Interaction Design and Architecture(s) Journal - IxD&A, N.64, 2025, pp. 28 - 62
DOI: 10.55612/s-5002-064-001sp

AI and Education Beyond 2030: Grand Challenges

Beverly Woolf, Danielle Allessio, Ivon Arroyo, Sai Gattupalli, Boming Zhang

University of Massachusetts, Amherst, MA, 01002, USA
bev@cs.umass.edu, danielle.allessio@gmail.com, ivon@cs.umass.edu,
sgattupalli@umass.edu, bomingzhang@umass.edu

Abstract. Artificial intelligence (AI) has made large changes in major industries and disrupted or reorganized many disciplines. As a result, traditional educational practices need to be reexamined to enable learners to develop new skills, to manage and utilize new technologies and to increase productivity in a rapidly changing world. A more flexible educational system is required to enable life-long and life-wide upskilling and reskilling. This article provides four grand challenges for AI and education to optimize digital learning, online resources and virtual classrooms. It suggests several problems to address, visions to spur the field forward and strategies that will make teaching and learning more effective. The article also considers ethical use of technology, decreased jobs in some sectors and the possibility that AI will exacerbate an existing deficit in diversity and equity among students.

Keywords: Artificial Intelligence, Education Technology, Personalized Learning, Ethical AI, Intelligent Tutoring Systems, Workforce Development

1 Introduction

Artificial intelligence (AI) has been integrated into almost every industry, e.g., healthcare, government, communication, and finance and is present in everyday interactions through smart phones, cars, homes and more, a testament to the immense potential of the technology and its ability to reshape nearly every sector. AI is critical in travel (directional way-finding), daily life (legal issues and civic responsibilities), and health care (medical and pharmacological information, self-care strategies, distance medicine). Yet AI-based education remains not well utilized; despite the immense theoretical potential along with a large amount of monetary investment, intelligent learning technologies have not yet delivered impactful results in real-world education, nor are deliverables institutionalized within traditional teaching environments.

One long-term goal for AI and education is to contribute to changing the current transmissive, regimented and school-factory model of education based on student age, location, and classroom, and to move towards a more social, diverse and equitable model of instruction. AI can't change the transmissive model itself, because the learning paradigm needs to change from "learning by knowing" to "learning by doing" [37]. The gap between knowing and doing can be a problem in education, management, science, and health. It can occur when someone has knowledge of something but doesn't apply it in real life. For example, one way to bridge the gap from knowing to doing is through experiential learning or encouraging students to learn through experiences that help them retain information. Traditional models of education (fixed classrooms, one-to-many lectures) struggle to accommodate diverse learning styles and preferences. AI can build in education flexibility, accessibility and personalized experiences. Personalized learning systems adapt instruction to each student's individual learning style and pace, complimenting learning and freeing teachers to spend more time on instruction and one-to-one counseling. Personalization by itself is not enough to change the learning paradigm, as human support is needed to monitor the development of competencies.

A Brief Overview of AI. Many grand challenges in education have been identified and numerous artificial intelligence (AI) technologies have been employed to tackle each challenge (Table 1). AI is a set of technologies designed to enable computers to perform tasks that mimic human intelligence, including to see and understand (process visual and written information), analyze data (organize and evaluate large amounts of data to solve problems), make recommendations (suggestions based on data analysis), learn and reason (use knowledge to solve problems), and act (achieve goals through actions). Generative AI tools, discussed in this article, use technologies to create original text, images, video and other content.

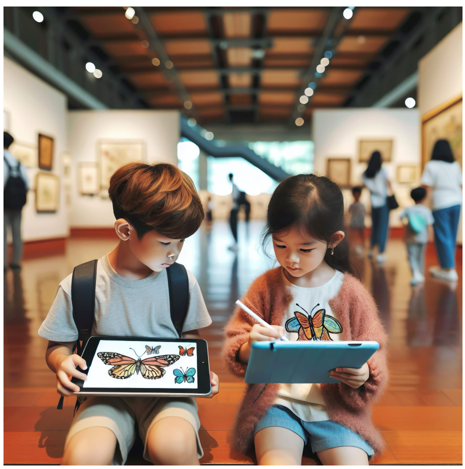

Fig. 1. Elementary school children draw distinct butterfly pictures on digital devices during a field trip to a museum. Individual graphics are stored in a class database.

These tools are built on machine learning (algorithms that enable systems to learn from data) and deep learning (neural networks to mimic human intelligence). Other AI subfields mentioned in this article include natural language processing (speech and text recognition, analysis, and generation), computer vision (used for facial, emotional, and gesture recognition), automated speech recognition (voice recognition and speech production devices like Siri and Alexa), and decision management (automate decision making in organizations to improve operational decisions). AI has

Table 1. Characteristics of AI and Education. Educational challenges (Column 1), the educational impact of solving each challenge (Column 2) and AI technologies used to tackle each challenge (Column 3). The AI technologies listed in Column 3 are not exhaustive; many others might be used. Legend: ML (machine learning), LLM (large learning models).

Educational Challenges	**Teaching and Learning Features** (How is teaching/learning impacted)?	**AI Technology for the Challenge** (Which AI technology is used)?
Pedagogical Innovations		
Personalize teaching	Recognize /infer student knowledge, behavior and emotion; Adapt responses, questions and content for each learner.	User models, ML, deep learning, LLM
Foster critical thinking	Students clarify their own thinking, analyze facts, evidence, observations, and arguments to form judgements; Students are in control of their learning.	Natural language processing, generative AI, deep learning, LLM,
Multimodal learning	Use multiple senses and teaching styles, e.g., kinesthetic, auditory, and visual to reinforce knowledge comprehension.	Computer vision, embodied cognition.
Address the Digital Divide		
Equitable distribution of education	Provide opportunities, transparency, accountability and support for underrepresented communities; Strive for social justive.	Big data, ML, deep learning, optimization, stealth assessment
Diverse, ethical inclusive education	Respect, value, and incorporate learners' diverse cultural backgrounds, experiences, and perspectives; Create cultural sensitivity and inclusivity.	ML, deep learning, LLM, data security, data privacy, avoid bias in AI databases and training, avoid exploitive labor practices.
Collaborative problem solving	Improve academic achievement, social and emotional development, and peer acceptance; Recognize that benefits extend over generations.	ML, deep learning, LLM, interfaces for critical thinking, creativity and collaboration.
Global Learning Communities		
Benefit more students	Provide access to high quality educational tools, resources (problems, answers, hints) produced at scale; Improve professional learning about using/integrating AI tools into schools.	Intelligent tutoring systems, adaptable systems, open sources AI systems.
Lifelong learning	Develop career platforms in which agents motivate users; make age, economic, and cultural considerations; Teach within practical/real-life contexts, promote self-efficacy.	AI-based agents as facilitators integrated into learning environments, virtual learning companions
Data-Driven Decision Making		
Predictive analysis	Data-directed decision-making to identify patterns and trends; Provide immediate responses, personalized learning paths; Guide institutional learning (budgeting, transparency).	ML, deep learning, big data, Computer vision, facial expression recognition, gesture recognition.
Model teaching and learning	A proxy of real learners (many backgrounds and different learning context); Stakeholders practice strategies with the model, identify pedagogical practices, test the quality and difficulty of new learning content and predict learner deficiencies.	Planning, user models, LLM, generative AI, learning science, cognitive science

many applications, including: self-driving cars, speech and facial recognition, digital personal assistants, and virtual customer service, and recommendation engines.

Perhaps an apt metaphor for current AI and education is a bonfire ready to ignite--along with the promised inordinate benefits come the dangers of unintended consequences and misuses [65]; along with great power comes great responsibility. Limitations exist for using AI in education; some AI systems are not consistently ethical (lack of privacy, diversity and inclusivity); some are of mixed quality (truthfulness of chatbots in classrooms can't be guaranteed); many are created in the developed world and thus don't represent the majority of global stakeholders; often these systeems do not provide good predictive power (of student grades); and are not effective with issues that require a user's subjective judgments (personal taste, ethics or moral dilemmas). This article discusses the consequences that the introduction of AI in education may have on individuals, groups and the global economy.

Computers, with or without AI, will not replace teachers; they can't provide face to face contact and don't provide the wealth of experience that a teacher has [40]. They are not as interactive as human teachers; they don't feel emotion, and can't care for children, discipline students nor provide safety for vulnerable people. However, they can enhance and enrich education and be used as tools in the classroom, as described in this article [40].

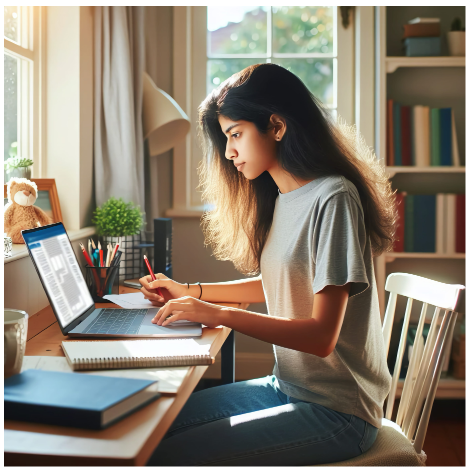

Fig. 2. Student completes assignments while a digital tutor selects problems and questions based on individual learning needs.

School leaders who put their heads in the sand and ignore the potential impact of AI on education are committing a serious disservice to students when it comes to job market competitiveness [52, 66]. Many challenges remain and we stand on the brink of a new era in a young field with an upward trajectory. Yet, new policies need to be put into place to safely leverage the vast potential of AI and education for the benefit of humanity, moving the field from prescriptive algorithms to a human-centric and impactful ecology [16]. Computers with AI will also make big changes in the workplace; countries with more advanced economies will face greater risks as about 60 percent of jobs in such countries will be impacted by AI [30, 37]. Workers who can harness AI will experience increased productivity and wages—and those who can't, will fall behind [37]. Roughly half the exposed jobs may benefit from integrating AI into existing jobs to enhance productivity. For the other half, AI applications will take over some key tasks currently performed by humans, which could lower labor demand, leading to lower wages and reduced hiring. In the most extreme cases, jobs will disappear. It is estimated that over 30% of current jobs require some type of AI skill and this number will increase sharply by 2030 [37].

This article identifies four (4) grand challenges for AI and education and considers affordances of new technologies, e.g., massive data sets might follow students during their lifetimes and make predictions about student learning and affect. Risks of AI and education include exacerbating the inequalities of students or having biases (e.g., referring to people in certain professions who look or talk like the dominant populations). This article contributes to ongoing discussions (involving researchers, teachers, parents, and other stakeholders) to redesign education and learning through people centered learning ecosystems [9, 66]. AI will help these community movements to assemble evidence of student learning and teacher leadership in ways that were heretofore impossible. It can support relationship-centered and inclusive education for marginalized students. This article suggests a framework for moving forward desirable AI affordances while also considering AI liabilities.

2 Pedagogical Innovations

A first grand challenge for AI and education is to provide pedagogical innovations for learners, e.g., as students work on computers, an AI partner or "more knowledgeable other" becomes an assistant teacher for each student, supporting students emotionally [5, 8, 10, 24, 25]. The integration of AI in education operates through fundamentally different technological approaches, each serving distinct purposes in the educational ecosystem [17, 66]. Understanding these differences is crucial for effective implementation and evaluation of improved learning.

