The critique of AI as a foundation for judicious use in higher education

Stefan Popenici
Academic Lead - Quality Initiatives, Charles Darwin University, Australia

Keywords
Artificial intelligence (AI); bias; discrimination; education; higher education; learning; surveillance; teaching.

Abstract
The integration of Artificial Intelligence (AI) into educational settings, especially after the launch of ChatGPT into the public space, created new challenges and massive disruption for schools and universities across the world. This paper aims to state and look beyond the hype on AI, marketing and myths that are obscuring some of the most significant challenges and analyse potential risks associated with the adoption of AI in education. It also aims to find practical ways of using AI for the benefit of students, teachers and institutions of education. The analysis is focused on the key ethical implications of AI, the impact on teachers, students, and the future of learning, as well as long-term societal implications.

Correspondence
stefanpopenici@gmail.com

Article Info
Received 25 June 2023
Received in revised form 4 July 2023
Accepted 4 July 2023
Available online 5 July 2023

DOI: https://doi.org/10.37074/jalt.2023.6.2.4
Ideological roots in defining human and artificial intelligence

The end of 2022, and especially the first months of 2023, will remain in history as the time when artificial intelligence (AI) invaded and colonised public imaginations across the world. The prominence of generative AI solutions such as ChatGPT, created and released for the public by OpenAI, made evident in an extraordinarily short period of time that the impact of AI on everyone’s life will be exponentially more significant than it already was (Sullivan et al., 2023; Rudolph et al., 2023b). This makes it even more important for education at all levels, and especially for higher education, to stop and seriously interrogate what artificial intelligence is and what stands behind this marketing formula that came into the world in 1956 at the Dartmouth Summer Research Project (McCarthy et al., 1955). Defining artificial intelligence is notoriously difficult, especially if we consider that AI is now attached to almost any technology that can be sold and take advantage of the popularity of the term. In a paper published in 2017, I defined AI “as computing systems that are able to engage in human-like processes such as learning, adapting, synthesising, self-correction and use of data for complex processing tasks” (Popenici & Kerr, 2017, p. 2). In other words, AI describes technological solutions that are able to complete tasks that are usually associated with human abilities. And here is the point where we have a source of confusion: while it is true that AI can complete tasks that are usually requiring human abilities, it does not also mean that AI is intelligent or able to think. The second problem related to a definition of AI is that here we have a narrow understanding of what intelligence means based on a narrow and ideological description of human intelligence. We can understand this better if we simply stop to think about what stands behind the words that create the marketing label that is artificial intelligence. Records show that John McCarthy was clearly aware of the fact that artificial intelligence is, most of all, a marketing concept, a magnet for research funding, as well as a political move to distance himself from cybernetics and Norbert Wiener. When he won in 1971, the A. M. Turing Award, McCarthy mentioned in his speech to the Association for Computing Machinery (ACM) that it is ironic that his most widely recognised contribution “turned out to be in the field of marketing, specifically in choosing a brand name for the field.” (Katz, 2020, p. 23).

The first part of the formula is the word “artificial”, which is an unusual choice. If we try to find a synonym for artificial, we realise that this is not a positive term as it essentially opposes what is natural. Synonyms for artificial have, in general, a negative connotation, with replacements such as “fake”, “fictitious”, “false”, “simulated”, “bogus”, “made-up”, “forged”, “not genuine”, etc. This should be the first clue that we have a certain view on what kind of intelligence we are considering here. It is not a natural or human intelligence.

