



Welcome to ChatGPT your OpenAI account

Login

SignUp

SHUTTERSTOCK



chat.openai.com



AI in Education

The leap into a new era of machine intelligence carries risks and challenges, but also plenty of promise

IN NEAL STEPHENSON'S 1995 science fiction novel, *The Diamond Age*, readers meet Nell, a young girl who comes into possession of a highly advanced book, *The Young Lady's Illustrated Primer*. The book is not the usual static collection of texts and images but a deeply immersive tool that can converse with the reader, answer questions, and personalize its content, all in service of educating and motivating a young girl to be a strong, independent individual.

Such a device, even after the introduction of the Internet and tablet computers, has

remained in the realm of science fiction—until now. Artificial intelligence, or AI, took a giant leap forward with the introduction in November 2022 of ChatGPT, an AI technology capable of producing remarkably creative responses and sophisticated analysis through human-like dialogue. It has triggered a wave of innovation, some of which suggests we might be on the brink of an era of interactive, super-intelligent tools not unlike the book Sundar Pichai, Google's CEO, calls artificial intelligence “more profound than fire or electricity or anything we have done in the past.” Reid Hoffman,

By **JOHN BAILEY**

the founder of LinkedIn and current partner at Greylock Partners, says, “The power to make positive change in the world is about to get the biggest boost it’s ever had.” And Bill Gates has said that “this new wave of AI is as fundamental as the creation of the microprocessor, the personal computer, the Internet, and the mobile phone.”

Over the last year, developers have released a dizzying array of AI tools that can generate text, images, music, and video with no need for complicated coding but simply in response to instructions given in natural language. These technologies are rapidly improving, and developers are introducing capabilities that would have been considered science fiction just a few years ago. AI is also raising pressing ethical questions around bias, appropriate use, and plagiarism.

In the realm of education, this technology will influence how students learn, how teachers work, and ultimately how we structure our education system. Some educators and leaders look forward to these changes with great enthusiasm. Sal Kahn, founder of Khan Academy, went so far as to say in a TED talk that AI has the potential to effect “probably the biggest positive transformation that education has ever seen.” But others warn that AI will enable the spread of misinformation, facilitate cheating in school and college, kill whatever vestiges of individual privacy remain, and cause massive job loss. The challenge is to harness the positive potential while avoiding or mitigating the harm.

What Is Generative AI?

Artificial intelligence is a branch of computer science that focuses on creating software capable of mimicking behaviors and processes we would consider “intelligent” if exhibited by humans, including reasoning, learning, problem-solving, and exercising creativity. AI systems can be applied to an extensive range of tasks, including language translation, image recognition, navigating autonomous vehicles, detecting and treating cancer, and, in the case of generative AI, producing content and knowledge rather than simply searching for and retrieving it.

“Foundation models” in generative AI are systems trained on a large dataset to learn a broad base of knowledge that can then be adapted to a range of different, more specific purposes. This learning method is self-supervised, meaning the model learns by finding patterns and relationships in the data it is trained on.

Large Language Models (LLMs) are foundation models that have been trained on a vast amount of text data. For example, the training data for OpenAI’s GPT model consisted of web content, books, Wikipedia articles, news articles, social media posts, code snippets, and more. OpenAI’s GPT-3 models underwent training on a staggering 300 billion “tokens” or word pieces, using more than 175 billion parameters to shape the model’s behavior—nearly

100 times more data than the company’s GPT-2 model had.

By doing this analysis across billions of sentences, LLM models develop a statistical understanding of language: how words and phrases are usually combined, what topics are typically discussed together, and what tone or style is appropriate in different contexts. That allows it to generate human-like text and perform a wide range of tasks, such as writing articles, answering questions, or analyzing unstructured data.

LLMs include OpenAI’s GPT-4, Google’s PaLM, and Meta’s LLaMA. These LLMs serve as “foundations” for AI applications. ChatGPT is built on GPT-3.5 and GPT-4, while Bard uses Google’s Pathways Language Model 2 (PaLM 2) as its foundation.

Some of the best-known applications are:

ChatGPT 3.5. The free version of ChatGPT released by OpenAI in November 2022. It was trained on data only up to 2021, and while it is very fast, it is prone to inaccuracies.

ChatGPT 4.0. The newest version of ChatGPT, which is more powerful and accurate than ChatGPT 3.5 but also slower, and it requires a paid account. It also has extended capabilities through plug-ins that give it the ability to interface with content from websites, perform more sophisticated mathematical functions, and access other services. A new Code Interpreter feature gives ChatGPT the ability to analyze data, create charts, solve math problems, edit files, and even develop hypotheses to explain data trends.

