that impact should be measured as 'learning outcomes', understood as content mastery. As education drifts from its industrial-era epistemic focus towards what Gert Biesta (2010*a*) calls subjectification and socialization, the impact of AIED should be understood more broadly.

The commodification of knowledge and expertise suggests that we should focus on the development of human socio-technically embedded agency. Instead of searching for robust evidence and causal impacts on 'learning outcomes' that could be used to market technology at scale, we should ask how the social functions of education are changing.

The answer to this question will suggest new designs for AIED but, more importantly, it allows us to see AIED and education differently.

Conclusion

To understand the impact of AIED, we need both new epistemology and new ontology. Experimental research was based on the Newtonian model of linear causality, where agency was replaced by deterministic interactions between inanimate objects. It successfully built on an empiricist epistemology that since Aristotle has separated knowing subjects from known objects. Because of this, it has been natural to think that impact is a property of a tool or an intervention. In this think piece, I have argued that AIED is fundamentally relational. It forms composites with humans, at the same time as expanding human capabilities for thought and action. The impact of AIED, therefore, should be measured as the expansion of these capabilities. Education, itself, is more than mastery of curriculum content, and the social and developmental functions of education cannot be reduced to content mastery. Education policies can be informed by research, but deeper integration and experimentation is needed to support intelligence in the policy process. Instead of adapting to an imagined future, we should take back our agency and make futures we have reason to value. A fundamental assumption underpinning evidence-based education and EdTech evidence frameworks has been that if evidence can be generalized, the interventions and products can be scaled up. Such scaling may be possible, perhaps at a higher level of abstraction than we expected, but it requires that we rethink what is technology and education.

References

Biesta, G. 2010*a*. *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. London, Routledge.

—. 2010*b*. Why 'what works' still won't work: From evidence-based education to value-based education. *Studies in Philosophy and Education*, Vol. 29, No. 5. Berlin, Springer Nature, pp. 491–503.

Bloom, B. S. 1984. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, Vol. 13, No. 6. Washington DC, American Educational Research Association, pp. 4–16. https://doi.org/10.3102/0013189X013006004 (Accessed 14 August 2025.)

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., and Siemens, G. 2024. A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, Vol. 21, No. 1. Berlin, Springer Nature, p. 4. https://doi.org/10.1186/s41239-023-00436-z (Accessed 14 August 2025.)

Cartwright, N. 2019. What is meant by 'rigour' in evidence-based educational policy and what's so good about it? *Educational Research and Evaluation*, Vol. 25, No. 1–2. London, Taylor & Francis, pp. 63–80. https://doi.org/10.1080/13803611.2019.1617990 (Accessed 14 August 2025.)

Cowen, N. 2019. For whom does 'what works' work? The political economy of evidence-based education. *Educational Research and Evaluation*, Vol. 25, No. 1–2. London, Taylor & Francis, pp. 81–98. https://doi.org/10.1080/13803611.2019.1617991 (Accessed 14 August 2025.)

Davies, P. 1999. What is evidence-based education? *British Journal of Educational Studies*, Vol. 47, No. 2. London, Taylor & Francis, pp. 108–121. https://doi.org/10.1111/1467-8527.00106 (Accessed 14 August 2025.)

Deaton, A. and Cartwright, N. 2018. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, Vol. 210. Amsterdam, Elsevier, pp. 2–21. https://doi.org/10.1016/j.socscimed.2017.12.005 (Accessed 14 August 2025.)

Dewey, J. 1991. *How We Think*. New York, Prometheus Books.

Gould, C. C. 1988. *Rethinking Democracy: Freedom and Social Cooperation in Politics, Economy, and Society*. Cambridge, Cambridge University Press.

Hammersley, M. 2007. *Educational Research and Evidence-Based Practice*. Thousand Oaks, SAGE Publications.

Heckman, J. J. 1991. Randomization and social policy evaluation revisited. *NBER Technical Working Paper No. 107*. Cambridge, National Bureau of Economic Research (NBER). https://doi.org/10.3386/t0107 (Accessed 14 August 2025.)

