Artificial Intelligence: Ethical Considerations and Caveats

Prof. Aly Abdul Samea Qoura
Mansoura Faculty of Education

المجلة الدولية للعلوم التربوية والتكنولوجية والتنمية

المجلد الثاني– العدد الثالث– مسلسل العدد (٤) – يوليو ٢٠٢٤م

ISSN-Print: 3009-7851 ISSN-Online: 3009-7444

موقع المجلة عبر بنك المعرفة المصري https://ijsetd.journals.ekb.eg

البريد الإلكتروني للمجلة E-mail IJESTD@foe.zu.edu.eg

Artificial Intelligence: Ethical Considerations and Caveats

Prof. Aly Abdul Samea Qoura
Mansoura Faculty of Education

Abstract

AI is a powerful force transforming lives, interactions, environments and societies. We use AI every day, mostly without noticing. Self-driving vehicles, digital assistants, chatbots, face recognition and drones are just a few examples of AI applications. This paper reflects on the ethical aspects of using artificial intelligence (AI) systems in educational contexts. On the one hand, the impact of AI in the field of education is promising where recent research suggests that AI will enhance the teaching–learning process, as well as the engine of what is already being called Education 4.0. However, there is an emphasis on the ethical dangers of AI in the educational context concerning transparency, privacy, data protection, human rights, and data ownership, to mention a few. This paper focuses on four main issues: 1. Professional ethics, 2. Benits of AI in education, 3. Ethical considerations and caveats of AI and finally, Measures and/or recommendations to mitigate AI ethical caveats.
Keywords: Artificial intelligence, ethics, morals, ethical, considerations, caveats

Introduction

A new era of education and research based on chatbots and artificial intelligence (AI) is quickly growing. However, the application of these new systems is associated with several challenges and limitations, mainly related to ethics. This paper explores the potential use of AI systems and chatbots in the academic field and their impact on research and education from an ethical perspective. The paper highlights the necessity of adaptation to the new reality of AI systems and chatbots. Raising awareness, adopting appropriate legislations and enforcing ethical values will strengthen research and protect educational systems. The presence of AI systems and chatbots in education needs to be considered as an opportunity for development taking into consideration the ethical values that are needed to govern the AI apps and tools.

Artificial Intelligence (AI) is a broad term used to describe a collection of technologies able to solve problems and perform tasks without explicit human guidance. Some of these include: machine learning, computer vision, natural language processing, robotics and deep learning. A general-purpose technology, AI uses data-driven algorithms to autonomously solve problems and perform tasks without human guidance. The algorithms that underpin artificial intelligence have existed

for quite some time, however exponentially growing volumes of data and the widespread availability of affordable computation mean that the world, can operate this revolutionary technology at a scale and speed never seen before (CSIRO, n.d.).

Artificial intelligence has been defined as: a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy (OECD, 2019). AI has also been considered as machine-based systems, operating on a large scale with varying levels of autonomy that can "make predictions, recommendations, or decisions influencing real or virtual environments." AI is relevant in any context where large volumes of data and information are processed. Today, it is a strategic technology that offers many potential benefits for citizens, the public good and the economy at large, provided it is human-centric, ethical, and respects fundamental rights and values.

AI systems are designed to operate with varying levels of autonomy. AI system lifecycle phases consist of: 1) planning and design, data collection and processing, and model building and interpretation; 2) verification and validation; 3) deployment; and 4) operation and monitoring (OECD, 2019, pp. 23–24). One of the most promising AI techniques is machine learning (ML), which is described as a set of techniques to allow machines to learn in an automated manner through patterns and inferences rather than through explicit instructions from a human. Behind ML is a technique referred to as 'neural networks', which is accompanied by growing computational power and the availability of massive datasets, also known as big data (OECD, 2019, pp. 27–28). In education for example, language learning applications rely on ML.

AI is a powerful force transforming lives, interactions, environments and societies. We use AI every day, mostly without noticing. Self-driving vehicles, digital assistants, chatbots, face recognition and personalized recommendations are just a few examples of AI applications. AI is here to stay and will continue to transform society substantially. The most pressing questions are how this transformation will take place and what the repercussions will be. The introduction of increasingly sophisticated AI enabled systems in many areas of everyday life comes with promises of improvements in social, physical and environmental well-being. Nevertheless, growing concerns about ethical challenges accompany the development of these cutting-edge tools, such as bias, inequity and loss of autonomy.

These technologies often seem to pose novel challenges to professional ethics and values, though whether they change issues

fundamentally is more debatable (Ferguson, et al., 2016). One such technology that seems to raise a host of new ethics issues is artificial intelligence (AI) as a whole and ChatGPT in particular. In the last five years there has been much excitement around AI in many realms of activity. But some notable failures have also triggered considerable concern around the ethics of data and AI. In response there has been a proliferation of ethics principles published, though critics still find that these are not enough to ensure ethical practice, suggesting the need for more reflection on the underlying issues. Most domains of work and life seem to be being impacted by AI, including many applications relevant to information professionals. So, it is important for information professionals to think how their existing ethics and values apply to the new technology.

This paper focuses on the ethical aspects of using artificial intelligence systems in educational contexts. On the one hand, the impact of AI in the field of education is addressed from the perspective of its use by teachers and students. On the other hand, there is an emphasis on the uncertainties caused by the fears of some who think, for example, that AI robots will replace human teachers. Finally, some critical questions about the ethics and challenges of AI in education are proposed, concluding with the need to add proposals with new research and political actions which could lead to the creation of an ethical observatory of AI for education.

In this context, recent research suggests that AI will be a great novelty in education and the teaching–learning process, as well as the engine of what is already being called Education 4.0 (Fidalgo-Blanco et al., 2022; Ramírez-Montoya et al., 2022). Some argue that the role of AI will range from initial education (infants) to higher education, including business training. It will enhance the student experience by providing the opportunity to create adaptive learning functions and models with personalized tools (Flogie & Aberšek, 2021). It should be noted that both teachers and students are increasingly immersed in the use of technologies and platforms that optimize, on the one hand, the transmission of knowledge and, on the other hand, the acquisition of new learning. In summary, AI will have different educational roles (Affde, 2021; & UNESCO, 2021).

The various uses of AI, as well as the ethical challenges related to them, require multidisciplinary and multi-stakeholder engagement, and international collaboration on developing AI governance. As the Technical University of Munch (TUM)'s strategic partnership initiative articulates: "No single university or country can master today's scientific challenges on its own." The Institute for Ethics in Artificial Intelligence

(IEAI) holds public events and supports and participates in several networks related to AI ethics. By joining forces with researchers and practitioners worldwide, the IEAI aims to address real-world challenges and contribute to the broader conversations and concerns surrounding ethics and AI on an international level. Among others, priorities include protecting individual interest and privacy, ensuring social justice and digital equity, reducing disparities, protecting the environment as well as defining accountability, responsibility and transparency.

Consideration of these challenges will help to ensure the development of advanced and "trustworthy" AI systems. Achieving the dual advantage of "ethical AI", where technological innovations capitalize upon opportunities to improve well-being, sustainability and justice and, at the same time, foresee, minimize and mitigate associated risks , requires the consideration of several aspects of the development of these systems: from the design, to the context of deployment, to the users and needed data.

Professional Ethics

The second section of this paper addresses professional ethics since the need arises for an AI ethical framework that governs the data system, design, deployment, and use of applications and tools related to AI.

Qoura (2017, p. 2–3) stated that Professional ethics encompass the personal, and corporate standards of behavior expected by professionals. Professionals and those working in acknowledged professions exercise specialist knowledge and skill. How the use of this knowledge should be governed when providing a service to the public can be considered a moral issue and is termed professional ethics.

Ethical conduct is more than merely abiding by the letter of explicit prohibition. Rather, it requires unswerving commitment to honorable behavior, even at the sacrifice of personal advantage. The conduct toward which educators should strive is embodied in six broad concepts: Integrity, Objectivity, Competence, Privacy, confidentiality, Fairness and Justice.