2.1 Technologies for Institutional and Personalized Responses

A first feature of pedagogical innovation involves supporting institutional teaching, e.g., specific materials for organizational-level implementation [50]. Learning Management Systems (LMSs) typically utilize database management systems, scheduling algorithms, and analytics tools operating at the institutional level [64]. These systems process structured data about curriculum, resources, and administrative tasks using traditional database technologies and rule-based algorithms) [34]. Resource allocation systems employ optimization algorithms to schedule classes and assign rooms, while assessment tracking platforms use statistical analysis tools to monitor institutional-level performance metrics [99].

In contrast to institutional systems, technologies providing personalized feedback employ fundamentally different approaches that operate at the individual student level

[96]. These systems require sophisticated real-time processing capabilities and adaptive algorithms to respond to individual learner needs. Research indicates their effectiveness varies significantly based on implementation context and student characteristics [51].

Natural Language Processing (NLP) Systems: Advanced NLP technologies form the foundation of personalized feedback systems, employing large language models to analyze and respond to student work in real-time [106]. Unlike institutional systems' simple pattern matching, these tools process complex linguistic structures to understand student responses at both syntactic and semantic levels. Studies suggest that modern NLP systems can provide context-aware feedback by analyzing not just correctness, but also the reasoning process evident in student responses [48]. However, research indicates the effectiveness of such systems varies across different subject domains and student populations, requiring careful calibration and ongoing refinement [64].

Adaptive Learning Algorithms: These systems employ dynamic machine learning models that continuously adjust to individual student learning patterns. Unlike institutional systems' fixed rule-based approaches, adaptive algorithms create detailed learner models that evolve based on student interactions [99]. The technology tracks multiple parameters including response patterns, time-on-task, help-seeking behaviors, and learning progression to customize instruction. Research demonstrates that effective adaptive systems must balance immediate performance optimization with longer-term learning goals [51].

Real-time Analysis Systems: Modern personalized learning platforms process continuous streams of student interaction data, requiring substantially different architectural approaches than batch-processing institutional systems [88]. These systems analyze multimodal data including: student response patterns and solution strategies; temporal aspects of problem-solving attempts; help-seeking behaviors and resource utilization; and engagement indicators and affective states. This real-time processing enables immediate instructional adjustments based on student needs [34].

Intelligent Tutoring Systems: Intelligent tutoring systems (ITSs) provide one-to-one online teaching through a variety of environments, e.g., modeling, simulation, gamification, and hypermedia. These systems can include learning peers, e.g., animated characters that communicate with students [97, 53]. They perform well in contexts where students would otherwise have no teacher and are nearly as effective as human tutoring [51, 96]. They can leverage real-time affective feedback from students to improve their learning experiences. AI systems can generate content, problems, summaries and feedback and work well in blended learning situations. These tutors have garnered positive results [51] and trained people in fields such as medical diagnosis [21], electronics [102], marine steam propulsion [41] and taught many other disciplines, e.g., mathematics, computer science, language arts, sports, economic history, social issues, history, and politics [75].

AI helps produce new resources for traditional teaching, e.g., it can generate problems, answers, hints, visual and interactive content summaries and synthetic video materials that are as effective as learning from human teachers' video recordings [88]. More attention needs to be paid to determine effective ways to train teachers and parents about using AI tools and integrating them in traditional teaching practices, e.g., professional learning to help teachers use it, gain confidence in using it, judge product quality and reliability, assess potential workload considerations, and learn to trust these systems (assuming that these tools deserve to be trusted). Other constraints are schools' support mechanisms for helping teachers work with such systems in their classrooms and potential privacy and ethical concerns around a system's ability to support inclusion, diversity and equity.

Integration and Orchestration: Recent advances have focused on integrating these various technological components into coherent learning environments. Research indicates that successful integration requires careful orchestration of different AI components to provide seamless support while maintaining pedagogical effectiveness [17]. Studies suggest that well-integrated systems can approximate the effectiveness of human tutoring in specific contexts, though their success depends heavily on proper implementation and support [96].

Limitations and Considerations: While these technologies show promise, research indicates several important limitations. The effectiveness of personalized feedback systems can vary significantly based on factors including: student preparation and prior knowledge; subject matter complexity; implementation quality; available computational resources; and teacher training and support. Understanding these limitations is crucial for effective deployment [66].

2.2 Personalizing Responses Based on Individual Learning Needs

Another feature of pedagogical intervention is to adjust content and difficulty levels based on individual student performance. One objective is to design effective knowledge acquisition tracks that match the learner's strengths and address weaknesses to ultimately meet his/her desired learning goal [64]. Personalized teaching is provided when students make mistakes (providing help) or when they appear confused during cognitive conflict [3]. Pacing problem solving and learning materials can personalize tutoring to reach more diverse learners and enhance accessibility.

However, many risks/limitations of personalized teaching need to be considered. For example, personalization might only marginally contribute to changing the transmissive approach to education as it is focused mainly on content learning and knowledge transmission and, partially, also on problem solving. This approach may favor development of some abilities but not necessarily competencies.

Large Language Models. Large language models (LLMs) respond to user requests with relevant content, and recognize, translate, predict, or generate text or other content to process, understand, and generate human language. LLMs are useful for analyzing, summarizing, and creating content across many disciplines. They use deep learning algorithms to analyze large amounts of data and perform natural language processing (NLP) tasks on that data. They are trained on massive amounts of text, learning to recognize patterns and relationships between words and phrases.

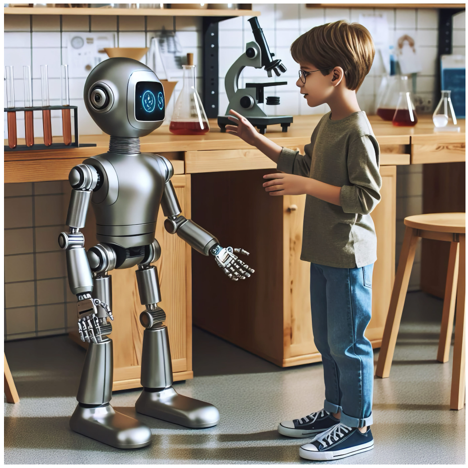

Fig. 3. A young boy interacts with a robot in a laboratory. Social robots provide learning experiences through social interaction with learners. Long-term human-robot interactions highlight the increasing popularity of using social robots in educational environments.

LLM are used to personalize education within educational partnerships, in which the human or LLM might benefit. For instance, students might work in conjunction with ChatGPT-4, to develop and refine assessment questions or explanations of learning content. LLMs are conversational and responsive to diverse student input and facilitate productive student thought rather than produce specific output. Such systems might be judged on the use of supportive and encouraging language or on giving topical hints phrased in a way a 7th grader would understand. LLMs can develop question answering systems [60], generate multiple choice questions [20], score essays [105], perform reasoning [110], support critical thinking [1] and solve math word problems [59]. They might cater to special needs, disabilities, and specific preferences of learners (e.g., readability, language, level of language, speaker speed, etc.). Limitless activity types can be created using a plethora of techniques.

Traditionally, creating personalized educational content for intelligent tutoring systems (ITS) has required a collaborative team of curriculum designers, subject matter experts, and instructional technologists. This process is both time-intensive and costly, especially when content requires regular updates to keep pace with curriculum changes, technological advancements, and evolving student needs. Large language models (LLMs) offer a transformative approach to reducing these high content creation costs. LLMs, trained on vast datasets across various subjects and disciplines, present a low-cost, rapid solution for generating educational materials like questions, explanations, feedback, and hints. For example, LLMs excel in math content generation, with models like OpenMathInstruct-1 scoring high on a dataset of 8.5K

Interaction Design and Architecture(s) Journal - IxD&A, N.64, 2025, pp. 28 - 62
DOI: 10.55612/s-5002-064-001sp

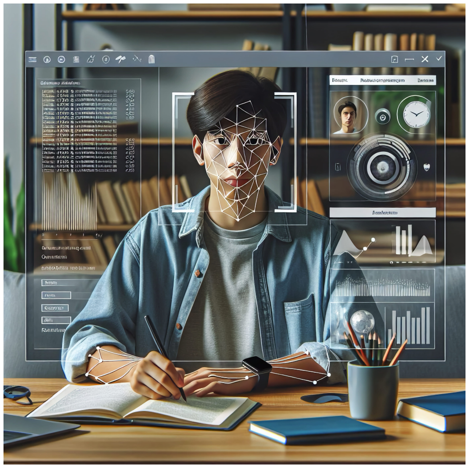

Fig. 4. A student’s facial expression and gestures are tracked while he studies. A personal tutor responds to his emotion, motivation and interests.

high quality linguistically diverse grade school math word problems created by human problem writers. This model also scores well on the MATH dataset benchmarks, which evaluate LLM responses to complex math problems [92]. For example, LLM can summarize text and produce questions and tests from that text [106]. This provides flexibility for students to choose what and when they learn, with or without human teachers. Recommender systems advise students which lectures or videos to view based on their learning and interests. Additionally, sentence-deletion methods for text simplification and knowledge extraction from textbooks have been tested and their effectiveness and the resulting discourse measured [86, 112].

LLM-generated hints enhance online math learning by providing adaptive, clear, and structured support, simplifying complex concepts, and guiding students through each step. Pre-service teachers in one study valued these hints for their accessibility and relatable connections to familiar concepts, which make challenging material more manageable [36]. This responsiveness enables ITS systems to deliver on-demand, personalized assistance, reducing human intervention and fostering scalable, inclusive educational support.

Recording Student Characteristics. AI and advances in big data analysis have enabled educational platforms to precisely record students’ academic and affective characteristics [48, 64]. For example, AI can track students’ learning, reason about their behavior and performance, provide real-time interventions, enhance student engagement, promote active learning and cater to diverse learning styles. Machine learning (ML) is the process of parsing data based on a sampling data set known as “training data” to generate meaningful patterns and structured knowledge. It can leverage student achievements data, aspirations, e.g., ML can help create recommendations for students as they select classes, even choose universities. Moreover, this technology can help instructors gain an understanding of how each concept was digested by each student [50]. In this way, instructors can adjust the teaching method to work well based on students' cumulative records, which may help students grasp course material better. This is done, in part, by observing past experiences available through data and by exploring the learners' features and similarities. For example, the system recommends the most appropriate content among numerous possibilities, including well-designed long-term curricula, and connects learners to performance evaluation. AI tools recognize and respond to learners' performance or emotional states using technologies that detect facial expressions, gestures, and other physiological signals to infer emotional states [85, 108, 109].

2.3 Foster Critical Thinking and Student Control

Another feature of the pedagogical innovations challenge is to enhance a student's critical thinking, which involves a learner's ability to analyze facts, evidence, observations, and arguments to form judgements. This requires that learners be self-directed, self-monitored, and self-corrective. One challenge is to foster development of the widest possible range of life competencies (among which critical thinking is one).

