

Approaches for Maintaining Ethics in AI-Drive Education Solutions and Foster Sustainability

Corina Marina MIREA*

Bucharest University of Economic Studies, Bucharest, Romania

*Corresponding author: corina.mirea@csie.ase.ro

Andrei BOBOCEA

Bucharest University of Economic Studies, Bucharest, Romania

andrei.bobocea@csie.ase.ro

Abstract. *This article introduces sequential strategies regarding the way to use artificial intelligence as a learning facilitator for both teachers and students while preserving ethics and the legislative norms launched. Learning solutions are proposed to be used both by students, to study at school, but also at home, and by teachers to meet all the needs of students. The intelligent systems that are in general proposed to be part of the educational process can be designed so that the norms of ethics and responsibility are taken into account and as many actors involved in the educational environment, whether they are teachers, students, people who run educational institutions or decide to country level to be as it embraces with a much higher level of confidence artificial intelligence learning systems. For this work, a qualitative analysis was carried out to answer the question of how learning systems implemented with artificial intelligence should be built so that ethical principles are not neglected. While ensuring a sustainable environment.*

Specifically, a review of academic articles was conducted to highlight the concerns and importance of artificial intelligence in education. An analysis was also made of documents produced by specialised institutions such as the Organisation for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organisation (UNESCO), U.S. Department of Education, World Economic Forum, European Commission, European Parliament, Council of the European Union, and The Institute for Ethical AI in Education on the topic of ethics in artificial intelligence and the guidelines outlined for building ethical systems and including sustainability.

In the conclusion section of the article, admissions are made that there is necessary support from the institutions, so it is still possible to make educational systems with artificial intelligence that respect ethical norms and pay attention to sustainability.

Keywords: ethics of artificial intelligence, sustainability, artificial intelligence in education, transparency, human-centered artificial intelligence, sustainable education, ethical framework for artificial intelligence, fairness.

Introduction

What the article proposes is to bring to attention the main elements of ethics in artificial intelligence that the official institutions propose in their documents and that the actors involved in the process of developing, adopting, proposing educational systems or applications that use artificial intelligence should take into account. Next, the sections of this article and their content will be briefly presented. The literature review part shows how the learning systems had a positive impact in the education system, how the applications with intelligent assistants used by the students led to good learning results. The appearance of assistants based on artificial intelligence generative AIs such as ChatGPT is rapidly gaining popularity, reaching 100 million users by early 2023. However, a single country started to develop rules on generative artificial intelligence. This means that

regulation of artificial intelligence is just beginning, but there is a need to be more aware of all the implications it brings (UNESCO, 2023).

Also in this part are the reasons why there should be ethical rules when it comes to the use of artificial intelligence along with the steps taken by organisations such as United Nations Educational, Scientific and Cultural Organisation or the Organisation for Economic Co-operation and Development Council and others to address the topic of ethics.

Special attention should be paid to the possible consequences of the implementation and development of artificial intelligence, which was also specified in the report of the World Economic Forum. Unfavourable outcomes of artificial intelligence are ranked sixth in the top 10-year risk ranking. This puts us in the position of being more aware and ensuring that AI-based systems are fully built to improve end-user processes and activity. As pointed out in the report, following strategies together with setting regulations can help reduce risks (World Economic Forum, 2024).

Besides the ethical aspect related to the use of artificial intelligence in all sectors of activity, special importance must be given to how it will affect the planet we live in and, by default, the environment. Keep in mind that there are elements of Planet Earth that are targeted for degradation: climate change, biodiversity, oceans, water, and air quality. Specifically, we can talk about the excess natural resources used along with the energy consumed by artificial intelligence equipment (Microsoft, 2019).

The results section is based on the analysis of official studies of worldwide organisations from the Organisation for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organisation (UNESCO) and United States of America. Department of Education, World Economic Forum, European Commission, the European Parliament, Council of the European Union and The Institute for Ethical AI in Education and finding common and particularly important principles that artificial intelligence systems should follow including especially applications in education.

What the article aims to do in the results section is to answer the question of whether there are frameworks for developing ethical artificial intelligence learning solutions.

Literature review

The educational environment has seen a real improvement following the use of artificial intelligence in e-learning systems, supported by the fact that students' involvement in the learning process has increased, the educational content adapts according to the learning mode, and therefore the learning outcomes are improved. Furthermore, adaptive learning systems have also been used to support decision making as they use data collected from students and can provide recommendations and set learning paths for them (Gligorea, et al., 2023).