Kucirkova, N., Schewe, O., Campbell, J., Lindroos Cermakova, A. and Pitchford, N. 2025. Developing evidence indicators for evaluating K12 EdTech: Towards a consensus on educational impact. *Humanities and Social Sciences Communications*, Vol. 12, No.1. Berlin, Springer Nature, pp. 947. https://doi.org/10.1057/s41599-025-05330-9 (Accessed 14 August 2025.)

Parra, J. D. and Edwards Jr, D. B. 2024. Challenging the gold standard consensus: Randomized controlled trials (RCTs) and their pitfalls in evidence-based education. *Critical Studies in Education*, Vol. 65, No. 5. London, Taylor & Francis, pp. 513–530. https://doi.org/10.1080/17508487.2024.2314118 (Accessed 15 August 2025.)

Pawson, R. and Tilley, N. 1997. *Realistic Evaluation*. Thousand Oaks, SAGE Publications.

Simpson, A. 2019. The evidential basis of "evidence-based education": An introduction to the special issue. *Educational Research and Evaluation*, Vol. 25, No. 1–2. London, Taylor & Francis, pp. 1–6. https://doi.org/10.1080/13803611.2019.1617979 (Accessed 15 August 2025.)

Sullivan, S. and Tuana, N. 2007. *Race and Epistemologies of Ignorance*. New York, State University of New York Press.

Tikly, L. 2015. What works, for whom, and in what circumstances? Towards a critical realist understanding of learning in international and comparative education. *International Journal of Educational Development*, Vol. 40. Amsterdam, Elsevier, pp. 237–249. https://doi.org/10.1016/j.ijedudev.2014.11.008 (Accessed 15 August 2025.)

Tuomi, I. 2004. Economic productivity in the Knowledge Society: A critical review of productivity theory and the impact of ICTs. *First Monday*, Vol. 9, No. 7. https://doi.org/doi:10.5210/fm.v9i7.1159 (Accessed 15 August 2025.)

—. 2022. Artificial intelligence, 21st century competences, and socio-emotional learning in education: More than high-risk? *European Journal of Education*, Vol. 57, No. 4. Hoboken, John Wiley & Sons, pp. 601–619. https://doi.org/10.1111/ejed.12531 (Accessed 15 August 2025.)

—. 2024. *Fostering Knowledge-Sharing within and among S4P Actors: Mutual Learning Exercise on Bridging the Gap between Science and Policy*. Brussels, Publications Office of the European Union. https://data.europa.eu/doi/10.2777/4531314 (Accessed 15 August 2025.)

—. 2025. What counts as evidence in AI & ED: Towards Science-for-Policy 3.0. *European Journal of Education Policy and Practice*, Vol. 1, No. 1. Amsterdam, Amsterdam University Press. Unpublished *(*Forthcoming).

9. Conclusion

This anthology reveals how the integration of AI in education poses profound philosophical, pedagogical, ethical and political questions. As this global AI ecosystem evolves and becomes increasingly ubiquitous, UNESCO and its partners have a shared responsibility to lead the global discourse towards an equity- and justice-centred agenda. The volume highlights three areas in which UNESCO will continue to convene and lead a global commons for dialog and action particularly in areas on AI futures, policy and practice innovation, and experimentation.

As guardian of ethical, equitable human-centred AI in education

This collection has underlined how AI futures in education require priority attention to the ways that entrenched inequalities of power, access and opportunity are being reshaped. As these systems evolve, they risk amplifying existing gender, class, language, geography and digital access disparities. This moment calls for courageous leadership that goes beyond technical fixes to fundamentally reimagine AI's role in education. Building on UNESCO's *Recommendation on the ethics of artificial intelligence*, its *Guidance for generative AI in education and research*, and twin *AI competency framework for teachers* and *AI competency framework for students*, UNESCO will continue to work with academic, research and policy communities towards designing systems that cultivate transformative ethical human agency, equity and social justice in digital learning spaces and create genuine possibilities for inclusion across all dimensions of difference. As we explore such creative possibilities together, UNESCO and its partners are also duty-bound to counterbalance the unchecked ascendancy of technocratic and corporate influence of AI in education, combat algorithmic discrimination and AI surveillance, and advocate for enforceable rights-based global standards in education.