Microsoft Bing Chat. An iteration of Microsoft’s Bing search engine that is enhanced with OpenAI’s ChatGPT technology. It can browse websites and offers source citations with its results.

Google Bard. Google’s AI generates text, translates languages, writes different kinds of creative content, and writes and debugs code in more than 20 different programming languages. The tone and style of Bard’s replies can be finetuned to be simple, long, short, professional, or casual. Bard also leverages Google Lens to analyze images uploaded with prompts.

Anthropic Claude 2. A chatbot that can generate text, summarize content, and perform other tasks, Claude 2 can analyze texts of roughly 75,000 words—about the length of *The Great Gatsby*—and generate responses of more than 3,000 words. The model was built using a set of principles that serve as a sort of “constitution” for AI systems, with the aim of making them more helpful, honest, and harmless.

These AI systems have been improving at a remarkable pace, including in how well they perform on assessments of human knowledge. OpenAI’s GPT-3.5, which was released in March 2022, only managed to score in the 10th percentile on the bar exam, but GPT-4.0, introduced a year later, made a significant leap, scoring in the 90th percentile. What makes these feats especially impressive is that OpenAI did not specifically train the system to take these exams; the AI was able to come up with the correct answers on its own. Similarly, Google’s medical AI model substantially improved its performance on a U.S. Medical Licensing Examination practice test, with its

accuracy rate jumping to 85 percent in March 2021 from 33 percent in December 2020.

These two examples prompt one to ask: if AI continues to improve so rapidly, what will these systems be able to achieve in the next few years? What's more, new studies challenge the assumption that AI-generated responses are stale or sterile. In the case of Google's AI model, physicians preferred the AI's long-form answers to those written by their fellow doctors, and nonmedical study participants rated the AI answers as more helpful. Another study found that participants preferred a medical chatbot's responses over those of a physician and rated them significantly higher, not just for quality but also for *empathy*. What will happen when "empathetic" AI is used in education?

Other studies have looked at the reasoning capabilities of these models. Microsoft researchers suggest that newer systems "exhibit more general intelligence than previous AI models" and are coming "strikingly close to human-level performance." While some observers question those conclusions, the AI systems display an increasing ability to generate coherent and

relevant sources for possible answers. The search engine (or librarian) doesn't create anything new but efficiently retrieves what's already there.

Generative AI is more akin to a competent intern. You give a generative AI tool instructions through prompts, as you would to an intern, asking it to complete a task and produce a product. The AI interprets your instructions, thinks about the best way to carry them out, and produces something original or performs a task to fulfill your directive. The results aren't pre-made or stored somewhere—they're produced on the fly, based on the information the intern (generative AI) has been trained on. The output often depends on the precision and clarity of the instructions (prompts) you provide. A vague or poorly defined prompt might lead the AI to produce less relevant results. The more context and direction you give it, the better the result will be. What's more, the capabilities of these AI systems are being enhanced through the introduction of versatile plug-ins that equip them to browse websites, analyze data files, or access other services. Think of this as giving your intern access to

Sal Kahn, founder of Khan Academy, said in a TED talk that AI has the potential to effect "probably the biggest positive transformation that education has ever seen."

contextually appropriate responses, make connections between different pieces of information, and engage in reasoning processes such as inference, deduction, and analogy.

Despite their prodigious capabilities, these systems are not without flaws. At times, they churn out information that might sound convincing but is irrelevant, illogical, or entirely false—an anomaly known as "hallucination." The execution of certain mathematical operations presents another area of difficulty for AI. And while these systems can generate well-crafted and realistic text, understanding why the model made specific decisions or predictions can be challenging.

The Importance of Well-Designed Prompts

Using generative AI systems such as ChatGPT, Bard, and Claude 2 is relatively simple. One has only to type in a request or a task (called a prompt), and the AI generates a response. Properly constructed prompts are essential for getting useful results from generative AI tools. You can ask generative AI to analyze text, find patterns in data, compare opposing arguments, and summarize an article in different ways.

One challenge is that, after using search engines for years, people have been preconditioned to phrase questions in a certain way. A search engine is something like a helpful librarian who takes a specific question and points you to the most

a group of experts to help accomplish your tasks.