The uses of artificial intelligence in the education sector have expanded from intelligent tutors to artificial intelligence-based systems that are in the role of being trained by users, to learning partners or decision consultants. They bring much greater efficiency to the educational process, providing support for various subjects taught at school. In addition, they engage with learners in the learning process by taking care of simple and basic tasks to create the right context for learners to apply their decision-making or logical thinking skills (Hwang, Xie, Wah, & Gašević, 2020).

The transition from Education 4.0, where artificial intelligence is predominant, to Education 5.0 where the focus is on the cognitive abilities of students as a result of combining artificial intelligence with human capabilities. Moving on to the predominant elements in

Education 5.0 besides intelligent assistants, learning analytics, lifelong learning, the integration of gamification platforms, the ethics of artificial intelligence is of great importance as systems based on artificial intelligence should ensure the respect of some principles in this sphere (Rane, Choudhary, & Rane, 2023).

To raise awareness of the risk and impact that artificial intelligence may have in the future, World Economic Forum has published several forecasts in The Global Risk Report on the risks that may arise globally and their impact in important areas. These focus on climate change, demographic shifts, technological developments, and geostrategy. The report presents areas that will present risks both in the short-term (2 years) and in the long-term (10 years) depending on their severity. Looking at the area of technology and the threats that come from it in the long term, the negative results that artificial intelligence can bring ranks 6th in terms of risk severity. Among the solutions proposed in the paper to mitigate the not-so-good results of artificial intelligence are public awareness along with much more education and knowledge of the field. More specifically, the education system needs to include information that familiarises people with the concepts of artificial intelligence, and, furthermore, key decision makers should be more informed about the power of artificial intelligence. Another solution is to set standards at both the local and national levels to ensure that developments in the area of artificial intelligence are aligned with current needs (World Economic Forum, 2024).

The Organisation for Economic Co-operation and Development Council's position on artificial intelligence focuses both on principles that are addressed to decision makers and on recommendations for policy implementation and cooperation (OECD, 2023).

In the document adopted on 23 November 2021 by the United Nations Educational, Scientific and Cultural Organisation on a set of recommendations related to ethics in artificial intelligence made to ensure that artificial intelligence is used to help humans and not against them and society. Moreover the recommendations made can be used both in formulating policies, regulations and to set actions of companies, institutions taking into account fundamental human rights. Thus in this regard, recommendations are made to address all players using artificial intelligence. The values included in the recommendations are: respect and protection of human rights and freedoms, environmental prosperity, ensuring diversity and inclusion, and last but not least living in an interconnected society where peace and justice prevail. At the same time, the principles we find are: proportionality and do no harm, safety and security, fairness and nondiscrimination, sustainability, right to privacy and data protection, human oversight and determination, transparency and explainability, responsibility and accountability, awareness and literacy, multistakeholder and adaptive governance and collaboration (UNESCO, 2022).

In order to make artificial intelligence-based systems that have humans as the main element, the authors of the article (Lepri, Oliver, & Pentland, 2021) proposed 3 solutions that should be considered. The first concern privacy and data ownership, the second would be transparency and accountability, and the third concerns fairness in decision-making processes that are governed by artificial intelligence.

In building an adaptive learning platform, elements such as the protection of personal data were taken into account, as it uses data and information about children using the platform, including as they advance in age and knowledge. In this regard, the following measures have been implemented: transparency in privacy policies, collection of only necessary data, and safe data storage. All this is done in order to respect the principles of design privacy. In order to support all of the above, several actions have been taken such as informing users, especially parents being a platform for children, about how and for what purpose their data are used. Another aspect would

be that only data that is really important is saved and the way in which certain information is requested is done so as not to put parents in delicate situations, for example, gender selection. As far as data storage is concerned, the platform stores user pictures locally for example (Verhagen, Dalibert, Lucivero, Timan, & Co, 2016).

Adaptive learning aims to customize the learning experience of individual students by leveraging artificial intelligence and other computer algorithms (Kaplan, 2021). The promise of personalization of the learning process is of great interest and its importance is well recognized both at an administrative level (Duncan, 2013) and in society as a whole. The global adaptive learning market was valued at USD 1.9 billion in 2020 and the same report projected a growth to USD 5.3 billion by 2025 (Markets and markets, 2020).

The aforementioned report was launched in June 2020 while OpenAI's GPT-3 model entered beta on 11 June 2020 with public access starting 30 November 2022. Therefore, it is reasonable to consider that the report could not have foreseen the effects of GPT-3 on adaptive learning and therefore possible that the adaptive learning market will exceed the projected values. As such, due to the considerable importance of adaptive learning and other AI-based educational solutions and the underlying economic interest, it is imperative that a proper legislative framework is put in place to ensure continued technological development and prevent or mitigate unwanted side effects.