As thought leader in reimagining curriculum and pedagogy

Contributions in this volume reveal how AI is already reshaping curricula, pedagogy, assessment and knowledge creation. These analyses underscore the need for leadership on emerging curriculum and pedagogical shifts and exploring strategies that prioritize critical thinking, metacognition, pedagogies of care and ethical reasoning, in ways that challenge reductive logics on rote learning, standardized assessment, cognitive offloading and complex risks associated with anthropomorphizing AI companionship. Here, attention to emerging possibilities for new pedagogies on inclusive cyber-social learning becomes imperative. UNESCO and its partners will continue to create open shared spaces for exploring such pedagogical possibilities.

As a platform for engaging pluralistic and contested dialogues

The philosophical tensions illuminated in this volume direct us towards convening deliberative forums, where diverse communities, disciplines and cultures engage with the contestations about AI's purpose, its unfolding influence and consequences in education, not just its implementation. Such deliberative forums need to amplify the voices and perspectives of marginalized and excluded communities, including indigenous, gendered, Global South and disability perspectives. The path forward demands more than cautious adaptation. It requires radical rethinking of how we design, govern and integrate co-constitutive human-machine systems in education amidst growing uncertainty, unpredictability and volatility. The nascent ideas on creating inclusive, socially just futures seeded in this volume will require sustained ethical commitment and political will at all levels of decision-making. UNESCO stands ready to take up this challenge.

References

UNESCO. 2022. *Recommendation on the ethics of artificial intelligence*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137 (Accessed 16 August 2025.)

—. 2023. *Guidance for generative AI in education and research*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693 (Accessed 16 August 2025.)

—. 2024. *AI competency framework for students*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391105 (Accessed 17 August 2025.)

—. 2024. *AI competency framework for teachers*. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391104 (Accessed 16 August 2025.)

About the authors

Carla Aerts, Futures of Education and AIEdTech Catalyser and Advisor

Carla is a distinguished thought leader in Education Futures and AIEd, bringing transformative vision and execution to the global education landscape. Following leadership roles at UCL Institute of Education and Cambridge University Press, she founded Tmrw Institute, uniting policy-makers, educators and innovators. Her expertise spans engagement with WEF's AI Network, Brookings Institution's AI Taskforce and Cambridge University's Digital Futures Education Initiative's board. Having mentored 100+ startups and led the Next Billion EdTech Prize, Carla specializes in digital transformation and innovation strategy. As Big Education Trustee and independent consultant, she guides C-suite executives and international organizations, while her speaking and writing shape the convergence of technology, learning and societal advancement.

Báyò Akómoláfé, Hubert Humphrey Distinguished Professor of American Studies (Macalester College); W.E.B DuBois Scholar, Schumacher Center for a New Economics; Inaugural Scholar in Residence, Aspen Global Leadership Network

Dr Báyò Akómoláfé is an internationally recognized posthumanist thinker, author and public intellectual rooted with the Yoruba people. He holds a Ph.D. in Clinical Psychology from Covenant University, Nigeria, and serves as the Hubert Humphrey Distinguished Professor of American Studies at Macalester College, Minnesota.

As founder of The Emergence Network, a global inquiry project exploring posthumanist parapolitics, Dr Akómoláfé develops innovative concepts that reframe human agency and responsibility in the Anthropocene. His widely celebrated expression 'the times are urgent, let us slow down' anchors his philosophical explorations of ontofugitivity, postactivism and more-than-human worlding.

Author of These Wilds Beyond our Fences and co-author of We Will Tell our Own Story, he holds prestigious fellowships including Global Senior Fellow at UC Berkeley's Othering and Belonging Institute and W.E.B. Du Bois Scholar in Residence at the Schumacher Centre. He is a Club of Rome member and recipient of multiple honours, including Portland's Key to the City.

A sought-after keynote speaker at institutions like Harvard, MIT and Duke University, his work transcends academia, inspiring artistic interpretations across multiple media and featured in award-winning documentaries.