One quick way to use AI to help teach critical thinking is for learners to repeatedly ask questions of a chatbot (e.g., ChatGPT, Bard) and to continue questioning the chatbot to refine their own critical thinking. The answers received from chatbots are full of facts and results generally expressed in syntactically correct English. AI tools also amplify human creativity in digital art software [113], music production tools [93] and platforms for individuals to express creative ideas, helping learners to flesh out ideas for text, graphics, ideas, alternatives, and to compare their creative products. Currently, AI has had a limited impact on students' critical thinking and problem solving [80].

AI also supports learners to be in control of their own learning, e.g., LLMs support learners to set their own agenda, ask their own questions and become engaged in natural conversations about nearly any topic [106]. Promising results in pre-training LLMs for educational settings enable tutors to discover patterns and relationships among topics, student learning styles and skills and knowledge [64, 106]. They can be tweaked to satisfy educational design goals, e.g., to prefer short responses [19] or longer responses, appropriate for giving feedback on essay drafts [69] Other design goals may include that the tutor should engage in 'Socratic-style' dialogues to guide students to arrive at the correct answer without directly providing a solution [76] or that the LLM should avoid hateful, abusive, or derogatory speech, particularly when used with children.

LLMs store terabytes of data (e.g., books, papers, or web pages) and use natural language to converse with humans and the more data fed, the better they can create new material. Chatbots or question-answering (QA) agents help students talk with data and support conversations between machines and humans. Yet these systems come with great risks; they may hallucinate and produce incorrect statements; and data is subject to algorithmic bias (in terms of gender, race, and culture). Most LLMs that incorporate neural networks are trained on huge amounts of text to aid in natural language processing. However, the quantity and quality of training data are critical to ensure diverse, ethical and inclusive instruction. Deployed unchecked, these models can exacerbate falsehoods, inaccuracies and biases of gender, race, or culture at scale and lead to the homogenization of thinking due to the monolithic nature of the models. The text (and graphics) produced by LLMs are difficult to explain (interpret), placing their accuracy and truthfulness in doubt. Notable large chatbots include ChatGPT-4, regarded as one of the most complex ever developed. It was originally trained on terabytes of data created in 2021 and comprises 175 billion parameters. It now has unrestricted access to the entire internet. Training the GPT algorithm requires several months due to its massive size.

Many logistic benefits are provided by AI language models in education; they support learners to make basic corrections to spelling and grammar, proof reading work for students writing projects, neatening up papers or write-ups, along with suggesting tools for improving wording or clarity and software helps students cite sources correctly and do research [83]. Image input technology (available with ChatGPT) enables software to read and identify elements of an image and turn it into text; a game-changer for students with visual disabilities.

However, students should not let chatbots generate whole essays; LLM answers from chatbots are not necessarily truthful nor accurate and learners must check every sentence produced for misstatements and falsehoods. People often become oblivious to the faults of LLM and allow these systems to prepare reports. In fact, learners should repeatedly supply new questions to refine their own thinking and check the truthfulness of system response. Relying on AI-generated content for answers or solutions may produce a worse grade than if students had written the document themself. Although the resulting documents might be technically correct, the base software generally isn't smart enough to write anything above a low average university paper [14]. Though the text might be quite insightful, teachers should advise students that the best way to avoid charges of plagiarism with chatbots is to write the text themselves. Plagiarism is a serious issue, and students should be penalized for doing it [56]. However, recent research indicates that language analysis tools (e.g., Turnitin) are not reliable in detecting ChatGPT-generated text [101]. These tools look for features such as unusual word choices, repetitive sentence structures, and a lack of originality. Also, teachers can tell if something was written by ChatGPT quite easily [36]. Even with a hyper-specific prompt, the resulting written material will be a generic essay that is easy to spot.

ChatGPT can't replicate human intellect, nor reason about its own data, nor incorporate new knowledge. It predicts the next word based on what is already seen. AI based applications, such as ChatGPT can support human creativity but they cannot be creative (i.e., they can only propose recombinations of previously generated elements). Competence-based education, in which application like ChatGPT partake, explores development of a paradigm aimed at the harmonious development of the learner that could favor transformation of the "factory-school" model [38]. AI will remain as assitants to teachers; humans need to remain creative (writing, composing, sharing) as an outlet for their expression, enabling them to be heard in different ways, increasing engagement and boosting self-confidence.

2.4 Multimodal Learning Interactions

Another featuere of the pedagogical innovations challenge is to integrate multimodal learning (not just cognitive learning) and to include different learning styles at the same time. Students learn best when educators provide multiple and multimodal learning styles using various media. This includes the use of multiple senses and teaching styles, e.g., kinesthetic, auditory and visual. AI applications contribute to multimodal teaching. This section examines how AI supports multiple modes of interaction beyond purely cognitive approaches, recognizing that effective learning engages multiple sensory and social channels.

Physical and Gestural Interaction. Research suggests that physical interaction plays a crucial role in learning processes [61, 71]. Computer vision systems form the foundation of motion tracking in educational environments, using specialized algorithms to capture and interpret students' movements and gestures in real-time [42]. These systems differ significantly from traditional educational technology, requiring sophisticated spatial recognition algorithms to understand three-dimensional movement patterns and translate them into meaningful learning data, see Section 5.1 [4, 6]. Haptic feedback systems complement this by providing tactile responses, creating a bidirectional channel of physical communication between learners and educational content [63]. Studies indicate these technologies require fundamentally different processing approaches than traditional cognitive-focused systems, as they must integrate multiple streams of spatial and temporal data in real-time [74, 115].

Affective Computing Integration. Emotional engagement represents a distinct mode of interaction requiring specific technological approaches. Modern emotion recognition systems employ facial expression analysis through specialized computer vision algorithms that track micro-expressions and facial muscle movements [4, 27]. This visual analysis works in concert with voice tone analysis systems that process audio inputs to detect emotional states through vocal patterns, pitch, and rhythm [24, 25]. Additionally, physiological signal interpretation adds another layer of emotional understanding by monitoring indicators such as heart rate variability and skin conductance [7, 48]. These integrated systems employ different architectures than cognitive learning tools, focusing on real-time signal processing and pattern recognition across multiple sensory inputs to build a comprehensive understanding of learner emotional states [5, 8, 10].

Social Learning Support. Social interaction technologies use distinct approaches from individual learning systems, focusing on the complex dynamics of group learning. Network analysis algorithms process interaction patterns among learners, tracking both direct communications and implicit connections formed through shared learning activities [114]. Collaborative filtering systems build on this network analysis to support effective group formation, using historical interaction data and learning outcomes to suggest productive learning partnerships [39]. Social graph analysis tools extend this functionality by monitoring the evolution of learning communities over time, tracking changes in interaction patterns and identifying emerging leaders and support networks within student groups [94, 116]. These tools require fundamentally different architectural approaches than individual learning systems, as they must process and analyze relationships and group dynamics rather than individual performance metrics [72].

Fig. 5. Students on a forest field trip use devices to make observations about leaves.

Environmental Integration. Learning occurs within complex physical and digital environments, requiring specialized technological approaches to support effective instruction. Internet of Things (IoT) device networks create a foundation for monitoring physical learning spaces, collecting data about environmental conditions that might affect learning outcomes [16]. These systems work in conjunction with environmental adaptation systems that can automatically adjust learning conditions based on collected data, optimizing factors such as lighting, temperature, and noise levels for different learning activities [46]. Context-aware computing builds on this environmental data to support situation-appropriate interactions, adjusting the presentation of learning materials and the nature of learning activities based on physical conditions and time of day [6, 22]. This integrated approach requires sophisticated sensor fusion algorithms and real-time adaptation systems that differ substantially from traditional educational technology architectures [13, 23].

Tracking Student Movement with Computer Vision. Computer vision is used to track learners' bodies, gestures and physical interactions while they learn. Students' actions and gestures are essential components of their learning, as ideas that are distributed among the mind, the world, the social context, facial expressions and hand gestures support essential communication during learning [74]. For example, students in classrooms explore mathematical relationships and concepts by manipulating real objects in the environment and communicating face-to-face with peers/teachers[4, 6]. Yet, there is a significant disparity between the current learning technologies employed in schools and effective practices of educators. For example, conventional K-12 mathematics technologies are typically not designed to facilitate embodied cognition; instead, they tend to align with more traditional perspectives on student learning (e.g., show video, give quiz).

Embodied Cognition. Embracing embodied learning implies moving the human-computer interaction (HCI) "off the keyboard" so students experience phenomena/relationships, engage in hands-on and socially-rich activities, and obtain support during moments of struggle and cognitive conflict [48]. The theory of embodied cognition assumes that sensory perceptions, motor functions, and sociocultural contexts shape the structure and development of thinking skills, including mathematical thinking [42, 116]. Learning involves the creation, manipulation, and sharing of meaning through bodily interactions [61]. For example, research on hand motion gestures to externalize ideas suggests that a motor encoding of math ideas exists, and thus, that mathematics teaching should use bodily motion, action (hands-on and body-on), and hand gesturing. Random movement does not lead to math learning, rather movement encouraged by math-embodied activities will capture, and/or express mathematical concepts to be learned [71].

Several learning technologies have addressed embodied learning, a shift that entails focusing on refining and deepening physical interactions with learning technology. For example, teachers have implemented tasks that are hands-on and cognitively demanding [22]. This challenge includes technologies tailored to support intelligent tutoring systems that trace students' knowledge by tracking students' performance and mental states (i.e., knowledge and affect) and embodied cognition and multi-modal interaction (students physically act on the environment in meaningful ways for more solid understanding/encoding of ideas).

This challenge facilitates bridging from concrete representations to abstract concepts, and involves students expressing solutions to math problems through their entire bodies. This may include using the body to measure objects, people, or spaces, and even physically walking out numerical mathematical solutions. Manipulatives or physical props play an important role, acting as bridges between familiar, intuitive prior knowledge and new concepts, supporting students to develop abstract mathematical thinking [13, 23].

For example, a new technology called Wearable Learning addresses the need for AI systems aligned with existing dynamic classroom activities, where students work collaboratively, use manipulatives and engage in discussions, rather than sit in front of computers [4, 6, 81]. This approach rethinks the design of educational technologies and moves students to engage their whole bodies, to gesture and interact with the world. Wearable Learning argues for development of AI systems that embed concepts from embodied cognition and learning theories.

Using AI for teaching math aligns with theories about how people learn. For example, embodied learning posits that sensory-motor skills are essential for learning. The exploration of cognition in research is expanding interactions to encompass wider viewpoints [33]. AI can guide students to encounter, discover, and rehearse perceptuo-motor schemas in relation to math concepts and relationships [73]. Cognition can be linked to student action and supports the idea that physical motion aids learning [63].

3 Address the Digital Divide

A second grand challenge for AI and education is to address the digital divide and develop strategies that enhance equitable education access for all people. Short term measures usually involve acquiring hardware and connecting with networks, e.g., providing devices, establishing WiFi hotspots, expanding technology access in public libraries and advocating for government policies to close the digital divide [72]. Long term strategies involve local communities, educational institutions, and policymakers for sustained efforts and effective solutions

3.1. Equitable Distribution of Education

One feature of adressing the digital divide is equitable distribution of resources to educate all people. Strategies are needed to achieve equitable distribution and to ensure transparency, accountability, and support for underrepresented communities, while also promoting social justice and collaboration across sectors, e.g., assessment of educational needs, targeted funding, community engagement, transparent allocation processes, equitable educational opportunities, affordable educational services, workforce development programs, periodic evaluation and international cooperation [72].