Professional ethics helps us to define our roles and responsibilities —both to ourselves and to our various constituencies. In our work, professional ethics means:

- upholding both the letter and the spirit of the principles, rules and guidelines applicable to all profession members;
- fostering accountability;
- eliminating unfair and disrespectful treatment of others;
- asking questions when we are confronted with ethical issues;
- encouraging open dialogue and discussion;
- candidly acknowledging, and learning from, our mistakes;

- feeling proud of what we have achieved and how we have achieved it; and
- being humble in considering what we can improve and how we can do it.

Morals versus Ethics

Chowdhury (2016) differentiated between morality and ethics. He believes that morality and ethics are part of a way of life and cannot be separated from all other aspects of life experiences. Moral education aims at promoting students' moral development and character formation. The theoretical framework of moral education is supported by moral philosophy, moral psychology and moral educational practices (Han, 2014). Beyond the scope of promoting rational pro-social skills or virtues, moral education of real human value should cultivate the meaningful and personally formative knowledge that significantly transcend or avoid natural and/or social scientific understanding and explanation (Carr, 2014).

Ethics sometimes override personal morals. That is the case when you believe that according to your morals you should do something but ethically you cannot. For example, consider a criminal defense lawyer. Though the lawyer's personal moral code likely finds murder immoral and reprehensible, ethics demand the accused client be defended as vigorously as possible, even when the lawyer knows the party is guilty and that a freed defendant would potentially lead to more crime.

Legal ethics must override personal morals for the greater good of upholding a justice system in which the accused are given a fair trial and the prosecution must prove guilt beyond a reasonable doubt.

Difference between Moral and Ethical

Morality is defined as conformance to a recognized code, doctrine, or religion, or system of rules of what is right or wrong and to behave accordingly. No system of morality is accepted as universal. What is moral and what is not moral differs sharply from place to place, group to group, and time to time.

Morals or moral values are generally associated with personal view of values. Which reflect beliefs relating to sex, drinking, gambling, etc. They can reflect the influence of religion, culture, family and friends. Ethics is concerned with how a moral person should behave. Ethical values are beliefs concerning what is morally right and proper as opposed to what is simply correct or effective.

In teaching ethics teachers and educators strive to achieve the highest ideals of their professional service. It refers to teacher's and educators basic professional responsibilities (individual, collective,

institutional) towards their profession, their academic discipline, their educational institution and to society as a whole.

Moral education is about an inner change, which is a spiritual matter and comes through the internalization of universal religious values (Islamic or Christian for example). Ethics is the branch of philosophy which tries to probe the reasoning behind our moral life.

While they're closely related concepts, morals refer mainly to guiding principles, and ethics refer to specific rules and actions, or behaviors

Moral refers to an individual's personal stance, principles and values of what is right and wrong, whereas ethics refers to rules, principles, norms and values that are external to the individual, such as the norms and values of a social community, occupational codes of conduct, religious principles, etc.

One controversial example from history is the dropping of the atomic bombs in World War II. The bombs were dropped on civilian populations, which is immoral, but the ethical argument is that ending the war sooner saved more lives than were lost in the bombings.

Ethics, for example, refers to those standards that impose the reasonable obligations to refrain from rape, stealing, murder, assault, slander, and fraud. Ethical standards also include those that enjoin virtues of honesty, compassion, and loyalty.

Ethics examines the rational justification for our moral judgments; it studies what is morally right or wrong, just or unjust. In a broader sense, ethics reflects on human beings and their interaction with nature and with other humans, on freedom, on responsibility and on justice

Ethics is what guides us to tell the truth, keep our promises, or help someone in need. There is a framework of ethics underlying our lives on a daily basis, helping us make decisions that create positive impacts and steering us away from unjust outcomes. Nest section tackles ethics as it relates to AI applications and repercussions.

Ethical Considerations

AI widespread use and the dangers that it may trigger, require a strong and serious framework of ethics. With the adequate ethics of AI, we can design and build AI that behaves ethically (i.e., Ethical AI). The ethical issue of AI generally refers to the morally bad things or problematic outcomes relevant to AI (i.e., these issues and risks that are raised by the development, deployment, and use of AI) that need to be addressed. Many ethical issues, such as lack of transparency, privacy and accountability, bias and discrimination, safety and security problems, the potential for criminal and malicious use, and so on, have been identified from the applications and studies.

According to Huang et al (2022), Ethical Issues caused by AI such as: (1) Transparency or explainability is one of the most widely discussed downside of AI. It is hard to explain and understand the inference procedure of ML, which is commonly known as the "black-box". The opacity of ML makes the algorithms or models mysterious to users and even developers. The lack of transparency not only leads to the explanatory problem, but also leads to difficulties in human monitoring and guidance of ML or AI. (2) Data Security and Privacy. The performance of current AI strongly depends on a huge amount of training data, which probably includes personal data and private data. The misuse of data, such as (personal) information leakage or tampering, are serious ethical issues that are closely related to every individual, encountered in the development and application of AI technology. (3) Autonomy, Intentionality, and Responsibility. Here, the autonomy of AI refers to an AI system's ability to operate without human intervention or direct control. Intentionality refers to the ability that an AI system can act in a way that is morally harmful or beneficial and the actions are deliberate and calculated. Responsibility indicates that the AI system fulfill some social rule and some assumed responsibilities. However, how much autonomy, intentionality, and responsibility should an AI system be allowed is a challenging question and issue. (4) Accountability. When an AI system or agent fails in a specified task and results in bad consequences, who should be responsible. The undesirable consequence may be caused by many factors, such as the programming codes, input data, improper operation, or other factors. (5) Human Rights Laws. Without training in human rights laws, designers and operators of AL apps may infringe and breach essential human rights without even realizing it.(6) Social Impact of Ethical AI Issues include Automation and Job Replacement.

Al Otaibi (2018) pointed out yet another set of AI ethical considerations including Unemployment , Inequality, Humanity (AI apps cannot be made to respond intelligently towards building relationships and in case of human beings they are limited in terms of attention and kindness.), Artificial Stupidity (It is very difficult to guard against the mistakes made by the machines. Especially in the area of security, finance and healthcare industries) , and Security (Many countries like USA, Japan, China, Iran, Israel, etc. are having automated drones and in the battle field they are very much destructive and can win the wars without the human presence. The magnitude of destruction they can cause to human beings is unimaginable)

Munich Center for Technology in Society Institute for Ethics in Artificial Intelligence (2020, p. 3) in their research brief on 'AI Ethics:

Why Does it Matter?' raised a number of questions like: Why has AI ethics become such an important topic today? Why do we need AI ethics? And what are the most important ethical dilemmas associated with AI right now? "The question (also) arises whether and where exactly AI creates genuinely new problems or whether problems that are already qualitatively known become more urgent". Hence, exploring the consequences of the deployment of AI systems in society is fundamental to ensuring such systems develop in a way that supports the well-being of humans, which should be the ultimate goal. "While AI-based decision-making opens up possibilities that were previously not thought of, it does not operate outside of the domain of ethical responsibilities". "Thus, it is extremely important that such a powerful 'tool' is accepted and understood".

The most important ethical dilemmas related to the development of AI reported by respondents included accountability, autonomy, privacy, fairness, (public) trust, explainability and interpretability. Hence, there is an urgent need to discuss real-world scenarios and "consider ethics at every stage of AI design, development and deployment".

In spite of a number of dilemmas and concerns related to AI, we have been experiencing great contributions achieved through its applications.

Benefits of AI **in Education**

Many bold claims have been made about the benefits that artificial intelligence (AI) could bring about for learners. Holemes et al (2021) provided some benefits of AI in education. The benefits stem from three fundamental factors.

AI can increase capacity within education systems and increase the productivity of educators. By reducing teacher workload and extending affordable and high quality learning opportunities, AI tools could address teacher recruitment and retention problems, and enable life-long learning provision globally.

AIEd can provide valuable insights that can enhance teaching and learning, and support learners' well-rounded development. As Professor Rose Luckin argues, "AI is a powerful tool that can open up the 'black box' of learning, providing a deep, fine-grained analysis of what pupils are doing as they learn, meaning their learning can be 'unpacked' as it happens." This functionality can be used to provide insights to both teachers and learners.