Joaquín T. Argüello de Jesús, Licensed Independent/Clinical Social Worker (LISW/LCSW), Doctoral Candidate, Language Literacy & Sociocultural Studies Department/College of Education and Human Sciences, University of New Mexico

DominiXicanoRriqueño decolonial, antiracist, bilingual community, clinical, School Social Worker raised on the 'Manito/a/x Trail;' Board affiliations: Arroyo Hondo Arriba Community Land Grant, Compostela Community & Family Cultural Institute; Language Literacy & Sociocultural Studies, University of New Mexico graduate research assistant: Sociology: Census evolution of race and ethnicity federal guidelines; Bilingual math (Algebra) learning through Indigenous Epistemologies: 'wonderment', animated, relationality, materiality and Translanguaging corriente of 1st graders; Current research: critical auto-ethnographic traditional merced/acequia rematriation survivance practices.

Payal Arora, Chair of Inclusive AI Cultures, Dept. of Media & Culture Studies, Utrecht University

Payal Arora is a Professor of Inclusive AI Cultures at Utrecht University and co-founder of Inclusive AI Lab. A digital anthropologist and award-winning author, her books include *The Next Billion Users* with Harvard Press and *From Pessimism to Promise* with MIT Press. Forbes named her the 'next billion champion' and the 'right kind of person to reform tech.' She is listed in the 100 Brilliant Women in AI Ethics 2025 and won the 2025 Women in AI Benelux Award. Her work has been featured in 250+ media outlets including *Financial Times*, *Wired*, *The Economist* and *Tech Crunch*, and has keynoted at 100+ events such as TEDx, COP26 and WEF.

Isak Nti Asare, Faculty Co-Director of the Cybersecurity and Global Policy Program at Indiana University-Bloomington

Isak Nti Asare is a faculty member at Indiana University's Hamilton Lugar School of Global and International Studies, where he serves as Assistant Dean for Undergraduate Education and Student Affairs, Co-Director of the Cybersecurity and Global Policy Program, and Executive Director of the Indiana University Cybersecurity Clinic. He has advised national and local governments on AI strategy and digital transformation. He holds degrees from Indiana University and the University of Oxford and is a Ph.D. candidate at the UCL Institute of Education.

Emily M. Bender, Professor, Department of Linguistics and Director, Computational Linguistics Laboratory, University of Washington

Emily M. Bender is a Professor of Linguistics and an Adjunct Professor in the School of Computer Science and the Information School at the University of Washington, where she

has been on the faculty since 2003. Her research interests include multilingual grammar engineering, computational semantics and the societal impacts of language technology.

In her public scholarship, including her recent book *The AI Con*, Bender brings linguistic insights to lay audiences to cut through the hype about 'AI' and facilitate understanding of the actual functionality of the systems being sold under that name. In 2022 she was elected Fellow of the American Association for the Advancement of Science (AAAS) and in September 2023 she was included in the first-ever TIME100AI list highlighting 100 individuals advancing major conversations about how AI is reshaping the world. She is also Past President (2024) of the Association for Computational Linguistics.

Kiran Bhatia, Research Lead, Localizing Responsible AI cluster, Inclusive AI Lab, and Research Affiliate, Department of Media and Culture Studies, Utrecht University

Kiran Vinod Bhatia is a digital anthropologist, user experience researcher and Responsible AI lead at Utrecht University's Inclusive AI Lab. With over a decade of fieldwork across the Global South, her research explores inclusive tech design, AI safety and digital well-being, centring the experiences of women, young people and marginalized communities.

Abeba Birhane, Founder, AI Accountability Lab (AIAL) and Research Fellow, School of Computer Science and Statistics, Trinity College Dublin

Abeba Birhane founded and leads the TCD AI Accountability Lab (AIAL). She received her Ph.D. from UCD in 2022 and is currently a Research Fellow at the School of Computer Science and Statistics in Trinity College Dublin. Her research focuses on AI accountability, with a particular focus on audits of AI models and training datasets – work for which she was featured in Wired UK and TIME on the TIME100 Most Influential People in AI list in 2023. Ms Birhane also served on the United Nations Secretary-General's AI Advisory Body and currently serves at the AI Advisory Council in Ireland.

Ching Sing Chai, Professor, Department of Curriculum and Instruction, Faculty of Education, The Chinese University of Hong Kong

Ching Sing Chai is a professor at the Chinese University of Hong Kong. He is currently the Associate Dean of Higher Education. His research interests are in the areas of Technological Pedagogical Content Knowledge (TPACK), Artificial Intelligence in education, teachers' beliefs, design thinking and students' learning with ICT.