The second part of the formula is much more complicated because the history of the term is rooted in some dark and toxic ideological positions. The contemporary understanding of what human intelligence is is shaped by a cousin of Charles Darwin, Sir Francis Galton, who was convinced that everything could be measured. How passionate he was about measurements becomes clear when we consider that he believed possible to build an “attraction gauge” (how much a person is infatuated with another). He explored how we can scientifically measure boredom, and he authored in 1872 the paper “Statistical inquiries into the efficacy of prayer”. More consequential is that Francis Galton was interested in human intelligence, and he translated his passion for measuring everything to this field, setting the foundations of a certain way of looking at human intelligence, which was later developed in the discipline of psychometrics. To understand better how Galton’s perspective on human intelligence influenced our current understanding of what human intelligence is, it is important to look at his other significant contribution to posterity: he also coined in 1883 the term eugenics, the reprehensible and toxic theory of rankings of human abilities based on racial and hereditary factors, which is commonly associated only with Nazi Germany and the Holocaust. This is just the effect of these ideas, but Galton founded the theory of “racial hygiene” and eugenics, which stands as the most influential pillar for the new studies and ideas of intelligence. The real impact of this reprehensible theory erupted a few decades later when Nazis adopted eugenics as an ideological foundation. These theoretical roots stand responsible for the unprecedented tragedy of the Holocaust. Historical evidence shows that Hitler was particularly interested in Galton’s ideas about intelligence, heredity and racial hygiene. Otto Wagener, the head of the Nazi Party’s Economic Policy Office from 1931 to 1933, specifically noted Hitler’s interest in Galton’s theory of eugenics and its applications in the United States. In his notes, Wagener quoted Hitler saying that “it is possible to a large extent to prevent unhealthy and severely handicapped beings from coming into the world. I have studied with great interest the laws of several American states concerning prevention of reproduction by people whose progeny would, in all probability, be of no value or be injurious to the racial stock” (Kühl, 1994, p. 37).

What is important to realise is the fact that these toxic ideas were not adopted only by an influential thinker that set some foundations for studies in this field at the end of the 19th century. The reality is that these theories remain influential across decades until the present. Most influential scholars in the field of intelligence pay tribute, most often with explicit enthusiasm, to a hereditary and discriminatory perspective on what human intelligence is and how it can be identified and measured. The most influential figures that shaped our current understanding of intelligence, from Galton and Karl Pearson, William James and John Dewey, to Lewis M. Terman, the famous professor of psychology at Stanford University with immense studies on human intelligence, stand close to a eugenic, hierarchical view of intelligence that is presented as determined by social status and hereditary mental abilities. All these thinkers and theories stand close to the disastrous and wrong assumptions set by Galton in his attempt to scientifically justify racism, social injustice and discrimination. In fact, Lewis Terman, who is undoubtedly one of the most influential scholars in the modern studies of intelligence, was an active member of eugenic societies in the United States and actively advocated for the forced sterilisation of those labelled as inferior in American society (Leslie, 2000). There is an obvious impact of these reprehensible ideas inspired by eugenics, with events that will forever mark the history of humanity with their power to fuel extreme violence.
perversion of thinking, and abysmal inhumanity. A more discrete and insidious effect of these theories is that they restricted collective imaginations and scientific research to a narrow view of what human intelligence is, how it manifests, and how and if it can be measured or nurtured.

It is important to note that the development of new technologies became the most fertile ground for eugenics and racial theories of intelligence. William Shockley is considered the “father of Silicon Valley”, a winner of the Nobel Prize in Physics in 1956, and the one who “brought silicon to Silicon Valley” (Isaak et al., 2016, p. 167). William Shockley is also a known racist, white supremacist, and strident supporter of eugenics. He was invited in 1965 to deliver an address at the first annual Nobel Conference, a conference on Genetics and the Future of Man that was held in the United States but was authorised by the Nobel Foundation. At that conference, Shockley revealed his racist ideology. He claimed that social policies were allowing genetic defectives to proliferate... To Shockley’s mind, only by systematic reduction of the African American population by sterilization and other methods of birth control could we improve our society. This would lead to survival of the fittest, and the fittest were the original European settlers into America. Racial discrimination was not prejudice, he claimed, but was justified based on statistics (Sussman, 2014, p. 236).