AIEd can deliver autonomous learning recommendations. The Education Endowment Foundation has shown that one to one tuition delivered by experienced professionals, in coordination with classroom teaching can result in the equivalent of five additional months of progress

for learners per year. Through autonomous learning recommendations, AIEd enables individualized instruction and personalized learning at scale.

Flores and Penalvo (2023, p. 38) assure that AI is beneficial in the education as it provides Universal access since it can make classrooms available to everyone, including those who speak different languages or those with visual or hearing disabilities. For example, "Presentation Translator", a free add-in for PowerPoint, helps create real-time captions when a teacher is teaching a class.

Intelligent tutoring systems: AI can be used to tutor students based on their difficulties. For example, air force technicians use an intelligent tutoring system called SHERLOCK to detect electrical system problems in aircraft. Virtual facilitators: With trending technology such as gesture recognition, a teacher can be replaced by a robot. AI, 3D gaming, and computer animation are used to create realistic virtual characters and interactions.

Intelligent content can be created from digitized guides, customized learning interfaces, and digital curricula through various media such as audio, video, and online wizard. And Automate administration tasks: Educators spend much time grading homework and exams. AI systems can grade multiple-choice questions and are close to accessing written answers.

In this context, educational AI applications, like those within broader learning analytics, can be broadly categorized into system-facing; learner-facing; or teacher-facing applications, although there may be overlap. Baker and Smith (2019, as cited in Ungerer and Slade, 2022) define these as:

- System-facing tools assist administrators and managers at an institutional level in generating information, for instance, about attrition patterns across units.
- Learner-facing tools provide software that students use when learning subject matter, namely adaptive or personalized learning management systems or intelligent tutoring systems (ITS).
- Teacher-facing systems automate tasks such as administration, student feedback, assessment and uncovering plagiarism, thereby reducing workload. AIED tools can also summarize student progress, enabling educators to provide them with relevant support when necessary.

OECD (2020) concluded that artificial intelligence is embedded in education in many technological innovations that provide learning analytics, recommendations and diagnosis tools in various ways and for various purposes. There are many examples of promising uses both in the

classroom and at the system levels, and address different stakeholders: students, teachers, administrators, parents, as well as policy makers.

AI could be used "in pursuit of beneficial outcomes for people and the planet, such as augmenting human capabilities and enhancing creativity, advancing inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and protecting natural environments, thus invigorating inclusive growth, sustainable development and well-being."

Zawacki-Richter et al. (2019) summarize the potential uses of AI in learning analytics. These are further explored below.

Profiling and Prediction: AI has the ability to accurately predict students' academic outcomes that might be highly beneficial. Predictive analytics solutions focusing on student retention involve the development of indicators to identify at-risk students and to predict dropout. AIED may also especially serve to enhance student support in open and distributed learning contexts.

Intelligent Tutoring Systems (ITS) whose functions are mainly linked to teaching course content, diagnosing strengths or gaps in students' knowledge and providing automated feedback, curating learning materials based on students' needs and facilitating collaboration between learners.

Assessment and Evaluation with particular focus on automated grading; feedback; evaluation of student understanding, engagement and academic integrity; and evaluation of teaching (Zhang & Asian, 2021). Studies suggest a significant correspondence between the marking of AES systems and that of human markers, implying that automated grading could diminish the time and cost involved in appointing assessors for large-scale marking.

Therefore, it is fair to say that many believe that AIED has significant potential to facilitate timely and personalized learning support to large groups of students, particularly in open and distributed learning environments.

Examples of Current applications of AI in Education are reported by Ungerer and Slade (2022):

ITS and ILE. Intelligent Tutoring Systems (ITS) and Intelligent Learning Environments (ILE) are among the primary expressions of AI in education, with growing adoption in a variety of institutional learning environments. An ITS replicates the role of a human tutor, monitoring learner progress and adjusting the focus and difficulty of instructional content based on the learner's interest and aptitude, as well as providing timely information and guidance. Intelligent learning environments operate in a similar manner, by incorporating aspects of ITS, and adaptive

educational hypermedia systems (AEH), which deliver interactive lessons in multiple mediums and in an adaptive manner.

Chatbots and assistants. Chatbots, also known as conversational agents or dialogue systems, are interfaces that allow people to have a conversation, simulated or otherwise, with a machine (Henderson et al., 2018). Chatbots, in this sense, can serve as valuable classroom aides, responding to discussion board posts and emails regardless of the availability of a human representative.

Grading and feedback. One of the roles often taken by classroom assistants is in relieving teachers of time-consuming basic tasks, such as grading. Multiple-choice questions have long been favored due to their ability to be quickly graded by a human or a machine, but multiple-choice questions can be more difficult and time-consuming to construct while being less suited to higher-order learning assessment than other formats. Software powered by AI has been designed to alleviate some of this tension. A tool called Gradescope has been employed by over 550 colleges in order to streamline and improve assessment and grading (Blumenstyk, 2018). Other tools allow for qualitative analysis of writing, and natural language processing can enable the automation of grading far beyond multiplechoice questions.

Student monitoring and support. Artificial intelligence can also augment teachers' ability to monitor and support their students. This includes tools like GoGuardian, which can replace traditional keyword-based and URL-level security monitoring software in school settings (Pierce & Hathaway, 2018). Using AI to analyze page content and learner text entry within their contexts, systems like this can alert administrators to students who might be accessing inappropriate materials or flagged behaviors, like content which might be indicators of suicidality (Friedman, 2019b; Pierce & Hathaway, 2018). The increased monitoring and analysis of student data can also increase the ability of teachers and administrators to discover when and how students might need additional support.

What is machine learning (ML**) in Education?**

Put simply, ML is: the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions (Faggella, 2018). Specifically, ML is: (A) subfield of artificial intelligence. Its goal is to enable computers to learn on their own. A machine's learning algorithm enables it to identify patterns in observed data, build models that explain the world, and predict things without having explicit preprogrammed rules and models (Maini & Sabri, 2017, p. 9).

Briefly, some of the different types of ML are:

- Supervised learning: Qualified people label or classify initial input data to train an algorithmic model to identify patterns and make predications when new data is given to it. The algorithm learns from experience that is guided by a human labelling the data.
- Unsupervised learning: In this type of ML, algorithms create their own structure (features) that can be used to detect patterns and classifications in unlabeled data. Unsupervised learning is used to explore and detect patterns when an outcome is unknown or not predetermined. It is possible that with large enough data sets, unsupervised learning algorithms would identify patterns in behavior or other phenomena that were previously unknown.
- Reinforcement learning: This has an algorithm interacting with a specific environment to find the best outcome through trial and error without training: 'The machine is trained to make specific decisions. … (It) learns from past experience and tries to capture the best possible knowledge to make accurate … decisions' (Ramzai, 2020).
- Deep learning: Associated with artificial neural networks (ANN) this type of ML is inspired by the way neurons connect in the human brain. It has numerous layers of algorithms that interact to model data and make inferences. There are multiple ANNs at lower levels of abstraction to effectively solve chunks of a problem and provide these partial solutions to ANNs at higher levels to derive a larger solution (LeCun, Bengio & Hinton, 2015).

Deep learning is being used to understand complex data such as natural language processing which involves complicated vocabularies or machine vision processing that has intricate pixel information. However, AI is known to have issues with classification bias and errors.

AI ethical Dangers

As noted above the apps and tools of AI pose many dangers and raise concerns. Ungerer and Slade (2022) contend that although it is believed that AI systems would support innovation, they may unintentionally also replicate traditional education approaches. Despite concerns that traditional assessment does not efficiently measure student learning and progress, assessment data are often used in predictive analytics systems. Educational institutions may consequently be kept from incorporating more reliable, authentic types of assessment (Berendt et al., 2020). Selwyn (2018a) expresses concern that certain benefits offered by human educators may be inadvertently lost when focusing on automating teaching. Educators' knowledge is often based on their own

learning; they connect with students at both a cognitive and social level; they use natural speech; they use their bodies during teaching; and they are able to improvise (Selwyn, 2018a). Many of the human attributes relevant to effective teaching, namely creativity, innovativeness and spontaneity, are absent in computer systems.