Thomas K.F. Chiu, Professor, Department of curriculum and instruction, The Chinese University of Hong Kong

Thomas Chiu is an Associate Professor of AI and STEM Education at The Chinese University of Hong Kong. A globally recognized scholar, he is named a Stanford University Top 2% Most Cited Scientist. He provides significant editorial leadership as Editor-in-Chief of *Interactive Learning Environments* and Associate Editor for three leading journals. His expertise is sought internationally, reflected in visiting scholar appointments across five continents. He drives innovation through the development of professional development initiatives and curricula for AI education, and AI applications for nurturing core skills. He regularly shares his expertise as a keynote and invited speaker at global academic and professional forums.

Bill Cope, Professor, Department of Education Policy, Organization and Leadership, University of Illinois at Urbana-Champaign

Bill Cope is a Professor in the Department of Education Policy, Organization & Leadership, University of Illinois, Urbana-Champaign. His and Mary Kalantzis' recent research has focused on the development of digital writing and assessment technologies, with the support of a number of major grants from the US Department of Education, the Bill and Melinda Gates Foundation and the National Science Foundation. He has co-authored or co-edited with Mary Kalantzis: *New Learning: Elements of a Science of Education*, Cambridge University Press, 2008 (3rd edition, 2022); *Literacies*, Cambridge University Press 2012 (2nd edition, 2016); and the two volume grammar of multimodal meaning: *Making Sense* and *Adding Sense*, Cambridge University Press, 2020.

Markus Deimann, Managing Director, ORCA.nrw / Ruhr-University Bochum

Markus Deimann is a German educational scientist who specializes in digital education, Open Educational Resources (OER) and the political dimensions of technology in higher education. He has held various academic leadership positions, including CEO of ORCA.nrw, and has published extensively on technology, ideology and educational policy. His research connects critical theory with practical innovation, exploring how imaginaries shape the future of universities. He also engages in public discourse through lectures, blogs and academic collaborations.

Robert Farrow, Senior Research Fellow, Institute of Educational Technology, The Open University

Dr Robert Farrow is Senior Research Fellow at the Institute of Educational Technology, The Open University (United Kingdom) where he leads the research programme Learning

in an Open World. He is Co-Director of the Global OER Graduate Network and Co-Editor of the Journal of Interactive Media in Education. His research focuses on technology enhanced learning and the design, implementation and evaluation of socio-technical learning systems, often with a focus on open practices.

Kalervo N. Gulson, Professor, Sydney School of Education and Social Work, Faculty of Arts and Social Sciences, The University of Sydney

Kalervo N. Gulson is a Professor of Education Policy at the University of Sydney. His research is located across social, political and cultural geography, education policy studies, and science and technology studies. Kalervo's work with the Education Futures Studio and the Centre for AI, Trust and Governance at the University of Sydney includes cutting-edge research, both in Australia and internationally, to understand how AI and EdTech are transforming education policy and implementation, and reimagining human-machine interactions and their implications for education governance. Kalervo develops participatory methods and works with educators, leaders, policy-makers, students and parents to build new ways to understand and evaluate how, and if, to use AI and EdTech. He also, through collaborative and commissioned work with government, NGOs and civil society actors, generates policy recommendations to guide the use of AI and EdTech towards more equitable education futures.

Andreas Horn, Head of AIOps, IBM

Andreas Horn leads IBM's AIOps service portfolio for EMEA, driving AI-powered transformation across enterprises. He specializes in strategic AI deployment for IT Operations, ensuring businesses leverage AI to enhance resilience, efficiency and innovation in this new era of AI. Beyond IBM, Andreas is a university lecturer, educating the next generation on practical AI applications that boost productivity and streamline workflows. His mission: turn AI from hype into real-world impact.

Mary Kalantzis, Professor, Department of Education Policy, Organization and Leadership, University of Illinois at Urbana-Champaign

Mary Kalantzis was from 2006 to 2016 Dean of the College of Education at the University of Illinois, Urbana-Champaign. Before this, she was Dean of the Faculty of Education, Language and Community Services at RMIT University, Melbourne, Australia, and President of the Australian Council of Deans of Education. She has co-authored or co-edited with Bill Cope: *New Learning: Elements of a Science of Education*, Cambridge University Press, 2008 (3rd edition, 2022); *Literacies*, Cambridge University Press 2012 (2nd edition, 2016); and the two volume

grammar of multimodal meaning: *Making Sense* and *Adding Sense*, Cambridge University Press, 2020.

