Generative AI and the future of education

Stefania Giannini

Assistant Director-General for Education
UNESCO

July 2023

The unrelenting pace of the digital revolution

The digital changes we are living through are thrilling, jarring, full of opportunity and, at the same time, terrifying.

Over the course of my career, I have witnessed at least four digital revolutions: the advent and proliferation of personal computers; the expansion of the internet and search; the rise and influence of social media; and the growing ubiquity of mobile computing and connectivity.

The sweeping changes brought by these revolutions can feel sudden and sometimes uninvited. They dramatically change the ways we live and how we teach and learn. Remarkably, many of us, and youth especially, now spend significantly more time immersed in digital spaces and interactions than in offline and offscreen exchanges – a proposition that seemed like science fiction just a generation ago. Developments with digital technology often seem only to accelerate, and the new worlds they create can feel unfamiliar and disorienting, even as we understand their potential to enrich our lives, improve our relationships, and open new horizons for education.

Not all people and not all countries have felt these recent technological revolutions in the same way, nor have they necessarily unfolded in a step-by-step progression. In many places, the mobile revolution has been the vehicle of personal computing, internet access, and social media – all four revolutions at once. A major disruption, however full of possibility.

Although most of us are all are still trying to come to terms with the sweeping social and educational implications of these earlier revolutions which are still unfolding, we have, in the

past several months, awoken to find ourselves abruptly entering yet another digital revolution – one which may make the others look minor by comparison. This is the AI revolution.

Language matters

Using improved computing power, synthetic neural networks, and large language modelling, AI technology is, if not cracking, at least feigning with remarkable dexterity the 'linchpin' of human civilization: *language*.

My formal academic training is in linguistics, so I have had ample opportunities to think about the structure, form, meaning, and power of language. Language matters. It is what distinguishes us from other animals. It is at the heart of identity and cultural diversity. It gives meaning to the world around us and inspires our actions. It is the basis of everything we do in education and in almost every other sphere of life. It lies at the root of love and of war. It can empower, and it can manipulate.

Until very recently, we had almost exclusive use and control of language. The fact that machines are now crossing so many language thresholds and so quickly should make us think and reflect. The processes that make these developments possible are important and deserve scrutiny, but their result is undeniable: machines can now simulate sophisticated conversation beyond narrow tasks.

We are coming to understand that our monopoly on advanced language – a natural ability, cultivated through education, and our species' most defining social trait – is no longer something we can take for granted. Recognizing this fact is forcing us to revisit the beliefs and assumptions that uphold our current education systems and, indeed, our wider societies.

AI applications that generate human-like language raise fundamental questions that concern education but spread far beyond: How will this technology change notions of who we are as humans? How will it reframe our understandings of human intelligence? How will it impact our relationships with each other?

We are also forced to consider the new technologies that study our languages and generate them, without explicit human direction and therefore unpredictably. Is it possible for technology that is proficient in language and learning to, at some point, develop sentience, knowledge of its own existence and desire greater autonomy? Is it wise to hand over millennia of knowledge to machines that appear to be capable of learning and performing beyond boundaries set by humans?

And what about our interactions with these machines: How should we 'treat' them? Is it appropriate for a non-human machine to speak to an adult as if it is another person? Is this appropriate for a child? What should we think when a chatbot assumes the voice of a living or long-dead historical figure on demand and without hesitation?

Implications for knowledge

Technology is never ideologically neutral. It exhibits and privileges certain worldviews and reflects particular ways of thinking and knowing. New generative AI models and utilities are no exception.

AI chatbots like ChatGPT enable a fundamentally different user experience than the AI technologies that support standard Google or other web searches. Search technology curates and ranks a menu of largely human-produced content in response to user queries. Large language model chatbots, by contrast, generate singular and, as such, much more authoritative-seeming responses using machine-produced content. AI chatbots function, therefore, like all-knowing oracles.

The answers provided by these AI chatbots do not trace to human minds. Rather, they stem from a maze of calculations so complex that it is not fully comprehensible even to the people who develop the technology. We have, in effect, an invention that gives human users singular responses to questions, but these responses cannot be traced to other people. Definitionally then, the responses lack humanity.

Machines that offer immediate, concise and seemingly definitive answers to knowledge questions can be helpful to learners, teachers, and others. But the technology can also usher in a world where machine knowledge becomes dominant, and proprietary AI models are elevated to global, and perhaps even revered, sources of authority. These models will project certain worldviews and ways of knowing and background others.