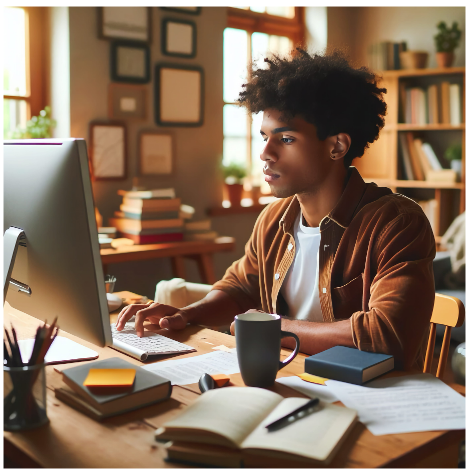

Fig. 6. A computer provides access to a vast amounts of information, interactive learning tools, multimedia content, and the ability to organize and manage study materials efficiently. However, its effectiveness depends on how it is used. Students must manage distractions and be mindful of potential downsides like screen fatigue, improper learning techniques and passive learning.

Research suggests that AI may help address certain diversity issues. For example, studies indicate that AI-powered tools can assist with cross-lingual translations and topic annotations, though effectiveness varies across languages and contexts. Students with special needs or disabilities constitute 15% of the out-of-school population and face complex barriers requiring special software [104]. For example, students with a sensory, physical or intellectual disability are 2.5 times *more likely* to have been out-of-school in comparison to their peers without disabilities, possibly due to lack of adapted infrastructure and materials and because of the known link between disability and poverty [43]. In this sense, assistive technology can determine participation in society or marginalization.

Research indicates that while many current AI education systems originate predominantly in developed nations, learners worldwide face varying educational contexts and challenges. Studies suggest that more inclusive development approaches, incorporating diverse global perspectives and contexts, may lead to more effective and equitable systems [17]. Human agency is required in the design and development of AI systems; learners must be placed at the epicenter of these designs, giving humans the ability to scrutinize them (e.g., by querying their biases and beliefs, changing the optimization goals and addressing the algorithmic bias). The human-computer loop should also revolve around less-than-ideal scenarios based on ill-defined datasets from imperfect settings and should engage all stakeholders in the design process.

Assessment of students' performance is key to equitable distribution of resources as student evaluation is critical to tracking the digital divide. A key problem in assessment is to estimate how well students have mastered each knowledge component/concept/skill. Assessment in traditional classrooms involves a cessation of teaching, i.e., teach, stop teaching, give a test, resume teaching. AI systems offer alternatives to traditional assessment through what Shute & Ventura (2015) describe as 'stealth assessments,' which analyze student interaction logs to help evaluate skills and knowledge in real-time. Assessment and evaluation can occur alongside learning activities, e.g., online teaching while the system tracks various factors such as performance, emotional states, and motivation [34]. However, the effectiveness of such approaches may vary depending on context, student population, and implementation quality and the accuracy and reliability of such tracking may vary significantly across different implementations. Intelligent platforms provide a variety of pedagogical styles (videos, explanations, narratives) and can assess a student's learning needs in real-time and immediately supply necessary activities. AI systems take advantage of data collected from traditional assignments (problem sets, laboratories), exercises and tests. Many AI systems incorporate assessment capabilities as part of their lesson optimization features [34], though the effectiveness of these tools appears to vary based on factors such as subject matter, student level, and implementation context.

3.2 Diverse, Ethical, and Inclusive

Another feature of adressing the digital divide is to incorporate diverse perspectives that contribute to more inclusive learning environments. When learners from varied backgrounds see their experiences represented, they may feel more engaged in the learning process [35, 72]. Culturally responsive education respects, values, and incorporates learners' diverse cultural backgrounds, experiences, and perspectives.

Cultural backgrounds may influence how students engage with and interpret educational content and can play a significant role in shaping learning experiences, though the extent and nature of this influence may vary across different contexts and individuals [35]. Cultural sensitivity and inclusivity require understanding and appreciating different backgrounds, traditions, and perspectives without imposing one's own beliefs or values. Inclusive practices promote a sense of belonging, equity, and equal opportunity for everyone, regardless of their cultural or social identity.

Providing diverse, ethical and inclusive education involves creating open dialogue spaces, ensuring diverse representations while designing and developing these systems, fostering cultural competences and encouraging open communication [72]. Recognizing the intersectionality of identities (e.g., race and gender) can break down bias and discrimination. Collaborative initiatives and fostering a culture of continuous learning encourages individuals to seek new perspectives and stay informed about different cultures.

Fig. 7. Three students collaborate in-person while a computer facilitates sharing. Multiple users add content simultaneously and use presentation software (slides) and online whiteboards to support brainstorming of visual ideas, drawing diagrams. Roles and responsibilities are clearly defined to ensure effective contributions; open communication and active listening are required.

Designers and developers of instructional systems need to support and engage learners' rich cultural identities and cultural signatures to express their cultural identities [72]. *Maker* stakeholders (researchers, developers, commercial companies) should intentionally think about meaningful ways to honor and actively engage learners' rich cultural backgrounds and include features and experiences that support and enable learners to draw on their cultural identities.

Many educational systems face ongoing challenges with equity, inclusivity, and accessibility. Approximately one in four 15-year-old students report feeling like outsiders at school [103], though experiences may vary significantly across different educational contexts and cultural settings. Accessibility is a large part of the problem, because education is not available to all students for many reasons. For example, 30% of students in the Dominican Republic report feeling like outsiders due to language differences in a country that is one of the most linguistically diverse in the world and where educational systems do not always account for this [79]. Many countries still practice segregation in their educational systems, especially for students with disabilities.

Studies have identified various ethical challenges, including documented cases of data breaches and privacy concerns [82]. Access to educational technology continues to show disparities across different socioeconomic groups and regions, though the extent and impact of these disparities may vary. *Maker* stakeholders are generally aware of these issues, e.g., several AI and education publishers require an ethics statement before papers are published. However, more effort is needed to ensure that basic system components are ethical and that students are mutually respected.

Ethical systems also require addressing *data privacy, bias, discrimination,* and *social impac*t. AI systems collect a great deal of information about learners, e.g., personal information, performance data during learning, and log data of interactions (questions posed, hints requested). Maker stakeholders need to ensure that their systems maintain data privacy in all communications, just as it is required for all teachers. Student data must be kept private in general communications and kept anonymous in publications. *Fairness* is another concern, requiring that computer tutors be fair to individuals, groups and communities. This means maker stakeholders must avoid favoritism during development, support academic integrity (data privacy) and not allow exploitative labor practices, inequitable access to technology or the potential for copyright infringement. They must also promote equality and non-discrimination among learners. *Bias* is a large issue for AI systems and comes in part from training large models; if a system is trained on a narrow segment of the population (e.g., a dataset constrained to a single gender, race, or culture) then this system's reasoning and responses will be biased (e.g., consistently using male pronouns to describe professionals). Also deep learning models developed by biased systems will propagate bias throughout tutoring problems, answers, and responses. To avoid educational *inequality,* AI systems should be trained on a diverse and inclusive set of learners (e.g., maker stakeholders should themselves be diverse in gender, race and culture) and deployed inclusively across nations.

Certain limitations in AI educational systems might potentially affect educational opportunity gaps and these effects could vary significantly across different contexts, with some regions potentially experiencing more pronounced impacts than others [17]. Factors such as connectivity, content quality, and local capacity appear to influence outcomes [70]. Good connectivity, quality content, and capacity building are necessary but not sufficient to support enhanced education in remote locations [70]. AI systems might propagate dangerous biases at scale and divert educational resources that could be put into more effective classroom use [17]. Other challenges include availability of quality data, fact-checking, learner modeling, and transparency of these systems. Political power is also a serious challenge as people who possess the technology might impact others negatively or exert control.

4. Global Learning Communities

A third grand challenge for AI and education is to support global learning communities. International partnerships can play an important role in supporting education, particularly in developing countries [72]. Well-designed and properly implemented partnerships may help demonstrate shared responsibility, strengthen international bonds, and work toward reducing inequality, though research suggests results vary significantly based on local contexts and implementation approaches.

4.1 Collaborative Problem Solving

One feature of global learning communities is that they bring together individuals with diverse skills and perspectives to address complex challenges [72]. Studies indicate that team diversity may enhance creative outcomes under certain conditions, while design thinking principles typically incorporate practices such as empathizing with end-users, defining problems, and prototyping - though effectiveness can vary across different contexts and cultures. Strategies include team diversity, clear problem definition, design thinking, brainstorming sessions, structured collaboration tools, role assignments and active listening.

Social and collaborative platforms are essential for creativity. Online collaborative spaces (e.g., Zoom, Slack) now support remote teams and individuals to work together. AI provides social support and gracious cues normally offered only by in-person collaborators. To work well, these platforms require high-quality communication (e.g., not requesting that users respond to messages or press buttons while expressing themselves creatively). Good collaborative interfaces will limit interference from the tool while users are working or expressing their creativity [114]. AI tools should support sensory, precognitive, cognitive, ergonomic, behavioral and social abilities, e.g., users can easily find a screen or template to enter a note. Many remote collaborative interfaces have these features and are now well deployed (e.g., Zoom, Google Meet, Microsoft Team).

Systematic frameworks can be provided in every country to remove barriers for all learners: ‘every learner matters and matters equally’ [2]. Inclusion is a prerequisite for sustainable societies and for democracies based on fairness, justice and equity. The benefits of inclusive education are hard to quantify, as they extend over generations, but they include improved academic achievement, social and emotional development, self-esteem and peer acceptance, as well as preventing stigma, stereotyping, discrimination and alienation [94]. Accessibility and inclusion are a foundational approach and the epicenter of education technology design. Computer solutions that do not address the deeper barriers of exclusion and inequity can only go so far towards improving learning outcomes.

4.2 Benefit More Students

Another feature of global learning communities is to extend educational reach beyond traditional methods and to provide high-quality education to almost eight billion people. This requires more than purely technology-centered solutions; it requires future-oriented visions of connected social and technological solutions [39]. One goal is to ensure that every learner has access to high quality AI tools that are suitable and beneficial. Another goal is to ensure that every learner has access to high quality AI tools that are suitable and beneficial. The effectiveness of global communities varies significantly based on factors such as infrastructure, teacher training, and local context. Significant challenges remain in ensuring equitable access and appropriate implementation. In prior studies, students using intelligent tutoring systems (ITSs) outperformed students from conventional classes in 92% controlled evaluations, and the improvement in performance was great enough to be considered of substantive importance in 78% of the studies [51]. Reports of improvement in student comprehension, engagement, attitude, affect, motivation, and academic results have contributed to investment in and research about these systems. ITSs are roughly equivalent to an improvement in test performance from the 50th to the 75th percentile and can improve the learning gains for average students by two standard deviations (2 SD), a measure of how far each student's score lies from the mean [51]. Two standard deviations indicate that values are clustered close to the mean or that a majority of students learned the material. This is stronger than typical effects from other forms of teaching and about twice as high as results from non-intelligent computer aided teaching systems.