Arafeh Karimi, Fractional CPO at Affexy

Arafeh Karimi is a global education and AI strategist architecting pedagogy-first, sovereign and relational AI ecosystems. Her Compassion by Design approach turns ethics into implementable policy and classroom practice, shaped through co-design with educators, communities and leaders. Drawing on extensive fieldwork across the Global South from refugee camps to rural schools, she builds inclusive, human-centred solutions that embed epistemic justice, care and neurodivergence-affirming practices. Arafeh's work weaves AI into the relationships, cultures and learning ecologies that enable communities to flourish.

Vukosi Marivate, ABSA Chair of Data Science, Professor, Department of Computer Science, University of Pretoria

Prof. Vukosi Marivate is Chair of Data Science and Professor of Computer Science at the University of Pretoria, where he leads the Data Science for Social Impact group. His research focuses on Machine Learning (ML), Artificial Intelligence (AI), and Natural Language Processing (NLP), particularly for African and other low-resource languages. He co-founded Lelapa AI, the Masakhane Research Foundation, and the Deep Learning Indaba. His work spans social challenges in science, energy, public safety, and utilities, aiming to create AI for Africans by Africans.

Baphumelele Masikisiki, Ph.D. candidate in Computer Science, University of Pretoria

Baphumelele Masikisiki is a doctoral candidate at the University of Pretoria specializing in Natural Language Processing (NLP), with a focus on advancing Artificial Intelligence applications in education. Her research explores how Large Language Models (LLMs), such as Llama, can improve Automated Essay Scoring (AES) by enhancing accuracy, robustness, and reliability. She applies transfer learning, in-context learning, and Mixture of Experts (MoE) techniques to improve scoring performance, while also incorporating machine learning interpretability methods to promote transparency and educator trust.

Beyond accuracy, her work emphasizes robustness to linguistic variation, mitigation of length bias, and the generation of rubric-aligned, pedagogically meaningful justifications. This research is among the first to investigate AES for essays written by African medical students, addressing both linguistic diversity and the underrepresentation of African educational contexts in NLP. Ultimately, her goal is to develop AES systems

that are accurate, fair, explainable, and context-aware, fostering meaningful feedback and advancing educational equity across diverse learning environments.

Mike Perkins, Associate Professor and Head, Centre for Research & Innovation, British University Vietnam

Assoc. Prof Dr Mike Perkins heads the Centre for Research & Innovation at British University Vietnam. Mike is one of the authors of the AI Assessment Scale, which has been adopted globally across schools and universities. His research focuses on GenAI's impact on education, and has explored various areas within this field. This has included AI text detectors, attitudes to AI technologies and the ethical integration of AI in assessments through the AI Assessment Scale. His work bridges technology, education and academic integrity.

Kaśka Porayska-Pomsta, Professor of AI in Education, University College London, UCL Institute of Education

Kaśka Porayska-Pomsta is full Professor of AI in Education at the University College London, UCL Faculty of Education and Society and the current Director of the UCL Knowledge Lab. She holds a Joint Honours Master's in Linguistics and Artificial Intelligence and a PhD in AI from the University of Edinburgh. Her research bridges engineering and social sciences, exploring how AI can support learning, development, and communication, while balancing the needs of learners and practitioners with the design and deployment constraints of AI in education. Diversity, inclusion, and trust in AI are central to her work. She serves on the management committee of the Bloomsbury Centre for Educational Neuroscience, the executive committee of the International Society for AI in Education, the advisory board of UCL EdTech Labs, the expert committee for the UCL-AWS Centre for Data Innovation, and the jury for the UNESCO King Hamad Bin Isa Al-Khalifa Prize for ICT in Education, bringing interdisciplinary expertise spanning AI engineering, learning sciences, cognitive research, and ethics.