The adherence to this toxic view of the world and discrimination based on old misjudgements and prejudices is also closely associated with the emergence of artificial intelligence as a formula and theory. John McCarthy, who coined the term “artificial intelligence” in 1955 (McCarthy et al., 1955), openly expressed his beliefs on the hierarchical structure of intelligence based on gender, underlining in an essay titled “Technology and the position of women” his beliefs. He states there that it would be a mistake to assume that women are not inferior to men, writing that:

The very highest level of potential in science and mathematics, which only one in a million men can attain, the fraction of women who can attain it may be biologically smaller... At present there are social movements and people with institutional power who regard there being fewer women than men at some level of some occupation as an injustice that must be corrected by quotas. This is a mistake and will not succeed because of differences in ability and motivations between males and females (McCarthy, 2006).

In June 2023, France24 published a documentary based on investigative journalism on Clearview AI, a company specialised in facial recognition that is scraping astonishing amounts of data on every person who had a picture taken and uploaded, with or without consent, on the Internet. In this documentary, Jessica Le Masurier and Romeo Langlois (2023) unveil not only the staggering breach of privacy and serious implications for citizens across the world but a surprisingly obvious link between this powerful company and white supremacists, fascists, and anti-democratic forces in the US.

We must consider in education that AI is placed at the convergence of two dangerous temptations, which both shape its development, influence, and dynamics. The first is the temptation of technology and its efficiencies to lead towards a certain view of the world, which is easily adopted by totalitarian, amoral tendencies. This connection was summarised by Herbert Marcuse in his analysis of Nazi dictatorship as “a striking example of the ways in which a highly rationalised and mechanised economy with the utmost efficiency in production can also operate in the interest of totalitarian oppression” (Marcuse & Kellner, 1998, p. 416). From a purely technological perspective, Nazi Germany was the most advanced at that time, creating the first man-made object in space (the infamous V2 rocket), making technological advancements that opened the space exploration program a few years after their defeat. Technological excellence was not making that regime less evil but worse and more destructive. This historical fact invites a serious consideration of the necessity to associate technological progress with ethical considerations and to maintain a critical perspective on technological development.

The second temptation for artificial intelligence is much more straightforward and evident. There is a documented tendency of AI to immensely enhance surveillance and inequality, bias, and discrimination and widen power imbalances. A disconnect from ethical considerations is dangerous for civil societies, democratic processes and educational aims for higher education.

Two dangerous myths about AI in education

One common misconception affecting the perception of AI systems and how they impact education is that data is an objective construct, atemporal and value-neutral, shaped only by exact and cold evidence and accurate representations of reality. In fact, the perception is that technology itself is an objective medium. Hence, AI is a technological solution that is operating based on factual, unbiased and clinical processes. If we think about how technology actually operates, we realise that there is not one point in the history of humanity when technology is no directly related to specific cultures and values, beliefs and biases, religious beliefs or gender stances. A study on gender bias in technology starts from these basic facts, noting that

first, and foremost, (there) is the notion that technology often shapes culture and its meanings. The second is that we have become so used to technology in our daily lives that we fail to see its implications. We argue that our familiarity with technology means that traditional methods of analysis will be unable to unveil the ideology that perpetuates gender bias as a mode of domination. Thus, a critical analysis of technology and society is required for technology to reach its emancipatory potential (Kilbourne & Weeks, 1997, p. 244).
Indeed, any informed and responsible use of technology, especially revolutionary solutions such as generative AI, requires a critical analysis that cannot start from the naive perception that we can have in this space, a vacuum of values and specific choices. Any technological solution and adoption involve a certain ideological choice and influence, consciously or not. If we accept the obvious fact that values and particular perspectives influence the development and applications of technologies, we have to consider a series of troubling facts, such as the vast disproportionate influence of men on the development and programming of AI. Currently, only a small percentage of women work now in AI: “The percentage of women working in AI today is approximately 30%” (WEF, 2023, p. 7). In Silicon Valley, the percentage is even smaller, just above 10%. To use just one example, we can consider that one research project conducted at the University of Cambridge has found that using AI to complete literature searches provides results with a bias favouring white, Western and male authors (Jordan & Tsai, 2022). This means that other voices and perspectives in research and the advancement of knowledge become de facto invisible. Implications for our common future are significant.