There are also concerns involving ethics, morals and values when machines make education decisions (Selwyn, 2018b). It is not clear how automated systems decide which students are first highlighted—i.e., struggling students or those who excel. Human teachers draw upon their empathy, knowledge and experience. It may be virtually impossible to automate decisions incorporating human conscience.

Human Rights Concerns. Although universities regularly use student data to monitor and predict student performance, students' and staff views about this process, social and ethical issues, ethical guidance and policy development are often not considered. Some of the ethical issues involved in using student data include transparency, consent, and the right to seek redress (Slade & Prinsloo, 2013). Chaudhari et al. (2019) highlight further concerns around data ownership, privacy, and digital exclusion resulting from algorithmic biases as the major challenges associated with AIED that should be addressed.

Data Ownership. Students should have control over their own data, for instance, in deciding whether it should reflect achievements throughout their lives. When students or consumers use online platforms or services, they must accept associated Terms and Conditions (TACs) and in doing so, often 'sign away' their rights and their ownership.

Data Privacy and Consent are of the main challenges in using AI-based intelligent agents in educational environments, because it involves private and personal information. Students may not realize the full implications of their actions when they provide consent for their data to be captured, especially when the accompanying TACs are presented in a complex, legal manner and they consequently lose control of their data (Chaudhari et al., 2019). Student data may be comprehensive and located on various administrative systems and commercial platforms, complicating the issue of who owns and controls that data.

Digital Exclusion Due to Algorithmic Biases. Irrespective of whether procedures for ensuring data privacy and transparency are in place, complete transparency is almost certainly unattainable because of the obscure nature of the machine learning algorithms supporting AIED (Burrell, as cited in Chaudhari et al., 2019). Even if they are open to scrutiny, algorithms may be difficult to interpret (Berendt et al., 2020). Willis et al. (2016) caution that the ideas and thinking underlying

algorithmic decision-making tend to be covert and rarely open to scrutiny.

A key problem is the often 'black box' components of ML. This means that the algorithmic 'decision making' processes between inputs and outputs is not transparent either because the algorithms are proprietary (the property of companies and governments who will not open these for independent review) or so complex in their operation, like ANNs, that the machine's decision-making processes are not wholly explainable even to the scientists who develop the systems (Campolo, Sanfilippo, Whittaker and Crawford, 2017). 'Black box' AI which has limited transparency can materially affect life opportunities — it can determine if someone gets a job interview, which learners gets categorized as 'at risk' of attrition or failure, or who has access to a particular curriculum pathway in an intelligent tutoring system. Many ethical and governance issues arise with the use of AI and the types of ML it utilizes.

The Ethics Guidelines for Trustworthy AI (2019) and Pedro et al. (2019) highlighted potential harms to learners due to the application of AI.

In the Human Agency and Oversight guideline Potential harms to learners are evident in decreasing Learners' agency if AI systems reduce independent and introspective thought, and lead to the underdevelopment of higher-order thinking skills and self-regulation, in addition, Learners could become over reliant on AI systems, hence diminishing their overall autonomy.

In view of the Technical Robustness and Safety guideline Potential harms to learners can be exemplified in that highly sensitive data relating to individuals' learning processes (including data on emotions) could be hacked by malicious parties, and AI systems could make inaccurate recommendations - e.g. on which university course to study, which career to pursue or on how to address a mental health condition - that could harm learners.

Potential harms to learners related to Diversity, non-discrimination and fairness are embodied in the fact that due to algorithmic bias, some groups of learners could be unfairly discriminated against. For example, a predictive analytics tool that predicts university students who are most at risk of dropping out of university may overstate the likelihood of pupils from lower socio-economic backgrounds dropping out and therefore lead to discriminatory actions. An AI system could also be more effective for male students due to being trained on datasets that included more males than females.

Concerning the Privacy and data governance guidelines: highly intimate data - including a learners' strengths, vulnerabilities, behavioral habits, and biometric information - could be inappropriately used to exploit, manipulate or oppress learners. Besides, Learners could lose control over who had access to their academic performance data, which could lead to employers or educational institutions having access to a disproportionately high level of information on applicants against their will

Lack of Transparency is manifested in that a learner could be denied a place on a particular university course, and have no means of seeking redress or an explanation due to the inherent opacity of the system. Furthermore, Accountability is also at risk since AIEd could erode accountability structures in school, and cause difficulty with deciding who is accountable for poor performance, or inappropriate or harmful outcomes.

Pedro et al. (2019) admitted that the concerns surrounding AI and its impact on education include access to educational systems through machine learning algorithms, potential discrimination, inadequate recommendations for certain groups of students, personal data concentration, liability, and the impact of AI automation on teacher jobs. The challenge lies in protecting personally identifiable information and privacy preferences, particularly for young learners who cannot provide express consent. Consequently there is a need to address the ethical concerns related to data collection, use, and dissemination.

The case of the Chatbots in Education

The integration of AI and chatbots into education and research has become more prevalent in recent years. Chatbots are automated conversational agents that use natural language processing and machine learning algorithms to interact with users in a human-like manner. However, the increasing use of AI and chatbots in these fields also raises ethical challenges that need to be addressed (Akgun & Greenhow, 2021).

Kooli (2023) conducted a literature review that aimed to explore the ethical challenges of using AI and chatbots in education and research, with a focus on the major observed issues. Chen and et al. (2020) provided potential benefits, including personalized learning, increased accessibility, and improved efficiency. The authors also discussed some ethical challenges associated with the adoption of AI in education. Zhang and Aslan (2021) made a comprehensive review of AI in education and its potential benefits, including improved student engagement and motivation, enhanced assessment and feedback, and increased efficiency and cost-effectiveness. Pedro et al. (2019) highlighted the advantages of the integration of AI in the educational field. The researchers noted that

AI in education can automate repetitive tasks, free up more time for complex cases, and upskill the workforce to become AI-ready.

The observed advantages and disadvantages of the use of AI in education need further investigation, especially with the emergence of the latest powerful AI based chatbot, named ChatGPT. The use of chatbots by students has become increasingly prevalent in recent years, as technology continues to advance and education moves towards online and hybrid models. While chatbots can provide students with quick and convenient access to information, they also present a number of dangers and ethical concerns. Recent research (Dwivedi et al., 2023 and Mendoza et al., 2020) analyzed the development and implementation of a chatbot designed to support student–teacher interaction. The chatbot was integrated into an online platform used in a university course and provided assistance to students regarding course content and assignments. The results show that the chatbot was positively evaluated by the majority of the students and was seen as a useful tool for facilitating communication with their teachers. This research concluded that chatbots have the potential to enhance student–teacher interaction and improve the overall learning experience.

A study by King (2023) on ChatGPT explored the ethical challenges associated with the use of chatbots in educational assessments. The authors highlighted the potential for cheating and the need to ensure the integrity of assessments when using chatbots. They also noted the importance of considering the ethical implications of AI systems and the need to ensure that chatbots are unbiased and fair. Thus, the use of chatbots in educational assessments has a potential for cheating. Students may use chatbots to cheat during exams or other assessments by inputting questions and receiving answers in real time. This undermines the integrity of assessments and can result in unfair advantages for students who use chatbots.

One ethical challenge associated with the use of chatbots in education is the potential for the technology to replace human interaction and expertise. This is particularly concerning in fields such as counseling and mental health, where students may seek emotional support from chatbots instead of trained professionals. A study by Pesonen (2021) found that although chatbots were perceived as useful and helpful by college students, they were not seen as a substitute for human support. Another ethical challenge is the potential for bias in chatbots. AI systems are only as unbiased as the data they are trained on. If the data used to train chatbots are biased, then the chatbot's responses may also be biased. This could result in unfair assessment outcomes and could potentially

perpetuate discrimination and inequality in the education field (Pedro et al., 2019)

Kooli (2023) concluded that Chatbots have emerged as a promising educational tool, with the potential to enhance the learning experience by providing personalized and immediate feedback to students. However, the use of chatbots in the educational field also raises ethical challenges that need to be addressed.