Paul Prinsloo, Professor Emeritus, University of South Africa (Unisa)

Paul Prinsloo is a Professor Extraordinaire at the University of South Africa (Unisa). He is a also a visiting professor at the National Open University of Nigeria (NOUN), a member of the Center for Open Education Research (COER) at the Carl von Ossietzky University of Oldenburg (Germany), a Senior Fellow of the European Distance and E-Learning Network (EDEN), a member of the the International Advisory Board (IAB) at the Open University Malaysia (OUM) and a Global Fellow at the Centre For Digital Education Futures (CENDEF)

at OUM. His research explores various issues in AI in education, digital identities as well as distributed and digital learning.

Mary Rice, Associate Professor of Literacy, Language Literacy & Sociocultural Studies Department/College of Education and Human Sciences, University of New Mexico (UNM)

Mary Rice taught English language arts and English as a second language to adolescents for 10 years in public schools. Her current teaching and research consider how roles and responsibilities emerge between people, things and wider environments when new technologies enter schools. She also considers which learners access which technologies in schools and how learners are framed with and by technologies. Mary is senior personnel for workforce development with UNM's photonic quantum computing virtual laboratory.

Jasper Roe, Assistant Professor in Digital Literacies and Pedagogies, Durham University School of Education

Dr Jasper Roe is a specialist in English Language Education, Higher Education, academic integrity and educational technology. Jasper is an Assistant Professor in Digital Literacies and Pedagogies at Durham University, United Kingdom, having previously held senior leadership roles including Head of Pre-University and Head of the Language School in Viet Nam and Singapore respectively. Since the early 2020s his research work has focused on Artificial Intelligence (AI) and education. He is a co-author of the AI Assessment Scale (AIAS).

Akash Kr. Saini, Ph.D. candidate, University of Illinois at Urbana-Champaign

Akash Kr. Saini is a Teaching and Research Assistant in the Learning Design & Leadership program at the College of Education, University of Illinois Urbana-Champaign. His scholarship bridges generative AI, assessment, higher education, design education, and education policy: he leads empirical studies on AI-generated feedback that boost learning outcomes and academic writing. He is currently experimenting with Retrieval-Augmented Generation systems that deliver ubiquitous and data-rich feedback. A portfolio of peer-reviewed articles, book chapters, and conference papers underpins his exploration of AI-powered feedback systems across university and K-12 classrooms.

George Siemens, Co-founder, Chief Scientist and Architect of Matter & Space, Professor and Director, Centre for Change and Complexity in Learning for UniSA Education Futures, University of South Australia

Professor George Siemens researches how human and artificial cognition intersect in knowledge processes. He is co-founder, Chief Scientist

and Architect of Matter & Space - focused on the systemic impact of AI on learning and wellness. He is the founding Director and Professor of the Center for Change and Complexity in Learning (C3L) at University of South Australia. He has delivered keynote addresses in more than 40. He has served as PI or Co-PI on grants funded by NSF, SSHRC (Canada), OLT (Australia), Intel, Boeing, Bill & Melinda Gates Foundation, and the Soros Foundation. He has received honorary doctorates from Universidad de San Martín de Porres and Fraser Valley University for his pioneering work.

Sam Sellar, Dean of Research for UniSA Education Futures, University of South Australia

Sam Sellar is Dean of Research (Education Futures) and Professor of Education Policy at the University of South Australia. Sam's research focuses on education policy, large-scale assessments and the datafication of education. He is currently co-investigator for an ESRC project investigating digital platforms in higher education (led by Janja Komljenovic, Lancaster University). Sam has published more than 70 books, book chapters and journal articles, and he is Lead Editor of *Discourse: Studies in the Cultural Politics of Education*. His forthcoming book is *Algorithms of Education: How datafication and artificial intelligence shape policy*, co-authored with Kalervo N. Gulson and P. Taylor Webb.

Bing Song, Senior Vice President, Berggruen Institute, and Director, Berggruen Institute China Center, China

Bing Song is Senior Vice President of the Berggruen Institute and Director of the Institute's China Center.

Prior to joining the Berggruen Institute, she was a senior executive with Goldman Sachs China for over a decade, and prior to Goldman, an experienced capital markets lawyer for many years. Earlier in her career, she undertook academic and policy research and published in the areas of administrative law, competition law and comparative procedural laws.