Methodology

In order to carry out this article, the qualitative research method was used. More specifically, for its application, documents issued by organisations and bodies around the world regarding ethics and sustainability in artificial intelligence, regulations, and proposed solutions for adoption were reviewed and read. Videos and reports elaborated on this topic worldwide were also viewed. The documents analysed are produced by the following organisations: Organisation for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organisation (UNESCO), U.S. Department of Education, World Economic Forum, European Commission, European Parliament, Council of the European Union and The Institute for Ethical AI in Education.

The main question that was asked and that led to the research is the following. *How should artificial intelligence learning systems be developed, implemented, and adopted in a way that still respects ethical norms and maintains a sustainable environment?*

Results and discussions

Based on the research done for this article, the following rules and principles in the ethical sphere are presented and not only that, but also are common to several officially elaborated documents. The principles and rules that are developed in this section lead to an answer to the question mentioned in the methodology.

Therefore, on the basis of the documents, the most important rules concerning ethics and sustainability in artificial intelligence systems have been extracted and detailed.

Artificial intelligence is encountered in applications we use on a daily basis, but there are few situations where the question is asked whether we can take for granted all the answers they give us. In these situations, the role of the user is not to be a filter when it comes to ethical issues, but it is the role of those who plan applications together with the institutions to set the rules and the

ethical and sustainability norms on which the applications should be based in order to achieve the desired results (UNESCO, 2023).

Artificial intelligence systems component for sustainable development

Developing an economy that is sustainable is among the main objectives that world organisations have made. And since artificial intelligence is one of the technologies that is making its mark on development and prosperity, the question of having artificial intelligence systems that promote sustainability automatically arises (Council of the European Union, 2020).

The sustainability theme is brought up by (The Montréal Declaration, 2018), which specifies that the realisation of systems that have artificial intelligence should be done with the environmental sustainability of our planet in mind.

Moving closer to what environmental sustainability could mean if done efficiently, the resources used and polluting emissions should not exceed what the environment can provide. And in this context, the natural resources that have been consumed cannot be replaced. The hardware used in artificial intelligence systems together with the infrastructure must be energy efficient and mitigate greenhouse gas emissions. They should also produce very little electronic waste and should not greatly affect the ecosystem and biodiversity. Of course, all this must be supported by the public and private sector to help minimise pollution and not waste natural resources (The Montréal Declaration, 2018).

The impact of artificial intelligence computers and applications on the environment can be both direct and indirect. More precisely, when we refer to the process through which the components of artificial intelligence computers go through, the effects start from production, then transport, the operations carried out until the end, and we are talking about direct effects. At the same time the indirect impact of artificial intelligence applications can be divided into two categories: a positive impact on the climate and on the areas where the applications are used and a negative impact, for example through carbon leakage (OECD, 2022).

The use of human-centered artificial intelligence

Starting from the fact that artificial intelligence should primarily come as a complement to the activities that humans carry out, especially in the field of education, it has been the basis from which many institutions have started to develop recommendations or regulations (UNESCO, 2023).

This human-centred approach supports the development of skills and helps improve life while respecting and protecting human rights, freedoms, and dignity without differentiating on the basis of colour, gender, nationality, religion, political choices, age, etc. Moreover, AI-based systems should not affect human beings in any way (economically, socially, politically, etc.), but should instead support even the most disadvantaged categories of people with disabilities (UNESCO, 2022).

What the document drafted by the Organisation for Economic Co-operation and Development advocates about placing humans in the spotlight when developing artificial intelligence systems is taken into account by those involved respecting freedom, dignity, personal data, equality, fairness, and labour rights internationally. Elements such as the ability to determine the human or others should also be realised depending on the context (OECD, 2023).

For the field of education, it must be taken into account that people have the main role, but also that this is taken into account in any development of technological applications. In this case,

artificial intelligence in education comes to support students and teachers as a tool through which their skills can be improved (UNESCO, 2021).

Protection of user data and limitations on the purpose for which it is used

To better understand both the importance of data protection and the main objective when it comes to data collection, the main international official documents should be considered.

The category of data to be protected includes both the data that the user of the AI-based system provides and the data that is obtained as the user uses the system. Although some data such as age, gender, political, or religious preferences are not obtained directly, they may be obtained indirectly from user activity, so users should be confident that they will not be used for discriminatory purposes. In addition, it is very important to establish rules and procedures for the use of data and who specifically has access to user data (European Commission, 2019).