To enhance the benefits of AI for more students, hidden human coaches might be used to improve the competency of these systems under special circumstances. For example, human operators might be included as "hidden coaches" who monitor interactive instruction during the training period for the ML components with different classes of learners. Hidden coaches would be situated with learners to guide and help the AI systems to complete their assigned tasks, remotely seeing the tutoring session through the eyes of the computer tutor and performing relatively simple tasks. Scientists would collect and analyze gathered data especially for different classes of learners and make critical creative decisions about which alternative actions the environment might take next. The hidden coaches do not take the place of actual human teachers; they're doing reconnaissance work to clear a path for the AI system to eventually improve its teaching.

Some intelligent instruction is available on open frameworks in which all code is freely available [90]. This supports transparency and democratization, as many scientific discoveries depend on both technological prowess and user studies. Such transparency also improves efforts towards understanding the ethical and fairness aspects of AI mass adoption. Open-sourcing is a trivial enabler and provides utility of AI tools in developing nations [17]. However, open-sourcing is not a magic enabler, as complex knowledge transfer is needed to build both operational and human capabilities to work with cutting-edge tools. Global collaborative ventures and earnest capacity building are required in AI to avoid falling back to techno-solutionism, or the idea that technology can solve all social, political, and access problems [29].

Fig. 8. Lifelong learners use informal learning tools and location-based delivery of content to identify plants.

However, many limitations and risks remain for using AI systems in educational settings. AI systems are not effective with questions that require subjective judgment or opinion, such as matters of personal taste, ethics, or moral dilemmas [78]. These types of questions often lack a single correct answer and AI systems do not yet discuss human qualities such as idea generation, intuition, cultural sensitivity, emotions and common sense. A serious limitation of ITSs is that they provide easy and frequent feedback and hint sequences, and some students will immediately turn to hints before attempting to solve problems or complete tasks [99]. Students might bottom out the hints – request as many hints as possible as fast as possible along with answers – to complete tasks faster. A second limitation is that these systems do not ask students to explain their actions [98]. If students are not engaged with online tutors it becomes more difficult for the AI system to gain a deeper understanding of students' learning. Students who can't explain their knowledge become opaque to AI systems. A third limitation is that ITSs are focused on instruction and not typically involved in student motivation or social-emotional learning [96]. All these limitations are being addressed in recent research with LLMs and solutions should be forthcoming, e.g. researchers are using LLMs to create question answering systems [60], multiple choice questions [20], essay scoring [105], supporting critical thinking [105] and solving math word problems [59].

4.3 Lifelong Learning

A final feature of global learning communities is to support learning continuously over the entirety of one's life (lifelong) and across all aspects of that life (lifewide). This refers to adapting resources to a persons' level of understanding, in ways that are highly relevant to each learner and doing so throughout life. AI can provide authentic learning opportunities that blur the distinction between learning and life, promoting the joy of learning. Many individuals do not participate in any meaningful learning after formal schooling ends and many others have only sporadic and highly interrupted patterns of engagement. These inequalities are highly dependent on an individual's age and stage of life, as well as patterned in terms of income, gender and social class [49, 55].

Career development in the information age may be measured as much by the acquisition and development of knowledge, as it is by the rank and title of each particular job [44]. In this context, "career," metaphorically, can be characterized as a repository of knowledge [12]. Animated and AI-based agents might motivate users based on age, economic, or cultural considerations [58, 91, 95]. Agents have taken on authentic role models as virtual learning companions and have promoted positive attitudes to build up a learner's self-efficacy [53]. They might request particular topics and knowledge components on behalf of their adult users provide complete models of their adult users; e.g., agents orchestrate their own interactions and adapt to the learner's characteristics (e.g., cognitive ability, previous skills, culture) and their needs (disabilities, learning difficulties). Learners can call upon virtual characters as virtual teachers and companions [15, 91]. These characters are not only knowledgeable, they also carefully reflect the characteristics of people they model. Agents might enhance teachers' professional development and best practices training for job advancement, recreation or instructional skills-based learning support lifelong learning (longitudinal), and ubiquitous (embedded) experiences [47].

Persistent interfaces adapt to learners across life transitions and stages. Early research suggests that AI systems may develop detailed learner profiles over time, though the extent and accuracy of such knowledge requires further investigation. The effectiveness of these tools varies based on factors such as user engagement, data quality, and system design.

5. Data-driven Decision Making

A fourth and final grand challenge for AI and education is to make strategic educational decisions based on data, e.g., student learning rate, questions asked, and inferred knowledge.

For example, data-driven tools might identify students at risk, challenge students who seem bored, guide teacher professional learning by identifying areas of curricula difficulty, support budgetary and policy change, and promote transparency. Data analytics in education is crucial for improving learning outcomes and processes [72]. For example, huge quantities and high-quality student data support examination of learning results, new instructional design, and reimplementation of platforms based on identifying (in)effective problems. Data-driven tools collect and organize data, analyze and interpret it, evaluate and improve it, and communicate data to teachers who can apply it while interacting with students. ML based on data utilizes different algorithms to analyze data and deliver a better model of teaching and learning.

5.1 Predictive Analysis

One feature of data-driven decision making in education is to use predictive analysis to determine and predict student and class outcomes using data. AI research provides novel approaches for predicting student outcomes using visual and tabular cues, though these predictions may vary across different educational contexts and student populations. Predictive analysis contributes to traditional classroom prediction methods, e.g., teacher observations, formative assessments, and student self-assessments. Predictive AI involves statistical analysis to identify student patterns, anticipate behaviors and forecast performance. It can anticipate a learner's response (e.g., whether a student will solve a problem, ask for hints, quit, guess) and provide response material before learners become discouraged or blocked [8, 104]. Early prediction of performance while students work online is crucial for providing timely and effective interventions, potentially improving student engagement and motivation and ultimately enhancing their learning [45]. Administrators use advanced analytics and machine learning algorithms to gain insights into class-level performance, to identify student difficulties, to make policy decisions, and to allocate resources more effectively [11]. For instance, studies by Yağcı et al. (2022) and Maghsudi et al. (2021) found correlations between specific factors (e.g., extracurricular participation and learning styles) and academic performance. This helps to identify areas for improvement for instructional systems, factors that contribute to dropout rates and interventions that support struggling students.

Fig. 9. Data-driven decision making based on statistical analysis is used to identify patterns in a student's learning, anticipate behaviors and forecast future events. Teachers share data with the student about her learning.

In one study examining student performance data, Battula (2023) found that linear regression achieved 88% accuracy in predicting student results. Educational data mining, while still an emerging field, attempts to identify relationships within educational data to better understand factors influencing academic achievement. For example, Yağcı et al. (2022) applied various machine learning algorithms to predict undergraduate midterm grades, achieving varying levels of success depending on the specific algorithms and contexts used. The performance of several ML algorithms (random forests, nearest neighbor, support vector machines, logistic regression, Naïve Bayes, and k-nearest neighbor) were calculated and compared to predict students' final exam grades. In this example, the dataset consisted of the academic achievement grades of 1854 students who took the Turkish Language-I university course [107]. Results show that the proposed model achieved a classification accuracy of 70–75% using only three types of parameters: midterm exam grades, department data and faculty data. Such data-driven studies establish learning analysis frameworks in higher education [50]. Clearly, such studies also help predict students at high risk of failure. However, predictive models have limitations and should be used as one of multiple tools in educational decision-making; different analytical approaches may be more appropriate in other settings or with different types of educational data.

In another example, a facial expression recognition system predicted student outcome in solving online math problems after analyzing only the first several seconds of each student's behavior. While students used an intelligent tutor platform, their behavior was captured in a video feed (videos of students interacting with the online system), along with timing information obtained from their learning log (student keystrokes) [27, 57, 108]. This multimodal approach predicted student exercise outcomes by analyzing the first k seconds of student video, where k was set to be as small as 5 sec and as large as 20 sec.

This team extracted facial affective embeddings from video frames using transfer learning and analyzed their temporal dependencies using a Transformer [84, 85]. The timing information about when students take certain actions (e.g., request a hint) was combined with video representations, enhancing the model's ability to predict student performance quickly. Initial studies suggest promising results compared to existing baselines when tested on specific student datasets. Moreover, due to the complex nature of student behavior, achieving high prediction performance has proven extremely challenging for these approaches. Even human observers demonstrate low accuracy predicting exercise solving outcomes solely from video, indicating the need for accessing additional information to improve prediction accuracy.

While these findings indicate potential for early outcome prediction, further research across diverse student populations and learning contexts is needed to validate preliminary results and to understand their broader applicability.

AI's ability to predict student performance is limited and depends on the nature of the available data. While AI can analyze patterns in historical data and make informed forecasts in certain domains, predicting complex and dynamic events about students' learning remains challenging. For example, family and home events, health, evolving circumstances, and the inherent uncertainty of human behavior make predicting student outcomes difficult. The effectiveness of these tools depends heavily on data quality, implementation context, and integration with existing educational practices.

5.2 Model Teaching and Learning

Another feature of data-driven educational decision making includes models that are generalizable across learning content, learners, and learning scenarios. One goal is to develop models that represent and analyze learners (e.g., knowledge level, learning pace), teachers (pedagogy, teaching style) and their interactions. This approach combines both cognitive science and learning science theory and data-driven modeling. For example, one research team is developing LLM-based simulated learner agents that effectively mimic the behavior of real learners [62]. The benefit of such a model is to provide a proxy of real learners from any background and under any learning context. Teachers and tutors can practice tutoring strategies with such a model and can identify best pedagogical practices and content designers can test the quality and difficulty of new learning content and assessment/practice questions and even evaluate new curricula. Key AI strategies are learning analytics, early intervention systems, individualized learning paths, predictive analysis, inclusion and equity metrics, and a culture of continuous improvement.

Understand Learner Deficiencies. One key limitation for existing learner models, is that they typically analyze and predict binary-valued learner responses to questions, i.e., the correctness of the response. As a result, this approach loses important information since it ignores the exact (mostly textual) content of questions and learner responses, especially for open-ended questions. One goal is to predict learner deficiencies and to generate open-ended answer responses that utilize this missing knowledge, hoping the student will recognize and focus on the text. For example, one research team combined learner models that produce latent learner knowledge representations with the generative nature of LLMs [32, 111]. In a computer science education setting this model effectively captured students' specific coding bugs and styles and was able to steer LLMs towards generating computer code, especially incorrect code, in a way that mimicked each individual learner. This is a step towards personalizing LLM output to each individual learner to cater to their background, knowledge levels, goals, and needs.

Make Errors. LLMs are trained on clean, high-quality textual data, which makes it difficult for them to understand human errors, let alone generate output with these errors in a controllable way. One research team set up an LLM to generate mathematically valid distractors as components of multiple choice questions, i.e., a possible solution that is a sensible result of flawed mathematical reasoning [31]. The group studied the problem of automatically generating incorrect options, utilizing the question's stem and key. The team found that LLMs are better at generating distractors, than understanding what errors real learners actually make.