The second common belief is somehow linked to the first position that leads to errors in the use and adoption of new technologies in education and revolves around the idea that algorithms do not discriminate, as they are “just maths”. In her book, “Weapons of math destruction: how big data increases inequality and threatens democracy”, Cathy O’Neil provides a convincing and well-justified counterargument to this erroneous position, noting that mathematics cannot offer alone protection against bias, misuse, and manipulations. The book documents that:

The math-powered applications powering the data economy were based on choices made by fallible human beings. Some of these choices were no doubt made with the best intentions. Nevertheless, many of these models encoded human prejudice, misunderstanding, and bias into the software systems that increasingly managed our lives. Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists. Their verdicts, even when wrong or harmful, were beyond dispute or appeal. And they tended to punish the poor and the oppressed in our society, while making the rich richer (O’Neil, 2016, p. 10).

There is consistent research and books that are providing examples of AI algorithms that discriminate, grotesquely amplify injustice and inequality, targeting and victimising the most vulnerable and exposing us all to unseen mechanisms of decision where we have no transparency and possibility of recourse. It is worth mentioning here the book “Automating inequality: How high-tech tools profile, police, and punish the poor”, by Virginia Eubanks (2018). In “Algorithms of oppression. How search engines reinforce racism”, Safiya Umoja Noble starts an excellent expert analysis by reminding us that “Part of the challenge of understanding algorithmic oppression is to understand that mathematical formulations to drive automated decisions are made by human beings”, and documents that “algorithmic oppression is not just a glitch in the system but, rather, is fundamental to the operating system of the web”. (Noble, 2018). This point is extremely important for any informed user of AI, especially in education, and the fact that discrimination and racial biases are part of the internal design rather than a simple “glitch”. In the book published in 2023, titled “More than a glitch: Confronting race, gender, and ability bias in tech”, Meredith Broussard (2023) substantiates the point that value neutrality in tech is a myth and, as it expertly proves that bias and discrimination are not a simple error but a matter of design, brings new arguments to focus our efforts on holding algorithms transparent and accountable.

It is impossible to make here a comprehensive selection of some of the most relevant research, books, studies or even journals reflecting the fact that algorithmic decision-making is inherently dangerous and toxic without constant and alert human supervision and interrogation. What stands relevant is that the blind trust in and adoption of new tech by educators, which was ubiquitous for the last decades in schools and universities across the world, becomes even more dangerous in the era of AI. The challenge ahead for education is to become users of AI for the benefit of our students and institutions rather than simple subjects of AI, providers of data and digital serfs controlled by an almighty Big Tech.

**AI and the aims of education**

A research paper submitted for preprint by a group of researchers from Princeton University, the University of Pennsylvania and New York University identifies professions that are most likely to be lost or degraded by the impact of AI. Researchers have found here that the vast majority of those most exposed the AI disruption are teachers in schools and higher education (Felten et al., 2023). This confirms what became obvious in the first months of 2023 after ChatGPT and other large language models and generative AI solutions became popular with the general public. The impact on education, students, teachers, schools and universities was not close to the main concerns of developers of AI or tech startups. Moreover, there are sufficient reasons to claim that the aims of educators stand very far from their motivations. For example, in a 2019 interview published by Forbes, Sam Altman, the CEO of OpenAI, the company that developed ChatGPT, makes the significant observation that “AI will probably most likely lead to the end of the world, but in the meantime, there'll be great companies” (Martin, 2019). Considering that OpenAI managed to secure immense funding and profits in a very short time, we can safely assume that “great companies” describes profitable companies here, and here is the key: Big Tech is driven by the aims of profits and power, control and financial gain. Institutions of education and teachers have very different aims: the advancement of knowledge and to nurture educated, responsible, and active citizens that are able to live a balanced life and bring a positive contribution to their societies. As noted in the book “Artificial Intelligence and Learning Futures”, institutions of higher education
were created to serve the common good and, with their concentration of academics working together for research and education, can advance knowledge and serve society with wise and innovative solutions for our critical challenges. Higher education is the space where new ideas can organically emerge, when the campus ethos is defined by intellectual effervescence, and moral engagement for a civil and advanced society. The general aim of universities is to disseminate knowledge and nurture more informed, ethical and educated citizens, able to bring a positive contribution for a civil society (Popenici, 2022, p. 3).