Results reported by Kooli (2023) indicated that Using chatbots or any other artificial intelligence tools to answer exam questions is a form of cheating and academic misconduct, and goes against the fundamental principles of learning and academic integrity. It could also severely affect the student's academic progress and knowledge-acquiring processes by hindering their critical thinking skills, creativity, and ability to apply the concepts learned to real-world situations. In addition, students may use chatbots to cheat on exams or assignments, which could compromise the reliability and validity of the assessment results. Receiving help from chatbots leads to unfair advantages and inaccurate evaluations of the students' knowledge and skills. This could result in a generation of students who are overly reliant on technology and unable to think critically or creatively.

Several reasons could be advanced against the delegation of the research assistance mission to chatbots. AI systems can be trained on biased data, leading to biased results. This can compromise the validity of research findings and harm communities. Additionally, these machines could lack the expertise and knowledge of human researchers, and may produce results that are inaccurate, incomplete, or irrelevant.

Chatbots lack the ability to understand context and nuance, which is essential in many areas of research. This can result in errors and the misinterpretations of data. Sometimes, chatbots lack the creativity and critical thinking skills of human researchers. This can limit the scope of research and the potential for new discoveries.

Finally, the use of chatbots in research raises ethical concerns, such as those related to data privacy, data security, and exploitation of participants. Therefore, chatbots should not be used as a substitute for human researchers, but as a complementary tool to aid in research, under the supervision and control of human experts.

Again, chatbots lack creativity and the ability to generate new ideas, which are essential for advancing scientific and academic research. Additionally, the results generated by chatbots often require human interpretation and evaluation to be meaningful and actionable. Finally, the use of chatbots in research raises ethical concerns about the accuracy and

reliability of the data collected, and the potential for chatbots to perpetuate harmful biases and discrimination.

Lund et al. (2023) highlighted some ethical dangers of OpenAI's ChatGPT which is a generative pre-trained transformer, that uses natural language processing to fulfill text-based user requests (i.e., a "chatbot"). They then discussed the app. in relation to its potential impact on academia and scholarly research and publishing. It is seen as a potential model for the automated preparation of essays and other types of scholarly manuscripts. However, the ethical implications of this technology and its underlying GPT-3 technology have not yet been fully considered (Gonzalez-Padilla, 2022).

Authorship, Copyright, and Plagiarism

Authorship attribution is a major issue concerning the generation of new knowledge by intelligent agents. There may be questions about the ownership of the content generated by the model (Schönberger, 2018). If a user provides input data to the model and the model generates content based on that input, it could be argued that the user owns the copyright to the generated content.

In terms of copyright, there may be concerns about the use of third-party content within the generated manuscripts (Beaza-Yates, 2022). Depending on the nature of the use, it may be necessary to obtain permission from the copyright owner or to rely on a defense such as fair use.

Plagiarism concerns arise from copyright issues, and AI has been previously known in journalism for plagiarism. As a result, the ethics of using ChatGPT to generate scientific articles have been widely discussed. Plagiarism is not limited to copying text but also includes paraphrasing text, methods, graphics, ideas, and any other product of intelligence that belongs to another person.

Citation Practices. In the time that citing sources demonstrates the writer's expertise in their field, showing that they are familiar with the existing research on a particular topic, the use of ChatGPT has been found to produce academic essays with missing references. While ChatGPT may be able to provide intext references in future versions, it is important to consider the potential consequences of relying solely on an automated tool like ChatGPT instead of reviewing the literature in depth (King, 2023).

Academic integrity – or using AI language models to cheat – has been by far the biggest potential issue of AI covered recently in the media. There have been widespread fears that students will use language models like ChatGPT to write essays, answer questions, and cheat on assignments. These fears seem particularly strong in secondary and

tertiary education, where many assignments are provided in written form. It is also still unclear to what extent using an AI constitutes "cheating". It is not, plagiarism as the output of the model is not copied from another source.

- AI models produce original output that cannot be detected by plagiarism detectors (and "AI detectors" are currently hit and miss).
- AI can be used to produce "fake news" or deliberately misleading information.
- English: If an AI can write an essay, what is the point of writing essays? how do essays help us to build knowledge and not just demonstrate it?
- Religion: What are the ethical and moral implications of academic integrity?
- Humanities (Legal studies): Is using an AI to write an essay cheating? Australia has laws against contract cheating (getting another person to produce an essay on your behalf). Is using an AI a form of contract cheating and therefore illegal?

Intellectual property. Large language models like ChatGPT incorporate huge amounts of other writers' work. Where writing is publicly available – such as out of copyright books or CC journalism and articles – it can be incorporated into the dataset. Even when writing is protected by copyright it can become part of the datasets.

There are also question marks over who owns the copyright to materials produced by AIs such as image generators and language models. AI language models and image generators may breach copyright by appropriating others' work. Does producing a piece of writing in English in the style of another author infringe their intellectual property rights?

Privacy and security. Privacy is a major concern in the development and use of AI systems. One of the most prominent examples of these issues can be found in the use of facial recognition technology. This technology, which is used in a variety of applications such as security, surveillance, and marketing, has been criticized for its potential to violate individuals' privacy and civil rights.

"Datafication" is a term used for turning all parts of our life into a data point to be fed into an AI algorithm. From location data to health data, shopping habits, likes, clicks, and views, almost every interaction we have with technology is fed into an algorithm somewhere.

One major ethical concern with "datafication" is that fact that the users become the products, and that the free-labor of the users is used to generate capital for the platform owners.

Further Para-educational Ethical Challenges of AI include

Malicious Use: Artificial intelligence, like human intelligence, will be used maliciously, there is no doubt.

For example, AI-assisted computer-hacking and lethal autonomous weapons systems, and "killer robots." are few examples.

Automating Ethics: One strength of AI is that it can automate decision-making, thus lowering the burden on humans and speeding up – potentially greatly speeding up—some kinds of decision-making processes. However, this automation of decision making will present huge problems for society, because if these automated decisions are good, society will benefit, but if they are bad, society will be harmed.

Moral Deskilling & Debility: If we turn over our decision-making capacities to machines, we will become less experienced at making decisions. In its most extreme form, if AI starts to make ethical and political decisions for us, we will become worse at ethics and politics. We may reduce or stunt our moral development precisely at the time when our power has become greatest and our decisions the most important.

AI Consciousness, and "Robot Rights" Morally speaking, we can anticipate that technologists will attempt to make the most human-like AIs and robots possible, and perhaps someday they will be such good imitations that we will wonder if they might be conscious and deserve rights—and we might not be able to determine this conclusively. If future humans do conclude AIs and robots might be worthy of moral status, then we ought to err on the side of caution and give it.

AGI and Superintelligence: If or when AI reaches human levels of intelligence, doing everything that humans can do as well the average human can, then it will be an Artificial General Intelligence—an AGI—and it will be the only other such intelligence to exist on Earth at the human level.

The advent of AGI or superintelligence will mark the dethroning of humanity as the most intelligent thing on Earth. We have never faced (in the material world) anything smarter than us before. As we encounter AGI and superintelligence, we ought to keep this in mind; though, because AI is a tool, there may be ways yet to maintain an ethical balance between human and machine.

AI-powered Addiction Smartphone app makers have turned addiction into a science, and AI-powered video games and apps can be addictive like drugs. AI can exploit numerous human desires and weaknesses including purpose-seeking, gambling, greed, libido, violence, and so on.

Addiction not only manipulates and controls us; it also prevents us from doing other more important things—educational, economic, and social.

Effects on the Human Spirit: AI will have effects on how humans perceive themselves, relate to each other, and live their lives. What will happen to the human spirit if or when we are bested by our own creations in everything that we do? Will human life lose meaning? Will we come to a new discovery of our identity beyond our intelligence?

In conclusion, if we instead find our humanity not in our brains, but in our hearts, perhaps we will come to recognize that caring, compassion, kindness, and love are ultimately what make us human and what make life worth living. Policies and measures to safeguard the use of AI are a dire need.

Recommendations and measures to safeguard Ethical AI

The best practice principles according to STM association (2021, p. 7) for ethical and trustworthy AI may be grouped in five categories that are considered to be highly specific and to-the-point. These recommendations are briefly reported below:

1. Transparency and Accountability can most clearly be achieved on the level of the data used in AI training and input, as well as in the use of AI technology in publisher's tools, processes and services. Publishers are both committed and uniquely positioned to support transparency in data – providing information on its provenance and ensuring data is available in a structured and consistent format – while clarity and transparency around the use of IP and copyright materials is required.
2. Quality and Integrity should be at the heart of the AI lifecycle, from the design and building of algorithms, to inputs used to train AI tools and services, to those used in the practical application of AI.
3. Privacy and Security include several principles that focus on data protection, data privacy and security can and should be used to respect and uphold privacy rights, data protection and ensure the security of datasets used in training or operating AI systems.
4. Fairness .In order to achieve fairness, data selection and the application of AI must be carefully analyzed, planned, reviewed, and continuously monitored. Feedback mechanisms should be developed that can report cases or concerns of bias.
5. Sustainable Development. The multi-disciplinary nature of AI systems makes them ideally positioned to address areas of global concern, such as the United Nations Sustainable Development Goals.

Pedro et al. (2019) also discussed the challenges and policy implications of introducing AI into education. The challenges include developing comprehensive public policies for sustainable development, ensuring inclusion and equity, preparing teachers for AI-powered education, developing quality and inclusive data systems, making research on AI in education significant, and addressing ethical concerns related to data collection, use, and dissemination. These challenges require international and national partnerships and public discussions on ethics, accountability, transparency, and security.

Mittelstadti (2019, p. 9) suggested a number of Recommendations that could contribute to safeguard AI applications. These include defining sustainable pathways since binding and highly visible accountability structures as well as clear implementation and review processes are needed at a sectoral and organizational level.

Supporting 'bottom-up' AI Ethics in the private sector because increased support and access to development settings should be made available to support multi-disciplinary bottom-up research and development in AI Ethics, particularly in commercial development contexts currently closed to external scrutiny.

Licensing developers of high-risk AI: To encourage long-term recognition of ethical commitments, it may be necessary to formally establish AI development as a profession with equivalent standing to other high-risk professions. It is a regulatory oddity that we license professions providing a public service, but not the profession responsible for developing technical systems to augment or replace human expertise and decision-making within them.

Pursuing ethics as a process, not technological solutionism: The rationale seems to be as follows: insufficient consideration of ethics leads to poor design decisions which create systems that harm users. This attitude is misguided. It is foolish to assume that very old and complex normative questions can be solved with technical fixes or 'good' design alone. Ethics is a process, not a destination. The real work of AI Ethics begins now: to translate and implement our lofty principles, and in doing so to begin to understand the real ethical challenges of AI

According to the IEIAE Interim report (2019) the following blueprint is intended to form the basis of an ethical framework that may safeguard or mitigate the AI- related harms and concerns laid out above.

The first recommendation intends to orientate the use of AI in Education towards benefits for learners.

- AI should only be used for educational purposes where there are clear indications that it will genuinely benefit learners either at an individual or collective level.

- AI should not be used for educational purposes in cases where the risks posed to learners are at all significant.
- To support the achievement of an optimal balance between underuse, and overuse/misuse of AI in Education, AI may only be used in cases where there is any risk of harm to learners - at an individual or collective level - if the following conditions are met: A) Decisive steps have been taken - and openly communicated in an understandable way - to mitigate the risks involved; B) The benefits to learners outweigh these risks; C) The risks to learners are not significant; D) Monitoring of conditions (A)-(C) is in place; E) If monitoring in (D) shows that conditions (A)-(C) are not fulfilled, the use of AI is to be improved or terminated.

A number of Key guidelines for Trustworthy AI were presented by High-Level Expert Group (2019, p. 2). Excerpts from these guidelines are as follows:

Ensure that the development, deployment and use of AI systems meets the seven key requirements for Trustworthy AI: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability.

- Consider technical and non-technical methods to ensure the implementation of those requirements.

Foster research and innovation to help assess AI systems and to further the achievement of the requirements; disseminate results and open questions to the wider public, and systematically train a new generation of experts in AI ethics.

- Communicate, in a clear and proactive manner, information to stakeholders about the AI system's capabilities and limitations, enabling realistic expectation setting, and about the manner in which the requirements are implemented.
- Facilitate the traceability and auditability of AI systems, particularly in critical contexts or situations.
- Involve stakeholders throughout the AI system's life cycle.
- Foster training and education so that all stakeholders are aware of and trained in Trustworthy AI.
- Continuously identify, evaluate, document and communicate these trade-offs and their solutions.

Some clear highlights emerged from the research brief of the Technical University of Munich (2020, p. 6):

- AI ethics as a field has a fundamental role to play in developing "trustworthy" AI tools. Priorities in the field include protecting individual interest and privacy, ensuring social justice and

environmental protection, as well as defining accountability, responsibility and transparency.

- There is a strong need for an interdisciplinary approach to research that moves out of the general/abstract and provides instruction for applying AI ethics in a way that is technically feasible
- Finally, universities have an opportunity to ingrain the importance of ethics in AI into the mindsets of our future developers and policymakers early on through integrating AI ethics and the importance of interdisciplinary, multi-stakeholder approaches into core curricula for engineers and social scientists alike.

The nine ethical pillars can be concluded from UNESCO (2021), Schoentgen and Wilkinson (2021) and Huang et al. (2022) are: Do No Harm is the first and foremost among these principles.In the event of possible occurrence of any harm to human beings, human rights and fundamental freedoms, communities and society at large or the environment and ecosystems, the implementation of procedures for risk assessment and the adoption of measures in order to preclude the occurrence of such harm should be ensured.

Ensure Awareness: in developing the awareness of individuals and groups so that they can take informed action and decisions related to the technology. Universities should proactively develop foundational knowledge about AI so that all stakeholders can understand what AI is, what it can and can't do, and where it is present in applications and systems especially when it operates in invisible ways.

Enforce Safety and security: Unwanted harms (safety risks), as well as vulnerabilities to attack (security risks) should be avoided, prevented and eliminated throughout the life cycle of AI systems. Thid can be done by the development of sustainable, privacy-protective data access frameworks that foster better training and validation of AI models utilizing quality data.

Safeguard Transparency: AI is often described as an opaque technology. It is commonly invisibly infused into computing systems in ways that can influence our interactions, decisions, moods and sense of self without us being aware of this (Cowrie, 2015). Technologists have described four technical ways in which AI systems can be made transparent, these are: • Traceability which refers to technical inspection of which norms have been implanted, for which contexts, and how norm conflicts are resolved by system. This can reveal biases which may have been built into a system. • Verifiability through formal mathematical techniques including a log of ethical reasoning that should be available for inspection. • Non-deception and honesty where systems are designed to accurately represent what the system is capable of doing to the person

using it. • Intelligibility which entails a clear requirement for a system to be able to explain its own reasoning to a user, at a level commensurate with human reasoning, when it suspects user confusion (IEEE, 2019).

Provide Explainability: involves the responsibility of manufacturers, vendors and procurers of AI technology to clearly elucidate: • what the technology should do, can and can't do • the educational and societal values and norms on which it was/is trained and acts • the learning and pedagogical theory and domain knowledge on which it is based • evidence of its efficacy for learning for diverse groups of students • arrangements for data collection, deidentification, storage and use including third party or other sharing agreements, and those for sensitive information such as biometrics or measures embedded in affective computing applications

Secure Human oversight and determination: Member States should ensure that it is always possible to attribute ethical and legal responsibility for any stage of the life cycle of AI systems, as well as in cases of remedy related to AI systems, to physical persons or to existing legal entities. Human oversight refers thus not only to individual human oversight, but to inclusive public oversight, as appropriate.

Ensure Accountability: Governance of AI will entail new ways of thinking about the interconnections and tensions between proprietary interests, (public and transparent) algorithmic auditability, regulatory standards, policy and risk assessment, legal obligations, and broader social, cultural and economic responsibilities. Thus, lack of transparency increases the risk and magnitude of harm when users do not understand the systems they are using, or there is a failure to fix faults and improve systems following accidents. Lack of transparency also increases the difficulty of ensuring accountability (IEEE, 2019, p. 27).

The learning analytics literature provides a good window into the scope and types of data that are being harvested and that can be combined to create big data for ML in higher education. These include (a) provided data which is intentionally given by individuals, for example, when they fill out a form; (b) observed data recorded automatically, for example, via online cookies or sensors for biometric (for example, facial recognition); (c) derived data produced from other data based on simple calculations that may provide proxy insights; and (d) inferred data produced by using analytic methods to find correlations between datasets used to categorize or profile people (Abrams, 2014).

Guarantee Trust. The trust in the development, deployment and use of AI systems is not only related to the inherent characteristics of the technology, but also related to the quality of the socio-technical system involving AI applications. Therefore, moving towards trustworthy AI not

only concerns the trustworthiness of the AI system itself, but also requires a holistic and systematic approach that covers the trustworthiness of all participants and processes that are the entire life cycle of the system.

Guard Dignity. Human dignity encompasses the belief that all people possess an intrinsic value that is tied solely to their humanity, i.e., it has nothing to do with their class, race, gender, religion, abilities, or any other factor other than them being human, and this intrinsic value should never be diminished, compromised, or repressed by other people nor by technologies like AI.

Conclusion

A new era of education and research based on chatbots and artificial intelligence (AI) is quickly growing. However, the application of these new systems is associated with several challenges and limitations, mainly related to ethics. This paper explores the potential use of AI systems and chatbots in the academic field and their impact on research and education from an ethical perspective.

Today, AI is a strategic technology that offers many potential benefits for citizens, the public good and the economy at large, provided it is human-centric, ethical, and respects fundamental rights and values.

Finally, some critical questions about the ethics and challenges of AI in education are proposed, concluding with the need to add proposals with new research and political actions which could lead to the creation of an ethical observatory of AI for education. The various uses of AI, as well as the ethical challenges related to them, require multidisciplinary and multi-stakeholder engagement, and international collaboration on developing AI governance.

Among others, priorities include protecting individual interest and privacy, ensuring social justice and digital equity, reducing disparities, protecting the environment as well as defining accountability, responsibility and transparency. Consideration of these challenges will help to ensure the development of advanced and “trustworthy” AI systems.

References

Affde. (Ed.). (2021). Impacto de la IA en la industria del e-learning para 2021 y más allá. Affde. Retrieved July 20, 2023, from https://www.affde.com/es/impact-of-ai-in-elearning-industry.html

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. AI Ethic, 2(10), 431–440. DOI:10.1007/s43681-021-00096-7

Alotaibi, S. (2018). Ethical issues and related considerations involved with artificial intelligence and autonomous systems. International Journal of Advanced Computer Science and Applications, 9(4), 35–40. https://thesai.org/Downloads/Volume9No4/Paper_7-Ethical_Issues_and_Related_Considerations.pdf

Australian Government Dept. Industry, Science and Resources. (2019). Australia's Artificial Intelligence Ethics Framework. Retrieved July 14, 2023, from https://industry.gov.au/publications/australias-artificial-intelligence-ethics-framework

Bryson, J. (2020). Reflections on AI. [Q&A with Professor Johanna Bryson]. TUM IEAI. Retrieved July 16, 2023, from https://ieai.mcts.tum.de/wpcontent/uploads/2020/05/Reflections-on-AI-Ethics-FINAL.pdf

Campolo, A., Sanfilippo, M., Whittaker, M., & Crawford, K. (2017). AI now 2017 report. New York: AI Now Institute. Retrieved July 20, 2023, from https://ainowinstitute.org/AI_Now_2017_Report.pdf

Carr, D. (2014). Metaphysics and methods in moral enquiry and education: Some old philosophical wine for new theoretical bottles. Journal of Moral Education, 43(4), 500–515. doi: 10.1080/03057240.2014.943167

Cath, C., Zimmer, M., Lomborg, S., & Zevenbergen, B. (2018). Association of Internet Researchers (AoIR) roundtable summary: artificial intelligence and the good society workshop proceedings. Philosophy & Technology, 31(1), 155-162.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. https://doi.org/10.1109/ACCESS.2020.2988510

Chowdhury, M. (2016). Emphasizing morals, values, ethics, and character education. The Malaysian Online Journal of Educational Science, 4(2)1–16. https://files.eric.ed.gov/fulltext/EJ1095995.pdf

Cormack, A. (2021). A pathway towards responsible ethical AI. Retrieved July 11, 2023, from https://www.jisc.ac.uk/reports/a-pathway-towards-responsible-ethical-ai

Cowie, R. (2015). Ethical Issues in Affective Computing. In R. A. Calvo, S. D'Mello, & J. Gratch, J. (Eds.), The Oxford handbook of affective computing (pp. 334–348). Oxford University Press.

Cox, A. (2021). The impact of AI, machine learning, automation and robotics on the information professions: A report for CILIP. Retrieved July 20, 2023, from https://eprints.whiterose.ac.uk/174522/

CSIRO (n..d.) Artificial intelligence. Retrieved July 19, 2023, from https://www.csiro.au/showcase/ai%20

Dwivedi, Y., Kshetri, N., Hughes, L., Slade, E., Jeyaraj, A., Kar, A., Baabdullah, A., Koohang, A., Raghavan, V., Ahuja, M. et al. (2023) "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://squ.elsevierpure.com/en/publications/so-what-if-chatgpt-wrote-it-multidisciplinary-perspectives-on-opp

Elkins, K., & Chun, J. (2020). Can GPT-3 pass a Writer's turing test? Journal of Cultural Analytics, 5(2), 1–16. https://doi.org/10.22148/001c.17212

European Commission. (2019). Ethics guidelines for trustworthy AI https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai)

Faggella, D. (2018). What is machine learning? Retrieved July 22, 2023, from https://emerj.com/ai-glossary-terms/whatis-machine-learning/

Fidalgo-Blanco, A., Sein-Echaluce, M., & García-Peñalvo, F. (2022). Método basado en Educación 4.0 para mejorar el aprendizaje: Lecciones Aprendidas de la COVID-19. RIED, 25(2), 49–72. https://revistas.uned.es/index.php/ried/article/view/32320

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center Research Publication. http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420

Flogie, A., & Aberšek, B. (2021). Artificial intelligence in education. In O. Lutsenko & G. Lutsenko (Eds.), Active learningTheory and Practice. IntechOpen. https://doi.org/10.5772/intechopen.96498

Flores, J., & Penalvo, F. (2023). Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4) Comunicar, XXXI(74), 35–44. https://doi.org/10.3916/C74-2023-03

Garret, N., Beard, N., & Fiesler, C. (2020). More than "If Time Allows": The Role of Ethics in AI Education. In the Proceedings of the 2020 AAAI/ACM AIES Conference on AI, Ethics, and Society, 272-278. New York, NY. 2020. https//cmci.colorado.edu.

Gillotte, J. (2019). Copyright infringement in ai-generated artworks. University of California, Davis, 53, 2655–2676. https://lawreview.law.ucdavis.edu/issues/53/5/notes/files/53-5_Gillotte.pdf

Green, B. (2021). The Contestation of Tech Ethics: A Sociotechnical Approach to Ethics and Technology in Action. Journal of Social Computing, 2(3), 209–225. Doi: https://doi.org/10.23919/JSC.2021.0018

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), 99–120.

Han, H. (2014). Analyzing theoretical frameworks of moral education through Lakatos's philosophy of science. Journal of Moral Education, 43(1), 32-53. doi: 10.1080/03057240.2014.893422.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E. Baker, T., Shum, S., Santos, O., Rodrigo, M., Cukurova, M., Bittencourt, I., & Koedinger, K. (2021). Ethics of AI in Education: Towards a Community-Wide Framework. International Journal of Artificial Intelligence in Education, **32**, pages504–526. https://link.springer.com/article/10.1007/s40593-021-00239-1

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of Artificial Intelligence ethics. ResearchGate.net. http://www.researchGate.net

IEEE. (2019). Ethically aligned design. A vision for prioritizing human well-being with autonomous and intelligent systems. Retrieved July 19, 2023, from https://standards.ieee.org/content/dam/ieeestandards/standards/web/documents/other/ead1e.pdf

Institute for Ethical AI in Education. (2021). The ethical framework for AI in Education. Retrieved July 13, 2023, from https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf

King, M. (2023) ChatGPT. A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education. Cel. Mol. Bioeng, 16, 1–2. https://doi.org/10.1007/s12195-022-00754-8

Kooli, C. (2023). Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions. Sustainability, 15, 1–15. https://doi.org/10.3390/

Lanier, J. (2018). Diez razones para borrar tus redes sociales de inmediato. Penguin Random House Grupo editorial. https://www.penguinlibros.com/es/ciencia-y-tecnologia/11811-libro-diez-razones-para-borrar-tus-redes-sociales-de-inmediato-9788499929170

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv. Retrieved July 16, 2023, from https://doi.org/10.48550/arXiv.1907.11692

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. Nature Human Behavior.

Lütge, C. (2020). Reflections on AI. [Q&A with Prof. Dr. Cristoph Lütge]. TUM IEAI Retrieved July 16, 2023, from https://ieai.mcts.tum.de/wpcontent/uploads/2020/10/Reflections-on-AIEthics_ChristophL%C3%BCtgeFinal.pdf

Maini, V., & Sabri, S. (2017). Machine learning for humans. Retrieved July 15, 2023, from https://everythingcomputerscience.com/books/Machine%20Learning%20for%20Humans.pdf

Mendoza, S., Hernández-León, M., Sánchez-Adame, L., Rodríguez, J., Decouchant, D., & Meneses-Viveros, A. (2020). Supporting student-teacher interaction through a chatbot. In Learning and collaboration technologies. Human and technology ecosystems (pp. 93–107). 7th International Conference, LCT 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22. Springer International Publishing.

Mittelstadti, B. (2019). Principles alone cannot guarantee ethical AI. Nature Machine Journal, 1, 1–16. Retrieved July 15, 2023, from https://www.researchgate.net/publication/337015694_Principles_alone_cannot_guarantee_ethical_AI

Munich Center for Technology in Society Institute for Ethics in Artificial Intelligence. (2020). AI Ethics: Why Does it Matter? Research Brief. Retrieved July 10, 2023, from https://ieai.sot.tum.de/wp-content/uploads/2020/10/Research-Brief_WhyAIEthicsMatter_Final-1.pdf

Nature Machine Intelligence Editorial Board. (2020). Next chapter in artificial writing. Nature Machine Intelligence, 2, 419. https://doi.org/10.1038/s42256-020-0223-0

Neppel, C. (2020). Reflections on AI. [Q&A with Dr. Clara Neppel]. TUM IEAI. Retrieved July 16, 2023, from https://ieai.mcts.tum.de/wpcontent/uploads/2020/06/Reflections-on-AIEthics_ClaraNeppel_Final.pdf

OECD. (Ed.). (2020). Trustworthy AI in Education: Promises and challenges.. Background paper for the G20 AI Dialogue, Digital Economy Task Force Meeting. Riyadh, 1-2 February 2020. https://www.oecd.org/education/trustworthy-artificial-intelligence-in-education.pdf

OECD. (Ed.). (2021). OECD Digital Education Outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots. OECD. Retrieved July 11, 2023, from https://www.oecd-ilibrary.org/education/oecd-digital-education-outlook-2021_589b283f-en

Oliver, N. (2020). Inteligencia Artificial, naturalmente. Centro de Publicaciones. Ministerio de Asuntos Económicos y Transformación Digital. Gobierno de España. https://bit.ly/3Olt5SE

O´Neil, C. (2017). Armas de destrucción matemática. Como el Big data aumenta la desigualdad y amenaza la democracia. Capitán Swing Libros. https://bit.ly/2Y0njfS

Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development. UNESCO: Paris, France. https://unesdoc.unesco.org/ark:/48223/pf0000366994

Pesonen, J. (2021) 'Are You OK?' Students' Trust in a Chatbot Providing Support Opportunities. In Learning and Collaboration Technologies: Games and Virtual Environments for Learning: 8th International Conference, LCT 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part II; Springer: Cham, Switzerland, 2021; pp. 199–215.

Price, H. (2020). Reflections on AI. [Q&A with Prof. Dr. Huw Price]. TUM IEAI. Retrieved July 16, 2023, from https://ieai.mcts.tum.de/wpcontent/uploads/2020/07/Reflections-on-AIEthics_HuwPrice_Final-pdf

Qoura, A. (2017). Ethics in the teaching profession: The Finnish Model. Port Said Faculty of Education conference, Egypt. Academia.edu. http://www.academia.edu

Ramírez-Montoya, M., Castillo-Martínez, I., Sanabria-Z., & Miranda, J. (2022). Complex thinking in the framework of education 4.0 and open innovation-a systematic literature review. Journal of Open Innovation, 8(1), 1–15. https://www.mdpi.com/2199-8531/8/1/4

Ramzai, J. (2020). Clearly explained: 4 Types of machine learning algorithms. Retrieved July 15, 2023, from https://towardsdatascience.com/clearly-explained-4-types-of-machine-learning-algorithms71304380c59

Schoentgen, A., & Wilkinson, L. (2021). Ethical issues in digital technologies. 23rd Biennial Conference of the International Telecommunications Society (ITS): "Digital societies and industrial transformations: Policies, markets, and technologies in a post Covid world". Online Conference / Gothenburg, Sweden, 21st-23rd June, 2021, International Telecommunications Society (ITS), Calgary. https://www.econstor.eu/handle/10419/238052

Schönberger, D. (2018). Deep copyright: Up- and downstream questions related to artificial intelligence (AI) and machine learning (ML). SSRN. Retrieved July 13, 2023, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3098315

Selwyn, N. (2018a). Six reasons artificial intelligence technology will never take over from human teachers. Retrieved July 15, 2023, from https://www.researchgate.net/deref/http%3A%2F%2Fwww.aare.edu.au%2Fblog%2F%3Fp%3D2948

Selwyn, N. (2018b). The rise of automated teaching technologies: We need to talk about robots. Retrieved July 15, 2023, from https://www.researchgate.net/deref/https%3A%2F%2Flit.blogg.gu.se%2F2018%2F06%2F12%2Fthe-ethical-dilemma-of-the-robot-teacher%2F

The Institute for Ethical AI in Education. (2020). Interim Report: Summary: Towards a Shared Vision of Ethical AI in Education. https://www.buckingham.ac.uk/ research/the-institute-for-ethical-ai-in-education

UNESCO. (Ed.) (2019). Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development. https://backend.educ.ar/refactor_resource/getBook/1097

UNESCO. (Ed.) (2021). Recommendation on the ethics of artificial intelligence.https://unesdoc.unesco.org/ark:/48223/pf0000380455_spa

Ungerer L., & Slade, S. (2022). Ethical Considerations of Artificial Intelligence in Learning Analytics in Distance Education Contexts. In P. Prinsloo, S., Slade, & Khalil M. (Eds.) Learning Analytics in Open and Distributed Learning (pp.105-120). Springer.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada. https://proceedings.neurips.cc/paper/2019/file/dc6a7e655d7e5840e66733e9ee67cc69-Paper.pdf

Zawacki-Richter, O., Marin, V., Bond, M., & Gouverneur F. (2019) Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(39). https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-019-0171-0

Zhang, K., & Aslan, A. (2021). AI technologies for education: Recent research &future directions. Computer and Education: Artificial Intelligence, 2, 100025. https://doi.org/10.1016/j.caeai.2021.100025