One strategy in using a generative AI tool is first to tell it what kind of expert or persona you want it to "be." Ask it to be an expert management consultant, a skilled teacher, a writing tutor, or a copy editor, and then give it a task.

Prompts can also be constructed to get these AI systems to perform complex and multi-step operations. For example, let's say a teacher wants to create an adaptive tutoring program—for any subject, any grade, in any language—that customizes the examples for students based on their interests. She wants each lesson to culminate in a short-response or multiple-choice quiz. If the student answers the questions correctly, the AI tutor should move on to the next lesson. If the student responds incorrectly, the AI should explain the concept again, but using simpler language.

Previously, designing this kind of interactive system would have required a relatively sophisticated and expensive software program. With ChatGPT, however, just giving those instructions in a prompt delivers a serviceable tutoring system. It isn't perfect, but remember that it was built virtually for free, with just a few lines of English language as a command. And nothing in the education market today has the capability to generate almost limitless examples to connect the lesson concept to students' interests.

Chained prompts can also help focus AI systems. For

example, an educator can prompt a generative AI system first to read a practice guide from the What Works Clearinghouse and summarize its recommendations. Then, in a follow-up prompt, the teacher can ask the AI to develop a set of classroom activities based on what it just read. By curating the source material and using the right prompts, the educator can anchor the generated responses in evidence and high-quality research.

However, much like fledgling interns learning the ropes in a new environment, AI does commit occasional errors. Such fallibility, while inevitable, underlines the critical importance of maintaining rigorous oversight of AI's output. Monitoring

their peers or students. Early uses include automated routine tasks such as drafting lesson plans, creating differentiated materials, designing worksheets, developing quizzes, and exploring ways of explaining complicated academic materials. AI can also provide educators with recommendations to meet student needs and help teachers reflect, plan, and improve their practice.

Parent assistants. Parents can use AI to generate letters requesting individualized education plan (IEP) services or to ask that a child be evaluated for gifted and talented programs. For parents choosing a school for their child, AI could serve as an administrative assistant, mapping out school options within driving distance of home, generating application timelines, compiling

Using natural language, you can ask generative AI to analyze text, find patterns in data, compare opposing arguments, and summarize an article in different ways.

not only acts as a crucial checkpoint for accuracy but also becomes a vital source of real-time feedback for the system. It's through this iterative refinement process that an AI system, over time, can significantly minimize its error rate and increase its efficacy.

Uses of AI in Education

In May 2023, the U.S. Department of Education released a report titled *Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations*. The department had conducted listening sessions in 2022 with more than 700 people, including educators and parents, to gauge their views on AI. The report noted that “constituents believe that action is required now in order to get ahead of the expected increase of AI in education technology—and they want to roll up their sleeves and start working together.” People expressed anxiety about “future potential risks” with AI but also felt that “AI may enable achieving educational priorities in better ways, at scale, and with lower costs.”

AI could serve—or is already serving—in several teaching-and-learning roles:

Instructional assistants. AI's ability to conduct human-like conversations opens up possibilities for adaptive tutoring or instructional assistants that can help explain difficult concepts to students. AI-based feedback systems can offer constructive critiques on student writing, which can help students fine-tune their writing skills. Some research also suggests certain kinds of prompts can help children generate more fruitful questions about learning. AI models might also support customized learning for students with disabilities and provide translation for English language learners.

Teaching assistants. AI might tackle some of the administrative tasks that keep teachers from investing more time with

contact information, and the like. Generative AI can even create bedtime stories with evolving plots tailored to a child's interests.

Administrator assistants. Using generative AI, school administrators can draft various communications, including materials for parents, newsletters, and other community-engagement documents. AI systems can also help with the difficult tasks of organizing class or bus schedules, and they can analyze complex data to identify patterns or needs. ChatGPT can perform sophisticated sentiment analysis that could be useful for measuring school-climate and other survey data.

Though the potential is great, most teachers have yet to use these tools. A Morning Consult and EdChoice poll found that while 60 percent say they've heard about ChatGPT, only 14 percent have used it in their free time, and just 13 percent have used it at school. It's likely that most teachers and students will engage with generative AI not through the platforms themselves but rather through AI capabilities embedded in software. Instructional providers such as Khan Academy, Varsity Tutors, and DuoLingo are experimenting with GPT-4-powered tutors that are trained on datasets specific to these organizations to provide individualized learning support that has additional guardrails to help protect students and enhance the experience for teachers.