Big Tech and new tech startups are established to secure profits and control, at least on the market. It is significant that OpenAI is an example of a startup that was established with the aim to “serve humanity”, “unconstrained by a need to generate financial return”, and just a few years later became entirely opaque in the design and use of their algorithms while securing billions of dollars in new funding. There is not even a point of convergence in the aims of those who currently control and build AI and the users of AI in education, a field that is undoubtedly going to suffer major changes in the near future as a result of its rapid development and adoption in teaching, learning and assessment. This is an important reason for universities and educators to stay alert and interrogate intentions and applications of AI, as well as keeping strict control on the AI inherent tendencies to discriminate and amplify biases.

The use of AI is also directly linked to a set of risks related to users’ privacy. The popularity of generative AI in 2023 obscured that all details, prompts and use of AI involve two clear dynamics. First, all information and data the user provides are used to train and develop the AI models. While it is absolutely acceptable for a teacher to contribute freely to the development of models that increase the profits of private companies, some students may not be aware that their work and ideas are used by a third party in ways that are not always clear. The second dynamic is much more significant: all data provided to an AI solution is potentially filed, used, aggregated, and connected to a user’s identity. Especially at a time when banks use data aggregated from the Internet to decide a credit score, insurance companies decide premiums based on information sold by data brokers and all our lives are influenced by data collected on individuals with and without their knowledge, teachers have a duty of care to protect the privacy of students and their future.

As we briefly detailed in previous paragraphs, algorithmic discrimination is a tangible reality which significant effects on large groups of people and is especially damaging for the marginalised and most vulnerable in society. This is directly associated with AI and its functions and is raising specific challenges in education as it inevitably requires data and information from the users. Any breaches of privacy and disclosure of sensitive data will have long-term impacts, which are impossible to evaluate due to the opaque nature of AI models. Even the most ardent and interested supporters of generative AI raise the alarm about the potential risks for privacy and data confidentiality: in June 2023, Google (Alphabet) warned its own staff to avoid sharing personal or professional information on AI chatbots, including on its own AI solution, Bard (Dastin & Tong, 2023).

There is also the real risk that learning is further pushed to the margins of the process in the current hype surrounding the potential of AI to improve education, assessment and teaching. This is summarised in an analysis of the adoption and use of ChatGPT in higher education, at a moment when it is tempting to use artificial intelligence to assess the originality of assignments: “A first AI circumvents a second AI and is assessed by a third AI. All that the humans do is press a couple of keys, and nobody learns anything” (Rudolph et al., 2023a, p. 354). AI presents obvious advantages in automating assessments, further personalising teaching, providing individualised assistance or replacing university administration, but it can also further alienate and dehumanise learning, to the point of technological potemkinisms as those described by Rudolph et al. (2023a). There is no evidence that universities across the world use the ChatGPT moment to radically change their governance and ideological models, and structurally change teaching and assessments to nurture key skills for the current challenges, students’ creative and critical thinking abilities (Rudolph & Tan, 2022), wisdom and social responsibility.