Enhance LLMs' Capability in Math. LLMs perform well on a wide range of natural language generation tasks, but less so on mathematical ones that require rigorous reasoning and precise calculations. One research team developed MathGPT, a modification to GPT's architecture to encode and decode mathematical expressions as operator trees [87]. MathGPT improves LLMs' ability to represent mathematical content and their ability to plan their reasoning processes. The team discovered ways to improve the validity of GPT's generated mathematical expressions.

Question Generation. Historically, educators thought that to create a "level playing field" for all learners, they would assign the same standardized question to each learner. However, this approach fails to account for the variation in individual learner characteristics (e.g., learning level, construct representation and language skills). Providing similar questions does not address the well-documented problem of decaying learner interest in math, which may prevent them from pursuing science related careers. Personalizing questions for each individual learner, i.e., changing the language complexity, consistency, topic, and other features of a math word problem while keeping its underlying mathematical elements fixed, presents a viable solution to this problem. One research team developed a tool for the controllable generation of new math word problems given a specified problem context, represented as keywords, and an underlying math equation [100].

6. Discussion and Conclusions

This article proposed four grand challenges for AI and education, including pedagogical innovations, addressing the digital divide, global learning communities and data-driven decision making. Several approaches and tools were described that show promise in educational contexts to support, not replace, teachers, administrators, policy makers and other education stakeholders. Several challenges were identified along with limitations and risks for each challenge and the broader impact of each challenge on stakeholders.

Research in this field is driven by the fact that education impacts each individual, both in terms of increased knowledge and earning power. For example, a typical worker in the United States with a bachelor's degree earns 80% more than does a high school graduate [18]. Additionally, a country that does not leverage the enormous payoff for investment in pre-schooling is deficient in its approach to education. The payoff in pre-schooling can be measured in improved college success, higher income, or even lower incarceration rates. Pre-schooling addresses some inequities that begin at birth (resulting from rich/poor parents) and improves the lot of disadvantaged children as they grow. Difference in cognitive performance between rich and poor is just as big at age 18 as it is at age 3, before students enter school. Thus, income inequality is passed down through generations. Humans not only learn and create, but they do so continuously while generating new ideas [65]. This elusive gift includes the possibility to invent everything and to go beyond the ability to solve isolated problems. AI techniques have already extended the success of today's learners in individual studies and might empower learners everywhere. While the findings in this article indicate potential for education, further research across diverse student populations and learning contexts is needed to validate these results and to understand their broader applicability.

AI techniques challenge, and possibly threaten, existing educational practices by suggesting new ways to learn [68]. Yet, technology can't impact education in isolation; rather it operates as one element in a complex social and political system that must consider content, pedagogy, and the environment that students, instructors, and technology co-create [77]. The AI community has historically designed systems to challenge human ability; AI systems were first developed to win against Go or chess masters and AI tools ultimately replaced radiologists, real-time translators and web designers. That old AI paradigm of developing technology to perform better than humans has been replaced by a paradigm to support humans to work better with AI systems, to keep the human-in-the-loop, shifting the focus from surpassing or replacing people to working and learning better together [65].

Many limitations of this review article should be considered. For example, this article describes what we may call "lab experiments" that were carried out on specific tasks in specific contexts (although carried out with real students in real classes). One overarching challenge is to move these experiments beyond the realm of isolated projects in which each research team develops idiosyncratic conceptual frameworks and methods [26] and to enable technologies to become largely diffused and adopted. One exception to the statement that this field has only produced "lab experiments" are LLMs that have not been realized and experimented with in educational contexts but, rather, have been adopted world-wide by teachers and learners. This discussion is not exhaustive and additional challenges and tools should be considered.

Intelligent instructional tools have not been combined in large scale nor in optimal ways for education; they often provide single fixes or add-ons to single classroom issues. Scientists must not fall back on techno-solutionism, in which technical solutions overlook human complexity; humans need to be kept in-the-loop. For example, *maker* stakeholders (researchers, data scientists, computer scientists) are needed to guide design and development of systems towards making them commercially viable. On the other hand, *user* stakeholders (students, teachers, parents, administrators) are needed once a tool is built to evaluate it, to identify (in)efficient or (in)effective components and to assess learning results. Cycles of design and development help evaluate the impact of AI on education, foster ownership and build trust (assuming these systems deserve trust), and move the field from prescriptive algorithms to human-centric and impactful partners.

Although many challenges remain, they are relatively unrestricted and can be addressed without stifling stakeholder creativity and innovation. Powerful AI tools hold great promise for enhancing education, improving human capabilities and skills, and shaping a better global world. Even before this research gets translated and deployed in society—via products and services—human factors, ethics, and diversity come into play. AI and education is a powerful new field and its omnipresence provides potential innovation and transformation. As we stand on the brink of this new era, we have immense responsibility to harness AI and education for the common good.

CrediT Graphics. All graphics were produced with the assistance of AI Microsoft Co-Pilot, a generative AI tool.

CRediT author statement. **Beverly Woolf:** Conceptualization, Methodology, Writing (original draft production), Execution. **Danielle Allessio**: Methodology, Validation, Writing – review and editing, Improvement of document, References.

Ivon Arroyo: Validation, Writing. **Sai Gattupalli:** Writing – review and editing, Critical Revision, References, Final approval. **Boming Zhang:** Writing – review and editing, References, Critical Revision, Final approval.

References

1. Adewumi T., Alkhaled L., Buck C., Hernandez S., Brilioth S., Kekung M., Ragimov Y., Barney E.: ProCoT: Stimulating Critical Thinking and Writing of Students through Engagement with Large Language Models (LLMs), http://arxiv.org/abs/2312.09801, (2023)
2. Ainscow M.: Every learner matters and matters equally: making education inclusive Conference High-level Policy Dialogue to Mark the 30th Anniversary of the Salamanca Statement, Paris, (2024) https://doi.org/10.4324/9781003504788-2
3. Aleven V., Roll I., McLaren B.M., Koedinger K.R.: Help Helps, But Only So Much: Research on Help Seeking with Intelligent Tutoring Systems Int J Artif Intell Educ, 26, pp. 205–223 (2016) https://doi.org/10.1007/s40593-015-0089-1
4. Arroyo I., Closser A.H., Castro F., Smith H., Ottmar E., Micciolo M.: The WearableLearning Platform: A Computational Thinking Tool Supporting Game Design and Active Play Tech Know Learn, 28, pp. 1815–1824 (2023) https://doi.org/10.1007/s10758-022-09601-1
5. Arroyo I., Cooper D.G., Burleson W., Woolf B.P., Muldner K., Christopherson R.: Emotion Sensors Go To School Artificial Intelligence in Education. pp. 17–24. IOS Press (2009) https://doi.org/10.3233/978-1-60750-028-5-17
6. Arroyo I., Micciollo M., Casano J., Ottmar E., Hulse T., Rodrigo Ma.M.: Wearable Learning: Multiplayer Embodied Games for Math Proceedings of the Annual Symposium on Computer-Human Interaction in Play. pp. 205–216. Association for Computing Machinery, New York, NY, USA (2017) https://doi.org/10.1145/3116595.3116637
7. Arroyo I., Royer J.M., Woolf B.P.: Using an Intelligent Tutor and Math Fluency Training to Improve Math Performance International Journal of Artificial Intelligence in Education, 21, pp. 135–152 (2011)
8. Arroyo I., Woolf B.P., Burelson W., Muldner K., Rai D., Tai M.: A Multimedia Adaptive Tutoring System for Mathematics that Addresses Cognition, Metacognition and Affect Int J Artif Intell Educ, 24, pp. 387–426 (2014) https://doi.org/10.1007/s40593-014-0023-y
9. Association for Smart Learning Ecosystems and Regional Development (ASLERD): Troyes Declaration (Timisoara 2.0) - Better Learning for a Better World 2.0: People Centred Smart Learning Ecosystems beyond 2030, https://www.aslerd.org/troyes-declaration/, (2024)
10. Baker R.S.J. d.: Modeling and understanding students' off-task behavior in intelligent tutoring systems Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1059–1068. Association for Computing Machinery, New York, NY, USA (2007) https://doi.org/10.1145/1240624.1240785
11. Battula S.: Student Performance Analysis and Prediction, https://www.analyticsvidhya.com/blog/2023/04/student-performance-analysis-and-prediction/, (2023)
12. Becher T.: 6. The Disciplinary Shaping of the Profession in Clark, B.R. (ed.) The Academic Profession. pp. 271–303. University of California Press (1987) https://doi.org/10.1525/9780520311329-009
13. Boggan M., Harper S., Whitmire A.: Using Manipulatives to Teach Elementary Mathematics Journal of Instructional Pedagogies, 3, (2010)

14. Brady J., Kuvalja M., Rodrigues A., Hughes S.: Does ChatGPT Make the Grade? Research Matters, (2024)
15. Bredeweg B., Arroyo I., Carney C., Mavrikis M., Timms M.: Intelligent environments Global Resources for Online Education Workshop, Brighton, UK (2009)
16. Bulathwela S., Pérez-Ortiz M., Holloway C., Shawe-Taylor J.: Could AI Democratise Education? Socio-Technical Imaginaries of an EdTech Revolution, http://arxiv.org/abs/2112.02034, (2021)
17. Bulathwela S., Pérez-Ortiz M., Holloway C., Cukurova M., Shawe-Taylor J.: Artificial Intelligence Alone Will Not Democratise Education: On Educational Inequality, Techno-Solutionism and Inclusive Tools Sustainability, 16, pp. 781 (2024) https://doi.org/10.3390/su16020781
18. Carnevale A.P., Strohl J., Ridley N., Gulish A.: Three Educational Pathways to Good Jobs: High School, Middle Skills, and Bachelor's Degree (2018)
19. Choi J.H., Garrod O., Atherton P., Joyce-Gibbons A., Mason-Sesay M., Björkegren D.: Are LLMs Useful in the Poorest Schools? TheTeacher.AI in Sierra Leone, http://arxiv.org/abs/2310.02982, (2024)
20. Chun J., Barley N.: A Comparative Analysis of Multiple-Choice Questions: ChatGPT-Generated Items vs. Human-Developed Items in Chapelle, C., Beckett, G., and Ranalli, J. (eds.) Exploring AI in Applied Linguistics. pp. 118–136. Iowa State University Digital Press (2024)
21. Clancey W.J.: Knowledge-based tutoring: the GUIDON program, MIT Press, Cambridge, MA, USA, (1987) https://doi.org/10.31274/isudp.2024.154.08
22. Clarke D., Roche A.: Using contextualized tasks to engage students in meaningful and worthwhile mathematics learning The Journal of Mathematical Behavior, 51, pp. 95–108 (2018) https://doi.org/10.1016/j.jmathb.2017.11.006
23. Cobb P., Yackel E., Wood T.: A Constructivist Alternative to the Representational View of Mind in Mathematics Education (1992) https://doi.org/10.2307/749161
24. D'Mello S.K., Bosch N., Chen H.: Multimodal-multisensor affect detection The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition - Volume 2. vol. 21. pp. 167–202. Association for Computing Machinery and Morgan & Claypool (2018) https://doi.org/10.1145/3107990.3107998
25. D'Mello S., Lehman B., Sullins J., Daigle R., Combs R., Vogt K., Perkins L., Graesser A.: A Time for Emoting: When Affect-Sensitivity Is and Isn't Effective at Promoting Deep Learning in Aleven, V., Kay, J., and Mostow, J. (eds.) Intelligent Tutoring Systems. pp. 245–254. Springer, Berlin, Heidelberg (2010) https://doi.org/10.1007/978-3-642-13388-6_29
26. Dede C.: Immersive Interfaces for Engagement and Learning Science, 323, pp. 66–69 (2009) https://doi.org/10.1126/science.1167311
27. Delgado K., Origgi J.M., Hasanpoor T., Yu H., Allessio D., Arroyo I., Lee W., Betke M., Woolf B., Bargal S.A.: Student Engagement Dataset Presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision (2021) https://doi.org/10.1109/ICCVW54120.2021.00405
28. Digital S.: Global Education Monitoring Report 2020: Inclusion and education: All means all, https://resource-centre.savethechildren.net/document/global-education-monitoring-report-2020-inclusion-and-education-all-means-all/
29. Engler A.: Five ways that Open Source Software shapes AI policy, https://www.brookings.edu/articles/five-ways-that-open-source-software-shapes-ai-policy/
30. Ernst E., Merola R., Samaan D.: Sciendo IZA Journal of Labor Policy, 9, (2019) https://doi.org/10.2478/izajolp-2019-0004