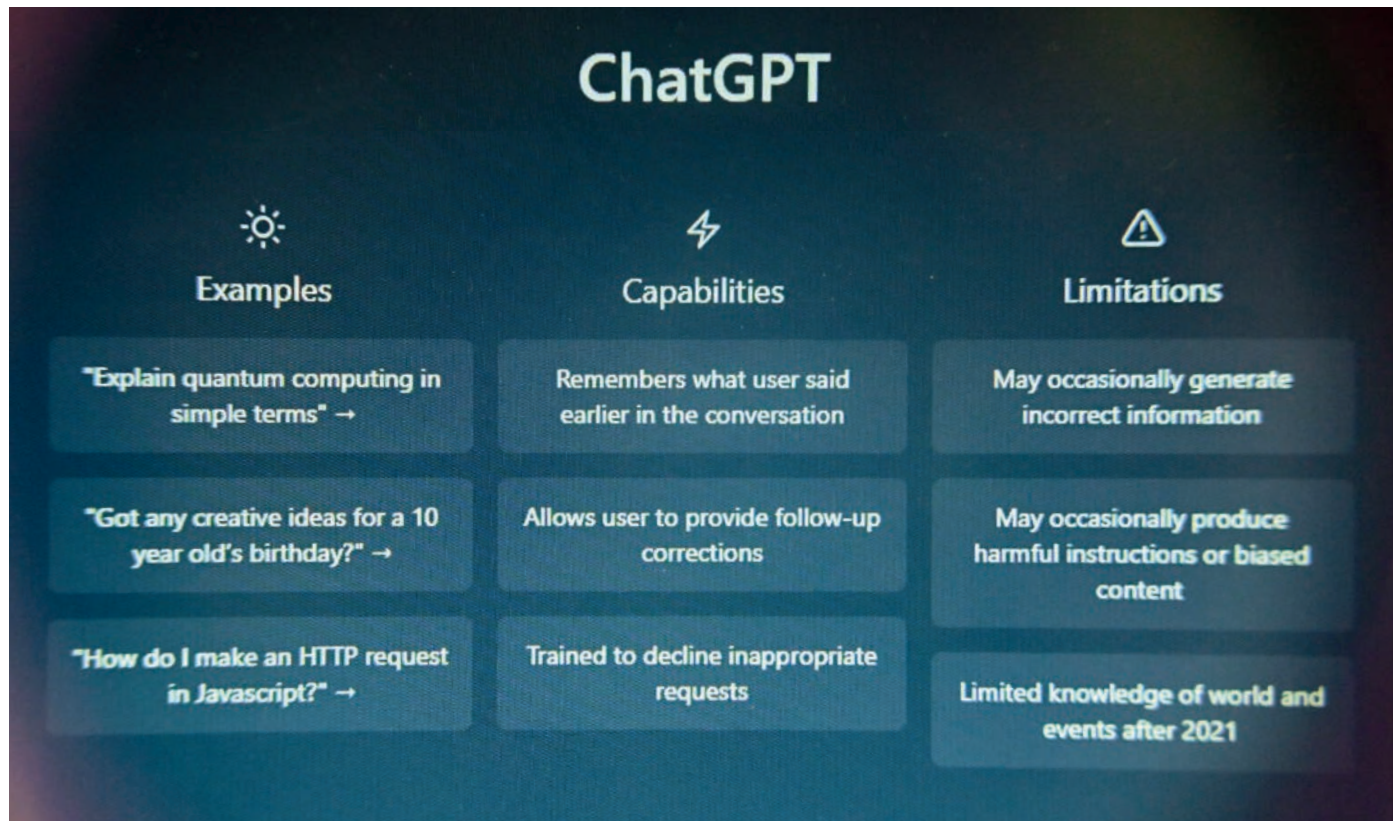
Google's Project Tailwind is experimenting with an AI notebook that can analyze student notes and then develop study questions or provide tutoring support through a chat interface. These features could soon be available on Google Classroom, potentially reaching over half of all U.S. classrooms. Brisk Teaching is one of the first companies to build a portfolio of AI services designed specifically for teachers—differentiating content, drafting lesson plans, providing student feedback, and serving as an AI assistant to streamline workflow among different apps and tools.

Providers of curriculum and instruction materials might also include AI assistants for instant help and tutoring tailored to the companies' products. One example is the edX Xpert, a ChatGPT-based learning assistant on the edX platform. It offers immediate, customized academic and customer support for online learners worldwide.