In 2024, she co-edited *Gongsheng Across Contexts – A Philosophy of Co-Becoming*, which explores philosophical foundations of *gongsheng,* a conception of the world as consisting of mutually embedding, co-existing and co-generating entities. Her edited volume *Wisdom and Intelligence – Artificial Intelligence Meets Chinese Philosophers*, published in 2021, marked the first systematic endeavour by Chinese philosophers to address challenges and opportunities posed by artificial intelligence systems. Currently, Bing is working on projects relating to consciousness, intelligence and Eastern philosophies.

Ilkka Tuomi, Chief Scientist at Meaning Processing Ltd.

Ilkka Tuomi is Chief Scientist at Meaning Processing Ltd., an independent research institute located in Helsinki, Finland. He is one of the pioneers of AI in Finland and author of several books, articles, and policy-oriented reports on AI and education, innovation theory, knowledge management, futures studies, and evidence-informed policy. He was senior thematic expert for the EC expert group that developed ethical guidelines for educators on the use of AI, and expert for UNESCO AI competence frameworks for teachers and students. At present he develops the Committee of Ministers recommendation on AI literacy for the Council of Europe and its AI&ED policy toolbox.

Yuchen Wang, Chancellor's Fellow, Strathclyde Institute of Education, University of Strathclyde

Yuchen Wang is a Chancellor's Fellow at the Strathclyde Institute of Education, University of Strathclyde. Her research interests include inclusive education, children and young people's participation, sustainable development, and technology. She has led projects funded by the Economic and Social Research Council, the Natural Environment Research Council, and the British Council. She was a co-editor for the award-winning book 'Artificial Intelligence and Inclusive Education: Speculative Futures and Emerging Practices', and a key contributor to the latest edition of Scotland's 'National Framework for Inclusion'.

Marloes Williams – van Elswijk, Project Manager, Kentalis International Foundation

Marloes Williams – van Elswijk is a project manager at Kentalis International Foundation where she develops and oversees projects and programmes aimed at improving equitable access to quality education for learners who are Deaf or hard of hearing in low-resource contexts. Her work focuses on designing evidence-informed programmes, fostering cross-sector collaborations, and supporting capacity-building initiatives through blended learning that promote inclusive educational practices.

Jiun-Yu Wu, Professor, Department of Teaching and Learning, Southern Methodist University

Jiun-Yu Wu is a Professor in the Department of Teaching and Learning at Southern Methodist University. His research bridges learning science, AI in education and data science, with a focus on multimodal learning analytics, human-AI symbiotic learning and technology-enhanced pedagogy. A globally recognized scholar, he has been named among the World's Top 2% Scientists by Stanford University.

(2021–2024) and serves as Editor for *Computers & Education and Associate Editor of Educational Psychology*. He has led major international collaborations across Asia and North America.

Nombuyiselo Caroline Zondi,

PhD candidate in Early Childhood Education, University of Pretoria

Nombuyiselo Caroline Zondi is a final-year Ph.D. candidate at the University of Pretoria, specializing in Early Childhood Education. Her research explores why many children struggle to read in their mother tongue, focusing on language policy, pedagogy, professional development and material alignment. A published children's author in Sepedi, isiZulu and Sesotho, she advocates for culturally grounded, mother tongue literacy. Inspired by the work of UNESCO and others in various literacy programmes across the globe, she also collaborates on AI, storybook and animation projects-believing teachers must not only teach, but also design the materials that shape learning. She welcomes partnerships that reimagine literacy for African children.

unesco
United Nations
Educational, Scientific
and Cultural Organization

AI and the future of education

Disruptions, dilemmas and directions

As AI systems increasingly influence our educational decisions, priorities and relationships, urgent collective reflection is needed on how we adapt and centre equity, care and human dignity. This UNESCO volume brings together leading thinkers, educators and policy-makers who engage with the divergent narratives on AI futures in education. It offers nuanced insights on wide-ranging topics, from coded inequalities to new pedagogical architectures. The anthology contributes to shaping a global commons for dialogue and action towards inclusive, equitable and ethical AI futures in education.

9 789231 007842

4 QUALITY EDUCATION

Sustainable Development Goals