31. Feng W., Lee J., McNichols H., Scarlatos A., Smith D., Woodhead S., Ornelas N.O., Lan A.: Exploring Automated Distractor Generation for Math Multiple-choice Questions via Large Language Models, http://arxiv.org/abs/2404.02124, (2024) https://doi.org/10.18653/v1/2024.findings-naacl.193
32. Fernandez N., Ghosh A., Liu N., Wang Z., Choffin B., Baraniuk R., Lan A.: Automated Scoring for Reading Comprehension via In-context BERT Tuning, http://arxiv.org/abs/2205.09864, (2023) https://doi.org/10.1007/978-3-031-11644-5_69
33. Foglia L., Wilson R.A.: Embodied cognition Wiley Interdiscip Rev Cogn Sci, 4, pp. 319–325 (2013) https://doi.org/10.1002/wcs.1226
34. Gao X., Li P., Shen J., Sun H.: Reviewing assessment of student learning in interdisciplinary STEM education International Journal of STEM Education, 7, pp. 24 (2020) https://doi.org/10.1186/s40594-020-00225-4
35. Gattupalli S., Castro F., Arroyo I.: Investigating Cultural Signatures in Game Designs of Middle School Students from Argentina, India, and the United States Presented at the June (2024) https://doi.org/10.22318/icls2024.138871
36. Gattupalli S., Lee W., Allessio D., Crabtree D., Arroyo I., Woolf B.: Exploring Pre-Service Teachers' Perceptions of Large Language Models-Generated Hints in Online Mathematics Learning (2023)
37. Georgieva K.: AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity IMF Blog (blog), International Monetary Fund, (2024)
38. Giovannella C.: "Learning by being": integrated thinking and competencies to mark the difference from AIs, Interaction Design and Architecture(s) IxD&A, 57, pp. 8–26 (2023) https://doi.org/10.55612/s-5002-057-001
39. Goodell J., Heffernan N.: Human-Centered Learning Engineering for the Emerging Intelligence Augmentation Economy Presented at the November (2022) https://doi.org/10.1007/978-3-031-19682-9_78
40. Henry Carrick: Will Computers remove the need for teachers in the future? (2015)
41. Hollan J.D., Hutchins E.L., Weitzman L.: STEAMER: An Interactive Inspectable Simulation-Based Training System AI Magazine, 5, pp. 15–15 (1984)
42. Hornecker E., Buur J.: Getting a grip on tangible interaction: a framework on physical space and social interaction Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 437–446. Association for Computing Machinery, New York, NY, USA (2006) https://doi.org/10.1007/978-3-031-19682-9_78
43. Huber C.: Matthew 25: Prayer for children with disabilities, https://www.worldvision.org/child-protection-news-stories/matthew-25-pray-for-children-with-disabilities, (2024)
44. Inkson K.: Understanding Careers: The Metaphors of Working Lives, SAGE Publications, Inc., Thousand Oaks, California, (2007) https://doi.org/10.4135/9781483328867
45. Joshi A., Allessio D., Magee J., Whitehill J., Arroyo I., Woolf B., Sclaroff S., Betke M.: Affect-driven Learning Outcomes Prediction in Intelligent Tutoring Systems 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019). pp. 1–5 (2019) https://doi.org/10.1109/FG.2019.8756624
46. Kamalov F., Santandreu Calonge D., Gurrib I.: New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution Sustainability, 15, pp. 12451 (2023) https://doi.org/10.3390/su151612451
47. Kapoor A., Burleson W., Picard R.W.: Automatic prediction of frustration International Journal of Human-Computer Studies, 65, pp. 724–736 (2007) https://doi.org/10.1016/j.ijhcs.2007.02.003
48. Karumbaiah S., Lizarralde R., Allessio D., Woolf B., Arroyo I., Wixon N.: Addressing Student Behavior and Affect with Empathy and Growth Mindset. International Educational Data Mining Society, (2017)

49. King J., Sabelli N., Kelly H.: Preamble on Policy Issues Global Resources for Online Education (GROE), Tempe Arizona, pp. 23–26 (2009)
50. Kucak D., Juricic V., Dambic G.: Machine Learning in Education - a Survey of Current Research Trends in Katalinic, B. (ed.) DAAAM Proceedings. vol. 1. pp. 0406–0410. DAAAM International Vienna (2018) https://doi.org/10.2507/29th.daaam.proceedings.059
51. Kulik J.A., Fletcher J.D.: Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review Review of Educational Research, 86, pp. 42–78 (2016) https://doi.org/10.3102/0034654315581420
52. Kweilin Ellingrud, Saurabh Sanghvi, Gurneet Singh Dandona, Anu Madgavkar, Michael Chui, Olivia White, Paige Hasebe: Generative AI and the future of work in America | McKinsey, McKinsey Global Institue, (2023)
53. Lane H.C., Cahill C., Foutz S., Auerbach D., Noren D., Lussenhop C., Swartout W.: The Effects of a Pedagogical Agent for Informal Science Education on Learner Behaviors and Self-efficacy in Lane, H.C., Yacef, K., Mostow, J., and Pavlik, P. (eds.) Artificial Intelligence in Education. pp. 309–318. Springer, Berlin, Heidelberg (2013) https://doi.org/10.1007/978-3-642-39112-5_32
54. Lane H.C., Hays M.J., Core M.G., Auerbach D.: Learning intercultural communication skills with virtual humans: Feedback and fidelity. Journal of Educational Psychology, 105, pp. 1026–1035 (2013) https://doi.org/10.1037/a0031506
55. Laurillard D., Kolokitha M., Mellar H., Selwyn N., Noss R.: Learning through life: The role of technology (2008)
56. Lee C.: AI Plagairism Changers: How Academic Leaders can Prepare Institutions, (2023)
57. Lee W., Allessio D., Rebelsky W., Satish Gattupalli S., Yu H., Arroyo I., Betke M., Bargal S., Murray T., Sylvia F., Woolf B.P.: Measurements and Interventions to Improve Student Engagement Through Facial Expression Recognition Adaptive Instructional Systems: 4th International Conference, AIS 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings. pp. 286–301. Springer-Verlag, Berlin, Heidelberg (2022) https://doi.org/10.1007/978-3-031-05887-5_20
58. Lester J.C., Ha E.Y., Lee S.Y., Mott B.W., Rowe J.P., Sabourin J.L.: Serious Games Get Smart: Intelligent Game-Based Learning Environments AI Magazine, 34, pp. 31–45 (2013) https://doi.org/10.1609/aimag.v34i4.2488
59. Levonian Z., Li C., Zhu W., Gade A., Henkel O., Postle M.-E., Xing W.: Retrieval-augmented Generation to Improve Math Question-Answering: Trade-offs Between Groundedness and Human Preference, http://arxiv.org/abs/2310.03184, (2023)
60. Liang Y., Wang J., Zhu H., Wang L., Qian W., Lan Y.: Prompting large language models with chain-of-thought for few-shot knowledge base question generation arXiv preprint arXiv:2310.08395, (2023) https://doi.org/10.18653/v1/2023.emnlp-main.263
61. Lindgren R., Johnson-Glenberg M.: Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality Educational Researcher, 42, pp. 445–452 (2013) https://doi.org/10.3102/0013189X13511661
62. Liu N., Wang Z., Baraniuk R., Lan A.: Open-ended Knowledge Tracing for Computer Science Education Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. pp. 3849–3862. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (2022) https://doi.org/10.18653/v1/2022.emnlp-main.254
63. Macedonia M.: Embodied Learning: Why at School the Mind Needs the Body Front. Psychol., 10, (2019) https://doi.org/10.3389/fpsyg.2019.02098
64. Maghsudi S., Lan A., Xu J., Schaar M. van der: Personalized Education in the AI Era: What to Expect Next?, http://arxiv.org/abs/2101.10074, (2021) https://doi.org/10.1109/MSP.2021.3055032