Regardless of the ways AI is used in classrooms, the funda-

the ethical issues involved in such cheating, students who use AI to do their work for them may not be learning the content and skills they need.

Bias in AI algorithms. AI systems learn from the data they are trained on. If this data contains biases, those biases can be learned and perpetuated by the AI system. For example, if the data include student-performance information that's biased



ADRIEN FILLON / ZUMA PRESS WIRE

The homescreen for OpenAI's foundation-model generative artificial intelligence, ChatGPT, gives users three sample commands and a list of functions and caveats. Introduced publicly in November 2022, ChatGPT can produce creative, human-like responses and analysis.

mental task of policymakers and education leaders is to ensure that the technology is serving sound instructional practice. As Vicki Phillips, CEO of the National Center on Education and the Economy, wrote, "We should not only think about how technology can assist teachers and learners in improving what they're doing now, but what it means for ensuring that new ways of teaching and learning flourish alongside the applications of AI."

Challenges and Risks

Along with these potential benefits come some difficult challenges and risks the education community must navigate:

Student cheating. Students might use AI to solve homework problems or take quizzes. AI-generated essays threaten to undermine learning as well as the college-entrance process. Aside from

toward one ethnicity, gender, or socioeconomic segment, the AI system could learn to favor students from that group. Less cited but still important are potential biases around political ideology and possibly even pedagogical philosophy that may generate responses not aligned to a community's values.

Privacy concerns. When students or educators interact with generative-AI tools, their conversations and personal information might be stored and analyzed, posing a risk to their privacy. With public AI systems, educators should refrain from inputting or exposing sensitive details about themselves, their colleagues, or their students, including but not limited to private communications, personally identifiable information, health records, academic performance, emotional well-being, and financial information.

Decreased social connection. There is a risk that more

time spent using AI systems will come at the cost of less student interaction with both educators and classmates. Children may also begin turning to these conversational AI systems in place of their friends. As a result, AI could intensify and worsen the public health crisis of loneliness, isolation, and lack of connection identified by the U.S. Surgeon General.

Overreliance on technology. Both teachers and students face the risk of becoming overly reliant on AI-driven technology. For students, this could stifle learning, especially the development of critical thinking. This challenge extends to educators as well. While AI can expedite lesson-plan generation, speed does not equate to quality. Teachers may be tempted to accept the initial AI-generated content rather than devote time to reviewing and refining it for optimal educational value.

Equity issues. Not all students have equal access to computer devices and the Internet. That imbalance could accelerate a widening of the achievement gap between students from different socioeconomic backgrounds.

Many of these risks are not new or unique to AI. Schools banned calculators and cellphones when these devices were first introduced, largely over concerns related to cheating. Privacy concerns around educational technology have led

Education's 2023 report describes the risks and opportunities posed by AI, but its recommendations amount to guidance at best. The White House released a Blueprint for an AI Bill of Rights, but it, too, is more an aspirational statement than a governing document. Congress is drafting legislation related to AI, which will help generate needed debate, but the path to the president's desk for signature is murky at best.

It is up to policymakers to establish clearer rules of the road and create a framework that provides consumer protections, builds public trust in AI systems, and establishes the regulatory certainty companies need for their product road maps. Considering the potential for AI to affect our economy, national security, and broader society, there is no time to waste.

Why AI Is Different

It is wise to be skeptical of new technologies that claim to revolutionize learning. In the past, prognosticators have promised that television, the computer, and the Internet, in turn, would transform education. Unfortunately, the heralded revolutions fell short of expectations.

There are some early signs, though, that this technological

It's likely that most teachers and students will engage with generative AI not through the platforms themselves but rather through AI capabilities embedded in software.

lawmakers to introduce hundreds of bills in state legislatures, and there are growing tensions between new technologies and existing federal privacy laws. The concerns over bias are understandable, but similar scrutiny is also warranted for existing content and materials that rarely, if ever, undergo review for racial or political bias.

In light of these challenges, the Department of Education has stressed the importance of keeping “humans in the loop” when using AI, particularly when the output might be used to inform a decision. As the department encouraged in its 2023 report, teachers, learners, and others need to retain their agency. AI cannot “replace a teacher, a guardian, or an education leader as the custodian of their students' learning,” the report stressed.