The long-term fixation of education on personalisation also requires the collection and aggregation of student data, and AI brings now new challenges to a project that was brilliantly analysed by Audrey Watters (2023) in her book, “Teaching machines: The history of personalized learning”. The rapid adoption of AI solutions by educators, administrators and students opens new possibilities for hyper-personalisation and data collection: “AI is bringing the promise of ‘super-charging’ personalisation of education, using data and complex algorithms to predict what is the most suitable content, teaching method, educational intervention, and pace of instruction for every student” (Popenici, 2022, p. 107). We cannot properly evaluate how super-personalisation and vast data collection will impact the future of our students, on their credit ratings, mortgages, medical services and so on, but we have the duty of care to protect them from exploitative and potentially damaging practices. The task of helping our students become informed and able users of various AI platforms becomes more important now.

**Conclusion**

The impact of generative AI is most visible in the area of learning, teaching, and especially academic integrity. The rapid adoption and unprecedented number of users in a very short time came as a “shock in education, like the COVID-19 pandemic” (Mills et al., 2023, p. 16). Noam Chomsky, who is most probably the most reputable professor of linguistics and cognitive science, can offer some guidance on the potential impact of ChatGPT, the most popular AI program for generative language. Chomsky succinctly defined the role of ChatGPT in education as “High-Tech Plagiarism” and “a way of avoiding learning” by students, noting in an essay written with Ian Roberts, a professor of linguistics at the University of Cambridge and Jeffrey Watumull, a philosopher and the director of artificial intelligence, that
ChatGPT exhibits something like the banality of evil: plagiarism and apathy and obviations... ChatGPT and its brethren are constitutionally unable to balance creativity with constraint. They either overgenerate (producing both truths and falsehoods, endorsing ethical and unethical decisions alike) or undergenerate (exhibiting noncommitment to any decisions and indifference to consequences). Given the amorality, faux science and linguistic incompetence of these systems, we can only laugh or cry at their popularity (Chomsky et al., 2023).

It is deceiving to say, dangerous to believe, that artificial intelligence is... intelligent. There is no creativity, no critical thinking, no depth or wisdom in what generative AI gives users after a prompt: it is just plausible text with good syntax and grammar, and this is all that it is. Intelligence, as a human trait, is a term that describes a very different set of skills and abilities, much more complex and harder to separate, label, measure and manipulate than any computing system associated with the marketing label of AI. AI is already tentatively used to replace teachers in higher education, and publications such as The Independent in the UK are spreading the hype with titles such as “Harvard’s new computer science teacher is a chatbot” (Cuthbertson, 2023). Harvard University presented this as “an evolution to tradition”, that “can support their learning at a pace and in a style that works best for them individually” (Hamid & Schisgall, 2023). Universities will further integrate AI in their courses and will use AI bots to replace teachers, not because it will help students develop skills that will be relevant and help them in the era of AI, such as independent and critical thinking, superior abilities to master knowledge with genuine creativity and meaningful contributions. The AI replacement of teachers is a process well aligned with the ideological models adopted by universities for the last decades: marketisation, maximisation of profits that can be secured by culling teaching employees and viewing learning as an assembly line process where information is delivered as a product and commodity, serving models narrowly suited for employment and the job market. This is a risk that is not yet on the agenda of politicians or decision-makers, and it is not part of the agenda of university administrators. The risk will not disappear and most probably will accelerate the current crisis of learning and teaching, and the crisis impacting enrolments and the public trust in higher education.

It is evident at this point that AI is an integral part of education, as it has been for a long time – more discreetly – in many other areas of our lives. Banning or ignoring generative AI in education is an unrealistic, ignorant, and dangerous option, which was unfortunately adopted by many educators, schools and universities when it became clear how widely used ChatGPT is. It is vital for educators to understand what AI is and what it is not, what is just hype and marketing, and make the difference between the real potential for beneficial use or selling points and propaganda. It is also important to identify real expertise or just a desire to join the hype or a simple lack of knowledge. This is a hard task, as the vast resources allocated to promote selling points of companies with vested interests in this field are building passionate defences of AI, usually associated with religious fervour. However, if universities and educators want to remain relevant in the future and have a real chance to reach the aims of education, it is important to consider the ethical and intellectual implications of AI, some not even mentioned in the current paper. This will be a field open for new and extraordinarily significant future research.

References