65. Mani G., Chen F., Cross S., Kalil T., Gopalakrishnan V., Rossi F., Stanley K.: Artificial Intelligence's Grand Challenges: Past, Present, and Future AI Magazine, 42, pp. 61–75 (2021) https://doi.org/10.1002/j.2371-9621.2021.tb00012.x
66. Matsuda M.: COMMENTARY: We must be proactive in guiding the influence of artificial intelligence on education, https://edsource.org/2024/we-must-be-proactive-in-guiding-the-influence-of-artificial-intelligence-on-education/708898
67. Maurer R.: AI Skills Will Be Crucial for Job Seekers, https://www.shrm.org/topics-tools/news/talent-acquisition/ai-skills-will-be-crucial-job-seekers
68. McArthur D., Lewis M., Bishary M.: The Roles of Artificial Intelligence in Education: Current Progress and Future Prospects Journal of Educational Technology, 1, pp. 42–80 (2005) https://doi.org/10.26634/jet.1.4.972
69. Meyer J., Jansen T., Schiller R., Liebenow L.W., Steinbach M., Horbach A., Fleckenstein J.: Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions Computers and Education: Artificial Intelligence, 6, pp. 100199 (2024) https://doi.org/10.1016/j.caeai.2023.100199
70. Mostow J.: Foreword What is AI? And What Does It Have to Do with Software Engineering? IEEE Transactions on Software Engineering, SE-11, pp. 1253–1256 (1985) https://doi.org/10.1109/TSE.1985.231876
71. Nathan M.J., Williams-Pierce C., Walkington C., Abrahamson D., Ottmar E., Soto H., Alibali M.W.: Embodied Design for Mathematical Imagination and Cognition (2021)
72. Negi A., Ghildiyal B.: Education 2030: A Roadmap to E-Learning Excellence International Conference on Knowledge Management in Higher Education Institutions (ICKHI 2024) (2024)
73. Nemirovsky R., Tierney C., Wright T.: Body Motion and Graphing Cognition and Instruction, 16, pp. 119–172 (1998) https://doi.org/10.1207/s1532690xci1602_1
74. Novack M.A., Goldin-Meadow S., Woodward A.L.: Learning from gesture: How early does it happen? Cognition, 142, pp. 138 (2015) https://doi.org/10.1016/j.cognition.2015.05.018
75. Nwana H.S.: Intelligent tutoring systems: an overview Artif Intell Rev, 4, pp. 251–277 (1990) https://doi.org/10.1007/BF00168958
76. Nye B.D., Mee D., Core M.G.: Generative Large Language Models for Dialog-Based Tutoring: An Early Consideration of Opportunities and Concerns (2023)
77. Oblinger D.: Game Changers: Education and Information Technologies, (2012)
78. Pamuk Z.: Artificial intelligence and the problem of judgment Ethics and International Affairs, 37, pp. 232–243 (2023) https://doi.org/10.1017/S089267942300014X
79. Phillips A.J.: Transnationalism and the Dominican Republic: The Effect on Student Identity and Achievement
80. Qawqzeh Y.: Exploring the Influence of Student Interaction with ChatGPT on Critical Thinking, Problem Solving, and Creativity IJIET, 14, pp. 596–601 (2024) https://doi.org/10.18178/ijiet.2024.14.4.2082
81. Rasul I., Castro F., Arroyo I.: Towards Embodied Wearable Intelligent Tutoring Systems Augmented Intelligence and Intelligent Tutoring Systems: 19th International Conference, ITS 2023, Corfu, Greece, June 2–5, 2023, Proceedings. pp. 298–306. Springer-Verlag, Berlin, Heidelberg (2023) https://doi.org/10.1007/978-3-031-32883-1_26
82. Regan P.M., Jesse J.: Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking Ethics and Information Technology, 21, pp. 167–179 (2019) https://doi.org/10.1007/s10676-018-9492-2
83. Robinson L.: ChatGPT Can Strengthen Student Writing, https://www.languagemagazine.com/2023/01/25/ai-can-strengthen-student-writing-not-weaken-it/, (2023)

84. Ruiz N., Yu H., Allessio D.A., Jalal M., Joshi A., Murray T., Magee J.J., Delgado K., Ablavsky V., Sclaroff S., Arroyo I., Woolf B.P., Bargal S.A., Betke M.: ATL-BP: a student engagement dataset and model for affect transfer learning for behavior prediction (2022) https://doi.org/10.1109/TBIOM.2022.3210479
85. Ruiz N., Yu H., Allessio D.A., Jalal M., Joshi A., Murray T., Magee J.J., Whitehill J.R., Ablavsky V., Arroyo I., Woolf B.P., Sclaroff S., Betke M.: Leveraging Affect Transfer Learning for Behavior Prediction in an Intelligent Tutoring System, http://arxiv.org/abs/2002.05242, (2021) https://doi.org/10.1109/FG52635.2021.9667001
86. Sachan M., Dubey A., Hovy E.H., Mitchell T.M., Roth D., Xing E.P.: Discourse in Multimedia: A Case Study in Extracting Geometry Knowledge from Textbooks Computational Linguistics, 45, pp. 627–665 (2020) https://doi.org/10.1162/coli_a_00360
87. Scarlatos A., Lan A.: Tree-Based Representation and Generation of Natural and Mathematical Language in Rogers, A., Boyd-Graber, J., and Okazaki, N. (eds.) Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 3714–3730. Association for Computational Linguistics, Toronto, Canada (2023) https://doi.org/10.18653/v1/2023.acl-long.205
88. Seo K., Tang J., Roll I., Fels S., Yoon D.: The impact of artificial intelligence on learner–instructor interaction in online learning International Journal of Educational Technology in Higher Education, 18, pp. 54 (2021) https://doi.org/10.1186/s41239-021-00292-9
89. Shute V.J.: The SAGE Encyclopedia of Educational Technology (2015)
90. Stafford D., Flatley R.: OpenStax, https://annurev.publisher.ingentaconnect.com/contentone/annurev/tca/2018/00000020/00000001/art00012
91. Swartout W., Artstein R., Forbell E., Foutz S., Lane H.C., Lange B., Morie J., Noren D., Rizzo S., Traum D.: Virtual humans for learning AI Magazine, 34, pp. 13–30 (2013) https://doi.org/10.1609/aimag.v34i4.2487
92. Toshniwal S., Moshkov I., Narenthiran S., Gitman D., Jia F., Gitman I.: OpenMathInstruct-1: A 1.8 Million Math Instruction Tuning Dataset, http://arxiv.org/abs/2402.10176, (2024)
93. Tsiros A., Palladini A.: Towards a Human-Centric Design Framework for AI Assisted Music Production Presented at the Proceedings of the International Conference on New Interfaces for Musical Expression (2020)
94. UNESCO: Global Education Monitoring Report 2020: Inclusion and Education-All Means All, (2020)
95. VanLehn K., Corbett A., Ramachandran S., Underwood J., Jensen C.: Intelligent Virtual Environments Global Resources for Online Education (GROE), Tempe Arizona, pp. 23–26 (2009)
96. VanLEHN K.: The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems Educational Psychologist, 46, pp. 197–221 (2011) https://doi.org/10.1080/00461520.2011.611369
97. VanLehn K., Zhang L., Burleson W., Girard S., Hidago-Pontet Y.: Can a Non-Cognitive Learning Companion Increase the Effectiveness of a Meta-Cognitive Learning Strategy? IEEE Transactions on Learning Technologies, 10, pp. 277–289 (2017) https://doi.org/10.1109/TLT.2016.2594775
98. Wang S., Xu T., Li H., Zhang C., Liang J., Tang J., Yu P.S., Wen Q.: Large Language Models for Education: A Survey and Outlook, http://arxiv.org/abs/2403.18105, (2024)
99. Wang Y., Heffernan N.T.: The “assistance” model: Leveraging how many hints and attempts a student needs Twenty-fourth international FLAIRS conference (2011)
100. Wang Z., Lan A., Baraniuk R.: Math Word Problem Generation with Mathematical Consistency and Problem Context Constraints in Moens, M.-F., Huang, X., Specia, L., and Yih, S.W. (eds.) Proceedings of the 2021 Conference on Empirical Methods in

Natural Language Processing. pp. 5986–5999. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic (2021) https://doi.org/10.18653/v1/2021.emnlp-main.484
101. Weber-Wulff D., Anohina-Naumeca A., Bjelobaba S., Foltýnek T., Guerrero-Dib J., Popoola O., Šigut P., Waddington L.: Testing of detection tools for AI-generated text Int J Educ Integr, 19, pp. 1–39 (2023)
102. White B.Y., Frederiksen J.R.: QUEST: Qualitative understanding of electrical system troubleshooting, BBN Laboratories, (1985)
103. Willms J.D.: Student engagement at school A sense of belonging and participation. Paris: Organisation for Economic Co-operation and Development, pp. 1–84 (2003)
104. Woolf B.P., Arroyo I., Muldner K., Burleson W., Cooper D.G., Dolan R., Christopherson R.M.: The Effect of Motivational Learning Companions on Low Achieving Students and Students with Disabilities in Aleven, V., Kay, J., and Mostow, J. (eds.) Intelligent Tutoring Systems. pp. 327–337. Springer, Berlin, Heidelberg (2010) https://doi.org/10.1007/978-3-642-13388-6_37
105. Xiao C., Ma W., Song Q., Xu S.X., Zhang K., Wang Y., Fu Q.: Human-AI Collaborative Essay Scoring: A Dual-Process Framework with LLMs, http://arxiv.org/abs/2401.06431, (2024)
106. Xiao C., Xu S.X., Zhang K., Wang Y., Xia L.: Evaluating Reading Comprehension Exercises Generated by LLMs: A Showcase of ChatGPT in Education Applications in Kochmar, E., Burstein, J., Horbach, A., Laarmann-Quante, R., Madnani, N., Tack, A., Yaneva, V., Yuan, Z., and Zesch, T. (eds.) Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023). pp. 610–625. Association for Computational Linguistics, Toronto, Canada (2023) https://doi.org/10.18653/v1/2023.bea-1.52
107. Yağcı M.: Educational data mining: prediction of students' academic performance using machine learning algorithms Smart Learning Environments, 9, pp. 11 (2022) https://doi.org/10.1186/s40561-022-00192-z
108. Yu H., Allessio D.A., Lee W., Rebelsky W., Sylvia F., Murray T., Magee J.J., Arroyo I., Woolf B.P., Bargal S.A., Betke M.: COVES: A Cognitive-Affective Deep Model that Personalizes Math Problem Difficulty in Real Time and Improves Student Engagement with an Online Tutor Proceedings of the 31st ACM International Conference on Multimedia. pp. 6152–6160. Association for Computing Machinery, New York, NY, USA (2023) https://doi.org/10.1145/3581783.3613965
109. Yu H., Gupta A., Lee W., Arroyo I., Betke M., Allesio D., Murray T., Magee J., Woolf B.P.: Measuring and Integrating Facial Expressions and Head Pose as Indicators of Engagement and Affect in Tutoring Systems Adaptive Instructional Systems. Adaptation Strategies and Methods: Third International Conference, AIS 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part II. pp. 219–233. Springer-Verlag, Berlin, Heidelberg (2021) https://doi.org/10.1145/3581783.3613965
110. Zellers R., Holtzman A., Bisk Y., Farhadi A., Choi Y.: HellaSwag: Can a Machine Really Finish Your Sentence?, http://arxiv.org/abs/1905.07830, (2019) https://doi.org/10.18653/v1/P19-1472
111. Zhang M., Baral S., Heffernan N., Lan A.: Automatic Short Math Answer Grading via In-context Meta-learning Presented at the Proceedings of the 15th International Conference on Educational Data Mining (2022)
112. Zhong Y., Jiang C., Xu W., Li J.J.: Discourse Level Factors for Sentence Deletion in Text Simplification Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp. 9709–9716 (2020) https://doi.org/10.1609/aaai.v34i05.6520
113. Zhou E., Lee D.: Generative artificial intelligence, human creativity, and art PNAS Nexus, 3, pp. pgae052 (2024) https://doi.org/10.1093/pnasnexus/pgae052

114. Zirar A., Ali S.I., Islam N.: Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda Technovation, 124, pp. 102747 (2023) https://doi.org/10.1016/j.technovation.2023.102747
115. Redish E.F., Kuo E.: Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology Sci & Educ, 24, pp. 561–590 (2015) https://doi.org/10.1007/s11191-015-9749-7
116. Porter S.G., Greene K., Esposito M.K.: Access and inclusion of students with disabilities in virtual learning environments: Implications for post-pandemic teaching International Journal of Multicultural Education, 23, pp. 43–61 (2021) https://doi.org/10.18251/ijme.v23i3.3011