Policy Challenges with AI

Policymakers are grappling with several questions related to AI as they seek to strike a balance between supporting innovation and protecting the public interest (see sidebar). The speed of innovation in AI is outpacing many policymakers' understanding, let alone their ability to develop a consensus on the best ways to minimize the potential harms from AI while maximizing the benefits. The Department of

wave might be different in the benefits it brings to students, teachers, and parents. Previous technologies democratized access to content and resources, but AI is democratizing a kind of machine intelligence that can be used to perform a myriad of tasks. Moreover, these capabilities are open and affordable—nearly anyone with an Internet connection and a phone now has access to an intelligent assistant.

Generative AI models keep getting more powerful and are improving rapidly. The capabilities of these systems months or years from now will far exceed their current capacity. Their capabilities are also expanding through integration with other expert systems. Take math, for example. GPT-3.5 had some difficulties with certain basic mathematical concepts, but GPT-4 made significant improvement. Now, the incorporation of the Wolfram plug-in has nearly erased the remaining limitations.

It's reasonable to anticipate that these systems will become more potent, more accessible, and more affordable in the years ahead. The question, then, is how to use these emerging capabilities responsibly to improve teaching and learning.

The paradox of AI may lie in its potential to enhance the human, interpersonal element in education. Aaron Levie, CEO of Box, a Cloud-based content-management company,

Policy Matters

Officials and lawmakers must grapple with several questions related to AI to protect students and consumers and establish the rules of the road for companies. Key issues include:

Risk management framework: What is the optimal framework for assessing and managing AI risks? What specific requirements should be instituted for higher-risk applications? In education, for example, there is a difference between an AI system that generates a lesson sample and an AI system grading a test that will determine a student's admission to a school or program. There is growing support for using the AI Risk Management Framework from the U.S. Commerce Department's National Institute of Standards and Technology as a starting point for building trustworthiness into the design, development, use, and evaluation of AI products, services, and systems.

Licensing and certification: Should the United States require licensing and certification for AI models, systems, and applications? If so, what role could third-party audits and certifications play in assessing the safety and reliability of different AI systems? Schools and companies need to begin thinking about responsible AI practices to prepare for potential certification systems in the future.

Centralized vs. decentralized AI governance: Is it more effective to establish a central AI authority or agency, or

would it be preferable to allow individual sectors to manage their own AI-related issues? For example, regulating AI in autonomous vehicles is different from regulating AI in drug discovery or intelligent tutoring systems. Overly broad, one-size-fits-all frameworks and mandates may not work and could slow innovation in these sectors. In addition, it is not clear that many agencies have the authority or expertise to regulate AI systems in diverse sectors.

Privacy and content moderation: Many of the new AI systems pose significant new privacy questions and challenges. How should existing privacy and content-moderation frameworks, such as the Family Educational Rights and Privacy Act (FERPA), be adapted for AI, and which new policies or frameworks might be necessary to address unique challenges posed by AI?

Transparency and disclosure: What degree of transparency and disclosure should be required for AI models, particularly regarding the data they have been trained on? How can we develop comprehensive disclosure policies to ensure that users are aware when they are interacting with an AI service?

believes that AI will ultimately help us attend more quickly to those important tasks “that only a human can do.” Frederick Hess, director of education policy studies at the American Enterprise Institute, similarly asserts that “successful schools are inevitably the product of the relationships between adults and students. When technology ignores that, it's bound to disappoint. But when it's designed to offer more coaching, free up time for meaningful teacher-student interaction, or offer students more personalized feedback, technology can make a significant, positive difference.”

Technology does not revolutionize education; humans do. It is humans who create the systems and institutions that educate children, and it is the leaders of those systems who decide which tools to use and how to use them. Until those institutions modernize to accommodate the new possibilities of these technologies, we should expect incremental improvements at best. As Joel Rose, CEO of New Classrooms Innovation Partners, noted, “The most urgent need is for new and existing organizations to redesign the student experience

in ways that take full advantage of AI's capabilities.”

While past technologies have not lived up to hyped expectations, AI is not merely a continuation of the past; it is a leap into a new era of machine intelligence that we are only beginning to grasp. While the immediate implementation of these systems is imperfect, the swift pace of improvement holds promising prospects. The responsibility rests with human intervention—with educators, policymakers, and parents to incorporate this technology thoughtfully in a manner that optimally benefits teachers and learners. Our collective ambition should not focus solely or primarily on averting potential risks but rather on articulating a vision of the role AI should play in teaching and learning—a game plan that leverages the best of these technologies while preserving the best of human relationships.

John Bailey is a strategic adviser to entrepreneurs, policymakers, investors, and philanthropists and is a nonresident senior fellow at the American Enterprise Institute.