Despite the promises of AI and other digital technologies to further diversify our knowledge systems, we may be moving in the opposite direction. This is particularly true if just one or two AI models and platforms, some of them already exercising near monopolistic powers, come to assert even greater dominance over our interface with knowledge. As AI technology continues to permeate our world, we must preserve and safeguard the diversity of our knowledge systems and develop AI technologies in ways that will protect and expand our rich knowledge commons. We cannot allow our varied systems for producing knowledge to atrophy, and we must guard against delinking knowledge creation form human beings. While machines may someday understand our morals and ethics, this day is not yet here. Aligning machine intelligence with human values is, as many scientists and philosophers have asserted, an urgent undertaking.

Implications for the future of education

Developments in generative AI raise fundamental questions for the future of education.

What will be the role of teachers with this technology in wide circulation? What will assessment look like now that AI utilities can perform very well on examinations that were, until very recently, widely considered un-hackable, such as tests to demonstrate mastery of specific subject areas, and exams to credential skilled professionals, including doctors, engineers, and lawyers?

As a university professor, I have long considered the teaching of writing to be one of the most effective ways to cultivate and demonstrate analytical and critical thinking skills. But generative AI invites me to question such assumptions, even as I continue to hold them.

In a world where generative AI systems seem to be developing new capabilities by the month, what skills, outlooks and competencies should our education systems cultivate? What changes are needed in schools and beyond to help students navigate a future where human and machine intelligence seem to be ever more closely connected – one supporting the other and vice versa?

It is possible that we will soon achieve artificial general intelligence – a milestone at which machines will surpass us not only in narrow areas such as playing chess, but also in much larger ones, such as recommending actions to mitigate the dangers of climate change. What then should education look like? What will be its purpose and role in a world where humans are not necessarily the ones opening new frontiers of understanding and knowledge?

These are daunting questions. They are forcing us to seriously consider concerns that we have, arguably, avoided for too long.

At their most basic level, these concerns relate to the sort of world we want to live in. Our education systems often take for granted what the world looks like – and will and should look like. Our formal learning systems are designed to help people develop the competencies needed to navigate and, we hope, thrive in this known world.

AI is forcing us to ask questions about the 'known-world' that we usually take as a starting point for education. Many of our old assumptions and norms, especially those concerning knowledge and learning, appear unlikely to sustain the 'weight' of this new technology. We can no longer just ask 'How do we prepare for an AI world?' We must go deeper: 'What should a world with AI look like? What roles should this powerful technology play? On whose terms? Who decides?'

Education systems need to return agency to learners and remind young people that we remain at the helm of technology. There is no predetermined course.

Slowing and regulating the use of AI in education

Since the start of this year, we have come to recognize with clarity what scientists have been saying for at least a decade: The pace of AI developments is only accelerating. Today, we are moving at a breathless pace – and largely without a roadmap. Moments to pause, reflect and ask questions can seem rare, but we must consider where we are going and if this is indeed what we want.

The speed at which generative AI technologies are being integrated into education systems in the absence of checks, rules or regulations, is astonishing. I am struck that today, in most national contexts, the time, steps and authorizations needed to validate a new textbook far surpass those required to move generative AI utilities into schools and classrooms. In fact, AI

utilities often required no validation at all. They have been 'dropped' into the public sphere without discussion or review. I can think of few other technologies that are rolled out to children and young people around the world just weeks after their development. In many cases, governments and schools are embracing a radically unfamiliar technology that even leading technologists do not claim to understand. There are very few precedents for this development. The internet and mobile phones were not immediately welcomed into schools and for use with children upon their invention. We discovered productive ways to integrate them, but it was not an overnight process.

Education, given its function to protect as well as facilitate development and learning, has a special obligation to be finely attuned to the risks of AI – both the known risks and those only just coming into view. But too often we are ignoring the risks.

Schools, and to a lesser extent universities, need to be places where we are sure about what tools we are using with young people and recommending to them. Although it is still early, we know that one of the primary and most readily apparent risks of AI is its potential to manipulate human users. We further know that children and youth are highly susceptible to manipulation, much more susceptible than adults. There are numerous examples of AI slipping out of guardrails put in place by their creators and engaging in all sorts of 'conversations' that are inappropriate for children and likely to adversely influence them. This is especially the case as these tools become more calibrated for influence, entertainment, and prolonged engagement, as is currently the case with social media.

We have numerous precedents for slowing, pausing, or ceasing the use of technologies we do not yet understand, while continuing to research them. The research is vital because it adds to our understanding of the technology and informs us when and how it might be safe to use and for what purposes.

The use of AI can be harnessed or limited as is the case for other technologies, even though it has become popular to suggest this is somehow not feasible. We have robust rules in many countries that control and restrict use of technology that is known to be dangerous or is still too new to justify a wide or uncontrolled release. While these rules may not always be perfect, they are quite effective.

As we take fuller stock of the proliferation of generative AI applications, we must keep safety issues at the front of our gaze. It will likely take time to develop the necessary checks. The regulatory bodies that review and validate textbooks and other educational materials took significant time and investment to establish and sustain. These processes, already in place in most contexts, provide early, if rudimentary, blueprints for systems and processes to check large language model AI technologies for compatibility with educational aims. Educational resources bound for use in schools and with schoolchildren are typically vetted, at a minimum, on four main criteria: (1) accuracy of content, (2) age appropriateness, (3) relevance of pedagogical methods, and (4) cultural and social suitability which encompasses checks to protect against bias. In many places, resources are further inspected by groups of teachers and school leaders as well as various civil society groups, prior to receiving institutional approval. AI models and applications that claim to have educational utility should be examined according to

similar criteria, and others, given their complexity and reach, before being deployed at scale. It is rather remarkable that they have largely bypassed scrutiny of this sort to date.

The education sector needs to make these 'qualifying' determinations on its own terms. It cannot rely on the corporate creators of AI to do this work. Such industry self-regulation would introduce an unacceptable conflict of interest. To vet and validate new and complex AI applications for formal use in school will require ministries of education to build their capacities, likely in coordination with other regulatory branches of government, in particular those regulating technologies. Going forward, we need a much better balance of AI experts developing technology and applications for use, and, on the other side, experts working for governments to review the safety of these applications and to carefully consider their potentials for misuse and how to minimize these potentials. Presently, there are very few experts on the safety side of this equation and even fewer that are operating with real independence and outside the organizations developing AI for commercial purposes. The recent 2023 AI Intelligence Index Report showed that fewer than one percent of AI PhD graduates go into government following graduation. This trend has remained unchanged for the past five years. The majority of these graduates go into industry, while about one quarter go into academia. There is simply not enough expertise on the regulatory side of the equation.

In our present context of uncertainty, novelty and weak safety checks, a more cautious approach to generative AI in education is a commonsense course of action.

A roadmap to chart the way forward

UNESCO is working with countries to help them develop strategies, plans, and regulations to assure the safe and beneficial use of AI in education. In May 2023, UNESCO organized the first global meeting of Ministers of Education to share knowledge about the impact of generative AI tools on teaching and learning. This meeting has helped UNESCO chart a roadmap to steer the global policy dialogue with governments, as well as academia, civil society and private sector partners.

We are not starting from scratch. The 2021 UNESCO Recommendation on the Ethics of AI is an essential reference, as is the 2019 Beijing Consensus on AI and Education and our 2021 AI and Education Guidance for Policy-Makers. Our 2019 publication, *I'd Blush if I Could*, looked at the gender aspects of AI chatbots, and we have been pleased that OpenAI and other companies appear to have followed our recommendation to avoid gendering chatbots as young subservient women.

UNESCO encourages countries to prioritize the principles of inclusion, equity, quality and, most vitally, safety when moving to utilize AI tools for education. This is line with the commitments countries have made as part of the Sustainable Development Agenda and, more recently, the 2022 Transforming Education Summit, the largest gathering of the international education community in a decade.

Assuming AI safety can be more fully understood and assured, we must be open and optimistic about the ways it can support, supplement, and enrich the vital learning that happens as part of interactions in the physical and social sites of formal education. Education is – and should

remain – a deeply human act rooted in social interaction. It is worth recalling that when digital technology became the primary medium and interface for education during COVID-19 school closures, education was severely diminished, even if this exceptional period allowed us to clarify some of the ways technology can be better employed for teaching and learning and make education more flexible.

New and emerging challenges of digital technologies in education

Digital technology has exhibited a disturbing track record of widening divides within and between countries in education and beyond. AI technology will most probably accelerate the automation of large numbers of jobs. It also appears likely to dramatically improve the productivity of select workers, especially those already in high-paying fields and professions. We need to resist AI further widening inequity that is already too wide in many societies. New technology implementations should prioritize the closing of equity gaps, not as an afterthought but a starting point. In the case of generative AI, we need to ask: Will its deployment, according to a specific plan and timeline, likely widen or narrow existing educational divides? If the answer is no, the plan and timeline should be revised. We should be resolute in our expectation that this new class of technology open opportunities for all and reassert our commitment to equitable education.

We should further be watchful against the potentials of newly powerful generative AI technology, alongside older digital tools and services, to undermine the authority and status of teachers, even as it demands more from them. We would be naïve to think that future AI utilities will not strengthen calls for further automation of education: teacher-less schools, school-less education, and other dystopic visions. Developments like these are sometimes carried out in the name of efficiency and often impact the most disadvantaged learners first. Digital automation of education has long been proposed as a 'solution' and 'fix' for communities where education challenges and deficits are most severe. In the months and years ahead, some will argue for the use of generative AI to bring 'high quality' education to places where schools are not functioning, and teachers are in short supply or paid so poorly they do not regularly show up for work. Frontier technology is not the solution in these challenging contexts, even if it might be a piece of it. Well-run schools, enough teachers, and teachers with the requisite conditions, training and salaries that allow them to be successful remain the main ingredients of a sustainable remedy.

Our emerging AI world has also surfaced a dilemma in terms of investment choices. To what extent should we direct investments, including public investments, towards building the capabilities of machines that act like intelligent humans, *or* towards building the capabilities of living people? In the recent past, we could be certain that terms such as 'learning', 'educating', 'training', 'coaching', 'teaching' concerned human beings. This is now less clear. The business of 'educating' and 'training' machines is big, global and growing. It is also increasingly an area of competition, between private companies and actors, as well as nation states. Billions of dollars are now being invested in generative AI companies, when they could be directed towards teacher development and making needed improvements to schools and other physical and social infrastructure that benefit children. It is conceivable that the investments directed to making AI smarter and more capable might someday surpass the investments directed towards educating children and other people. While it is easy to get excited about *machines*

that can read and write; *people* that can read and write remain far more important. Today, at the dawn of our AI era, more than 700 million people are non-literate. We know that good schools and teachers can resolve this persistent educational challenge – yet, we continue to underfund them. Even if AI starts to exceed humans in a wide range of intellectual abilities, educating people will remain important and developing literacy most of all.

Rethinking education to shape the future

In our environment of AI acceleration and uncertainty, we need education systems that help our societies construct ideas about what AI is and should be, what we want to do with it, and where we want to construct guardrails and draw red lines. Too often we only ask how a new technology will change education. A more interesting question is: How will education shape our reception and steer the integration of new technology – both technology that is here today and technology that remains on the horizon? Our education systems can define a trajectory and establish norms for how we understand world-changing technology – and, by extension, how we allow it to influence us and our world.

This is perhaps the 'raison d'être' of education: to help us make informed choices of how we want to construct our lives and our societies. The central task for education at this inflection moment is less to incorporate new and largely untested AI applications to advance against the usual targets for formal learning. Rather, it is to help people develop a clearer understanding of when, by whom, and for what reasons this new technology should and should not be used. AI is also giving us impetus to re-examine what we do in education, how we do it, and, most fundamentally, why.

Now is the time to rise to these challenges. As AI experts remind us, our continued well-being and perhaps even survival may be at stake. Our work must be infused with urgency as we endeavor together to ensure that our education systems play a key role in getting humanity's transition into an AI world right.

Published in 2023 by the United Nations Educational, Scientific and Cultural Organization

7, place de Fontenoy, 75352 Paris 07, France

© UNESCO 2023

This paper can be cited with the following reference: Stefania Giannini, 2023, Reflections on generative AI and the future of education. © UNESCO 2023

This work is available under the Creative Commons Attribution-ShareAlike 3.0 IGO licence (CC BY-SA 3.0 IGO; https://creativecommons.org/licenses/by-sa/3.0/igo).

Credit for the illustration on the first page: © UNESCO/Rob Dobi

For further information, please contact: adg-ed@unesco.org

ED/ADG/2023/02