For education, the use of artificial intelligence tools must go hand in hand with providing data to students and teachers. Student data should be protected by regulations that ensure that it is used ethically, nondiscriminatory, fair, and transparent manner (UNESCO, 2021).

Detailed explanations plus evidence must be provided to show those using artificial intelligence in education that the data used will benefit students directly or indirectly. Furthermore, it must be made clear to those developing the system that the data provided to the system and its processing are part of the benefit to students. All of this must be done in light of the fact that artificial intelligence systems can be accused of being able to monitor users (The Institute for Ethical AI in Education, 2021).

Transparency as a feature of artificial intelligence-based systems

Specified in the European Commission document, transparency in artificial intelligence systems is represented by both traceability, which is the ability to track processes, and explainability along with communication. More specifically, transparency arises in the context of always needing to know why intelligence systems have made a certain decision (European Commission, Directorate-General for Education, Youth, Sport and Culture, 2022).

Explainability also involves detailing the decisions and actions of a system. Furthermore, the focus should not be on how the system works, but rather on the meaning of the action taken. The same applies to transparency, which should not be seen as a component of the source code, but as a result of its use (European Parliament, European Parliamentary Research Service, Scientific Foresight Unit (STOA), 2020).

Although artificial intelligence systems can be taken into account when providing recommendations, when providing decision support and prediction in education, the ultimate responsibility remains the human factor. In education we can talk about the teacher or the decision makers in a school, etc. In implementing applications to be used in education, the developers must clearly specify which points they have not covered in order to make the system explainable and, furthermore, the system will not compromise the work of those responsible for the educational act through the results provided (The Institute for Ethical AI in Education, 2021).

Fairness and Bias

According to the European Commission, one of the four principles underlying artificial intelligence systems from an ethical point of view is fairness. It states that the way artificial intelligence-based systems are built and deployed must include fairness. In this context, fair systems should not subject people to discrimination or unfair prejudice and should not treat them with contempt in

public. The choices people make should not be a reason to be misled or miscategorised (European Commission, 2019).

The same vision of what intelligence systems in education should accomplish and deliver has been endorsed by the Office of Education Technology. Thus, applications that use artificial intelligence and are integrated into education should put first the main actors in education, i.e. students and teachers followed by studying students and finding their best ways to support learning. Also, the policy of protecting the data of students, teachers. Also important is the ability of the systems to be correct, to make as few errors as possible, and to be able to justify the decisions made. Last but not least, the efficiency of the applications used in the field is mentioned (U.S. Department of Education, Office of Educational Technology, 2023).

Conclusion

Current intelligent learning systems with artificial intelligence that are in the process of development or are in the planning stage can be provided with a number of functionalities to meet the ethical requirements that official bodies worldwide have recommended.

Starting from the initial question from which this article started and having the elements presented in the results section, it can be stated that there is necessary support for those who implement learning applications that use artificial intelligence in order to respect ethical rules and human rights regarding data protection, to be transparent, fair, and to limit bias of any kind. We can say that at the same time there is also the possibility of respecting environmental sustainability.

Creating education applications that use artificial intelligence and empower teachers to change the paradigm of learning while respecting ethics and having as little impact on the environment as possible is not something that will happen in the very distant future. All the recommendations made available together with the regulations adopted by institutions around the world are clear proof that developers of learning systems and others involved in the realm of education are starting a path that has already been initiated.

Taking into account the principles selected and argued in the results area of this article as a starting point for innovative educational systems that change the way students learn and teachers teach. And learning experiences for students are made as inclusive as possible. In addition to this, there will be a greater openness on the part of other stakeholders to adopt as soon as possible secure systems that put students and teachers at the centre, that ensure fairness, data security, transparency, and sustainability.

Apps in education that use artificial intelligence are meant to help students and teachers break down digital barriers, and they will make a difference if they are trusted and if they do not negatively affect the environment in which we carry out our activities.

People in positions of developers or designers of learning platforms should look at these regulations not as imposed barriers or as theoretical elements that will only be discussed. It should be seen from the perspective that in the process of developing safer systems for users they can widen the area from which they can seek guidance. Attention must be paid to the responsible development of systems that contribute to sustainability. In addition, bringing systems into education that put the learner and the teacher first contributes to creating a more stable ecosystem whose results are seen in the long term in the following generations.

This article can be considered as a good start for discussions and debates on how the proposed framework could be implemented in the results area. As limitations of the article that have been identified, it can be mentioned that more concrete examples of systems in which ethical rules have been integrated could have been mentioned. Moreover, a set of recommendations could

have been made for those who want to create a channel of collaboration with the institutions that have issued the principles in order to make a more elaborate planning of the system.

As future perspectives for the extension of the present article, the ethical implications of using artificial intelligence algorithms for decision making in education can be analyzed. The limitations that artificial intelligence may have in explaining how it makes certain decisions can also be explored. Future detailed research will also be in the area of artificial intelligence systems using a reusable energy infrastructure. It is also desired that the following articles present case studies on the carbon footprint of artificial intelligence.

References

- Council of the European Union. (2020, December 11). Draft Council conclusions on Digitalisation for the Benefit of the Environment. Retrieved from <https://data.consilium.europa.eu/doc/document/ST-13957-2020-INIT/en/pdf>
- Duncan, A. (2013, December 17). *Enabling the Future of Learning* | whitehouse.gov. Retrieved February 18, 2024, from Obama White House: <https://obamawhitehouse.archives.gov/blog/2013/12/17/enabling-future-learning>
- European Commission. (2019). ETHICS GUIDELINES FOR TRUSTWORTHY AI High-Level Expert Group on Artificial Intelligence. Retrieved from <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- European Commission, Directorate-General for Education, Youth, Sport and Culture. (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators. *Publications Office of the European Union*. Retrieved from <https://data.europa.eu/doi/10.2766/153756>
- European Parliament, European Parliamentary Research Service, Scientific Foresight Unit (STOA). (2020). The ethics of artificial intelligence : issues and initiatives : study Panel for the Future of Science and Technology. Retrieved from [https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU\(2020\)634452](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)634452)
- Gligorea, I., Cioca, M., Oancea, R., Gorski 4, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Education Sciences*, 13(12).
- Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100001. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2666920X20300011>
- Kaplan, A. (2021). *Higher Education at the Crossroads of Disruption: The University of the 21st Century*. Emerald Publishing Limited. Retrieved from <https://www.emerald.com/insight/publication/doi/10.1108/9781800715011>
- Lepri, B., Oliver, N., & Pentland, A. (2021). Ethical machines: The human-centric use of artificial intelligence. *iScience*, 24(3), 102249. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2589004221002170>
- Markets and markets. (2020, June). *Adaptive Learning Market Growth Drivers & Opportunities*. Retrieved February 18, 2024, from MarketsandMarkets: <https://www.marketsandmarkets.com/Market-Reports/adaptive-learning-market-257528889.html>

- Microsoft. (2019). How AI can enable a Sustainable Future. Retrieved from <https://www.pwc.co.uk/sustainability-climate-change/assets/pdf/how-ai-can-enable-a-sustainable-future.pdf>
- OECD. (2022). Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint. *OECD Digital Economy Papers*, 341.
- OECD. (2023). Recommendation of the Council on Artificial Intelligence. *OECD/LEGAL/0449*. Retrieved from <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- Rane, N., Choudhary, S., & Rane, J. (2023). Education 4.0 and 5.0: Integrating Artificial Intelligence (AI) for Personalized and Adaptive Learning.
- The Institute for Ethical AI in Education. (2021). The Ethical Framework for AI in Education. Retrieved from <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-Annex-Developing-the-Ethical-Framework-for-AI-in-Education-IEAIED-.pdf>
- The Montréal Declaration. (2018). MONTRÉAL DECLARATION FOR A RESPONSIBLE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE. Retrieved from https://monoskop.org/images/d/d2/Montreal_Declaration_for_a_Responsible_Development_of_Artificial_Intelligence_2018.pdf
- U.S. Department of Education, Office of Educational Technology. (2023). Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations. Retrieved from <https://tech.ed.gov/ai-future-of-teaching-and-learning/>
- UNESCO. (2021). AI and education: guidance for policy-makers. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- UNESCO. (2022). Recommendation on the Ethics of Artificial Intelligence. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- UNESCO. (2023, February 06). *Ethics of AI: Challenges and Governance [Video]* . Retrieved from YouTube: <https://www.youtube.com/watch?v=VqFqWIqOB1g>
- UNESCO. (2023). *Guidance for generative AI in education and research*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000386693/PDF/386693eng.pdf.multi>
- Verhagen, J., Dalibert, L., Lucivero, F., Timan, T., & Co, L. (2016). Designing values in an adaptive learning platform. *2nd workshop on Ethics & Privacy in Learning Analytic during the 6th International Learning Analytics & Knowledge Conference*. Retrieved from <https://shs.hal.science/halshs-01650224v1/document>
- World Economic Forum. (2024). The Global Risk Report 2024. Retrieved from https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf