A Critical Analysis of the European Union's Considerations on the Ethical Use of Artificial Intelligence in Education

Fabio Morandín Ahuerma

Meritorious Autonomous University of Puebla (BUAP)
fabio.morandin@correo.buap.mx
ORCID 0000-0001-6082-2207

The First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
Isaac Asimov (1920-1992).

▼ p. 116

INTRODUCTION

This document presents ethical guidelines for the use of artificial intelligence (AI) and data in education, developed by an Interdisciplinary Committee within the framework of the European Commission's Digital Education Action Plan 2021-2027, as outlined in the document: "Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators" (European Union, 2022). The guidelines clarify common misconceptions about AI and describe its potential benefits in optimizing educational processes for both students and teachers. However, the main focus is on establishing ethical requirements and considerations for the implementation of AI in education. Furthermore, the document provides six guidelines for educational agents and institutions, emphasizing the need for adaptive and personalized learning, the use of progress indicators to guide student learning, special attention to students with unique needs, responsible use of AI in assessment, and the importance of clearly defining the evolving role of human teachers.

▼ p. 117

In summary, the guidelines underscore the imperative of harnessing AI's transformative potential in education while proactively addressing ethical risks related to bias, transparency, privacy, and the preservation of human agency. Finally, it is noted that, although the European Union's ethical guidelines for the use of AI in education represent a commendable effort to proactively address complex issues in this area, digital and media literacy gaps in Latin America pose significant obstacles to responsibly implementing AI in education. Disparities in access to technology and the quality of education can reinforce existing inequalities. Despite these challenges, with collaborative efforts, a more inclusive and ethical application of AI in education can be achieved through the European initiative, prioritizing approaches that empower teachers and student communities.

THE DOCUMENT

Presented in October 2022 by Bulgarian politician Mariya Gabriel, then European

Commissioner for Innovation, Research, Culture, Education, and Youth, the document *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators* (European Union, 2022) is part of a broader European Commission initiative known as the Digital Education Action Plan 2021-2027 (European Commission, 2021). This initiative outlines the European Union's (EU) strategic perspective on managing digital transformation in a sustainable manner within the education and training sectors. Among the main initiatives of the Action Plan, the Innovation Commission suggests several measures aimed at fostering the creation of a cutting-edge digital educational environment. One of these measures focuses specifically on addressing the ethical challenges and considerations surrounding the implementation of AI and data management in the education and training sector. The two strategic priorities of the Action Plan are essentially the development of digital infrastructure and the enhancement of individuals' digital capacities.

▼ p. 118

An interdisciplinary committee composed of representatives from the EU, UNICEF (United Nations Children's Fund), UNESCO (United Nations Educational, Scientific, and Cultural Organization), and the OECD (Organisation for Economic Co-operation and Development), along with 24 advisors and the international research and consulting firm Ecorys, developed a 40-page illustrated document with five essential goals: a) clarifying common misconceptions about artificial intelligence (AI); b) providing practical examples of how AI and data analytics can optimize the educational process, benefiting both students and teachers, while supporting the infrastructure of the educational system; c) presenting considerations and establishing the ethical requirements necessary for AI implementation in the educational field; d) offering guidelines aimed at educators and school administrators; and e) presenting a glossary defining key terms related to AI and data management, thereby facilitating a better understanding and application of these technologies in education.

Among the participating experts are Agata Majchrowska, Aleksander Tarkowski, Ari Alamäki, Deirdre Butler, Duuk Baten, Egon Van den Broek, Guido Noto La Diega, Hanni Muukkonen van der Meer, Inge Molenaar, Jill-Jênn Vie, Josiah Kaplan, Juan Pablo Giraldo Ospino, Julian Estevez, Keith Quille, Lidija Kralj, Lucilla Crosta, Maksim Karliuk, Maria Wirzberger, Matthew Montebello, Stephan Vincent-Lancrin, Tapani Saarinen, Tobias Rohl, Viola Schiaffonati, Vitor Hugo Mendes da Costa Carvalho, and Vladislav Slavov (European Union, 2022).

However, it is necessary to conduct a critical analysis of two fundamental sections of the *Guidelines*. These sections are "Ethical Considerations and Requirements Accompanying the Ethical Guidelines," which consists of seven ethical considerations or axiological maxims, and "Guidelines for Educational Agents and School Administrators," which includes six guidelines and six areas of application.

▼ p. 119

An axiological maxim is a fundamental principle or value that guides the behavior and decisions of an individual or society in relation to what is considered valuable or desirable. The word "axiology" comes from the Greek *axios*, meaning "value," and refers to the field of philosophy that studies ethical beliefs and moral value judgments. Axiological maxims are concise, general statements that express the ethical, moral, or value-based principles that an individual or community considers essential in their belief system, such as the value of justice, freedom, respect, equality, honesty, solidarity, or any other maxim deemed important (Morandín-Ahuerma *et al.*, 2023b).

AXIOLOGICAL AXIMS PROPOSED BY THE EU FOR AI IN EDUCATION

Capacity for human action and vigilance

> Human agency relates to an individual's capability to become a competent member of society. A person with agency can determine their life choices and be responsible for their actions. Agency underpins widely used concepts such as autonomy, self-determination, and responsibility. (EU, 2022, p. 18)

The document *Ethical Guidelines*... points out that the concept of human agency is the intrinsic ability of an individual to effectively participate in society.

▼ p. 120

This concept implies that a person can not only identify and choose between different life options but also assume responsibility for their choices and actions. Agency is fundamental, as it underpins and supports key notions in the ethical and social spheres, such as autonomy, which refers to a person's ability to make independent decisions; self-determination, which emphasizes individuals' right to control their own destiny; and responsibility, which involves recognizing and accepting the consequences of one's actions. In essence, this concept highlights the importance of individual agency within the framework of social participation and ethical decision-making (EU, 2022, p. 18).

This is a crucial element in determining the human agency of those responsible for applying AI in education. Four key considerations support the ethical use of AI and data in teaching: a) learning and assessment as justified choices, seen as a rational decision-making process; b) human agency, also referred to as agency (Floridi, 2023); c) fairness, understood as equal opportunities; and d) humanity, which, although a potentially problematic term depending on the approach, is assumed to mean beneficence and non-maleficence.

Thus, the EU document assumes that those implementing AI application strategies in education will do so under the aforementioned concepts, particularly beneficence and non-maleficence.

▼ p. 121

Transparency

> Transparency includes traceability, explainability and communication. (EU, 2022, p. 18)

Regarding AI systems, transparency refers to the ability to understand and track how algorithms make decisions or reach specific conclusions. This characteristic is essential for fostering trust in AI technology, allowing users, developers, and regulators to comprehend the internal processes guiding automated actions. However, this is not always the case. Not necessarily due to intentionality, but rather because systems, when processing large amounts of data, can reach "their own" conclusions without human intervention.

Traceability. Transparency in AI relies on several pillars, including traceability, explainability, and communication. Although the EU document does not explicitly detail these concepts, traceability can be understood as the ability to document, track, and record the decision-making process of an AI system. This implies that all steps—from input data, the algorithms used, to the final decision—can be identified and reviewed. This allows for the interpretation of the sequence of events and decisions within the system, enabling the evaluation of AI processes as a whole.

Explainability. Explainability relates to the ability to interpret AI processes and decisions in a comprehensible manner. An AI system is explainable if its operations, decisions, and the logic behind them can be understood by humans. Explainability is necessary to assess the accuracy, effectiveness, and even fairness of AI systems, allowing users to understand why a specific decision was made and how it was reached.

▼ p. 122

Transparency. Transparency refers to how information about AI processes and decisions is communicated to users and developers. Effective communication is essential to ensure that information is accessible and understandable to those without specialized technical knowledge. This helps to dispel doubts and concerns about AI by providing clear data on its operation, limitations, and potential biases.

Inclusion

> Diversity, non-discrimination, and fairness including accessibility, universal design, the avoidance of unfair bias, and stakeholder participation, which allows use regardless of age, gender, abilities, or characteristics - with a particular focus for students with special needs. (EU, 2022, p. 18)

This maxim highlights the incorporation of diversity, non-discrimination, and equity in the development and use of technologies that have a profound impact on individuals, communities, and society as a whole. It translates into greater inclusion and accessibility, facilitating the active participation of a wider range of people in the digital community. Additionally, it reduces social and digital divides, thereby diminishing

inequalities in access to and use of technologies. In the educational sphere, it calls for all students to have access to high-quality learning opportunities, regardless of their individual characteristics (Floridi *et al.*, 2018).

▼ p. 123

Privacy and data governance

> Privacy and data governance including respect for privacy, quality and integrity of data, and access to data. (EU, 2022, p. 18)

Privacy and data governance are not merely technical concepts; they become fundamental pillars in building an ethical and responsible educational future. They encompass a set of principles and practices aimed at ensuring the proper handling of personal information, while also protecting the privacy of students, educators, and other stakeholders within the educational ecosystem.

AI, with its capacity to collect large amounts of sensitive data—such as learning progress, interactions, and assessment results—demands a strong commitment to the protection of personal information. It is about ensuring that the access, use, and disclosure of these data are carried out responsibly and with the individuals' consent, safeguarding their confidentiality (Véliz, 2020).

The accuracy, reliability, and consistency of the data used and generated by AI systems are crucial for ensuring that educational decisions based on these data are valid and beneficial. Data quality allows AI analyses to be trustworthy, while integrity ensures that the information has not been altered or corrupted, maintaining its accuracy over time.

▼ p. 124

Robustness and technical safety

> Technical robustness and security, particularly resilience against attacks, protection and overall safety, accuracy, reliability, and reproducibility. (EU, 2022, p. 19)

Artificial intelligence applied to education (AI) holds immense potential to transform teaching and learning. For this potential to be fully realized, the technology must be robust, secure, and trustworthy. Trust in AI means ensuring the integrity of educational processes, protecting students' personal data, and building a digital ecosystem where everyone feels safe and respected.

The security of AI systems is essential to generating trust among students, educators, and administrators. This security must protect personal and educational data from unauthorized access and ensure the technology's resilience to attacks and vulnerabilities. This means that the system must have the capability to restore and protect itself in the face of any attack. The ability of AI systems to resist and recover from cyberattacks and technical failures is referred to as "overall system resilience." This ensures the continuity and stability of processes even under adverse circumstances.

For example, on July 19, 2024, a global failure occurred in the Windows operating system, affecting Microsoft 365 applications and services, leading to widespread disruptions across several sectors, including airlines, banks, hospitals, and media outlets. The root cause of this incident was identified as a defect in a software update from CrowdStrike, a cybersecurity company whose technology is integrated into many systems dependent on Microsoft services (Satariano *et al.*, 2024).

▼ p. 125

This is an example of digital dependency and, ultimately, the vulnerability present in any digital system, which, in this case, caused the shutdown of several strategic systems worldwide for a few hours. This incident serves as a stark reminder that despite technological advancements, digital infrastructure remains susceptible to failures that can have far-reaching consequences. The interconnection of systems and reliance on third-party components, as in the case of CrowdStrike's software, can amplify the impact of a single error, causing a cascade of disruptions across multiple sectors.

Moreover, this event highlights the importance of having robust contingency plans and backup systems to mitigate the effects of such failures. The ability to quickly recover from a cybersecurity incident or technical failure is essential to ensuring operational continuity and minimizing economic and social losses.

In the context of AI in education, the accuracy, reliability, and reproducibility of AI systems are crucial for ensuring that AI-based educational interventions and assessments are trustworthy. Adhering to high standards of security and technical robustness is indispensable for complying

with data protection and privacy regulations, such as the General Data Protection Regulation (GDPR), a European Union law regulating personal data protection, which came into effect on May 25, 2018 (EU GDPR, 2016; Zekos, 2022).

This law is considered the most robust privacy and security regulation in existence and aims to protect the fundamental rights and freedoms of individuals in the digital age. The law includes: explicit consent for the use of personal data; the right to access one's own data; rectification, erasure, and the "right to be forgotten"; the right to object to the use of data for specific purposes, such as profiling; and the right to request a copy of one's data.

▼ p. 126

For their part, companies operating within European territory are required to appoint a data protection officer in certain cases, notify relevant authorities of data breaches, implement appropriate security measures to protect data, and provide transparent and easily accessible information to individuals about the processing of their data. To enforce the law, each EU member state must establish an independent supervisory authority to oversee GDPR compliance. In the event of violations, data controllers may face fines of up to 20 million euros or 4% of their global annual turnover for breaching one or more data protection rules (EU GDPR, 2016).

Accountability

> Accountability includes auditability, the minimization of negative effects and their reporting, as well as commitments and compensations (EU, 2022, p. 19).

Trust is strengthened when AI systems are auditable, allowing third parties to examine their operation and ensure they comply with ethical, regulatory, and technical principles. This transparency fosters trust among users by demonstrating a commitment to integrity and responsibility.

Therefore, preventing and managing risks serves to avoid harm and act effectively when issues arise. The proactive identification, minimization, and reporting of the negative effects of AI begin by detecting potential problems during development and implementing mechanisms for rapid responses to mitigate any adverse impacts during the process.

▼ p. 127

Data theft is also a common concern, and both children and young people can be easily targeted, whether in domestic or school environments. This highlights the need for robust security systems that protect their personal and academic information, as well as clear, up-to-date protocols to respond to security incidents. Moreover, it is essential to educate children and young people about the risks of the digital world, teaching them how to protect their personal information, identify potential threats, and navigate the internet safely. Collaboration between parents, educators, and institutions is key to creating a secure digital environment and protecting minors from the dangers of data theft (UK, 2023).

In 2022, 1,981 schools in the United States, spanning 45 districts, fell victim to cyberattacks. These attacks allowed criminals to access sensitive information such as grades, health records, attendance, and special education data, with the intent of extortion or identity theft. Besides the financial costs and loss of productivity, the emotional and socioeconomic impact on students was and remains significant (Curry, 2023).

Five measures to keep children and young people safe from data theft include:

a) Privacy Settings: Teach children and young people how to properly configure privacy settings on social media, apps, and devices, limiting the information they share publicly and controlling who can access their profiles.

b) Secure Passwords: Explain the importance of creating strong, unique passwords for each account, avoiding obvious personal information, and using combinations of letters, numbers, and symbols.

c) Security Software: Install and keep security software, such as antivirus, antimalware, and firewalls, updated on devices to protect against online threats.

d) Phishing and Scam Education: Teach children and young people to recognize suspicious emails, messages, and websites attempting to steal personal information, and to avoid clicking on links or downloading files from unknown sources.

▼ p. 128

e) Supervision and Communication: Maintain open communication with children and young people about their online activity, set time limits for

device use, and be vigilant for any unusual behavior or changes in their online habits (UK, 2023).

While privacy is a right for every individual, parents must stay vigilant about what their children consume online and who may be contacting them. Recreational activities, such as online gaming, can lead to interactions with unknown individuals behind avatars, potentially exposing minors to risks like cyberbullying, grooming (the process by which an adult builds a relationship of trust with a minor online for malicious purposes), or exposure to age-inappropriate content. For this reason, parents should get involved in their children's online activity, set clear boundaries, and foster open communication about the potential dangers in the digital world (Faraz *et al.*, 2022).

Schools, on the other hand, must implement security measures in their systems and networks to protect the information of students and staff. They should also educate students on the risks and best practices for using the internet and information technologies. Furthermore, schools must establish clear protocols for responding to security incidents and work in collaboration with parents to ensure the safety of minors in the digital environment.

▼ p. 129

GUIDELINES

Adaptability

> Use learning technologies that adapt to the capacity of each learner (EU, 2022, p. 22).

Artificial intelligence (AI) technology holds the potential to transform teaching and learning into a more personalized, efficient, and student-centered experience. A prime example of this is intelligent tutoring systems, which tailor education to the individual needs and levels of each student.

AI systems use intelligent algorithms to analyze each student's progress, identify their strengths and weaknesses, and provide personalized materials and activities that match their learning pace and style. AI can help students receive real-time feedback on their work, enabling them to identify areas for improvement and correct errors more effectively. While teacher support remains indispensable, AI tools can significantly streamline this process, especially for teachers managing large groups (Shah, 2023).

Intelligent tutoring systems empower teachers by providing tools to monitor each student's progress individually. This allows teachers to dedicate more time and attention to those students who need it most and offer more effective teaching (Okado *et al.*, 2023).

▼ p. 130

AI not only benefits students but also teachers, enabling them to work more efficiently and deliver higher-quality education. Additionally, it can help reduce the educational gap by offering all students equal learning opportunities, regardless of their background or socioeconomic status.

However, the EU document emphasizes the need to critically evaluate the effectiveness, reliability, and fairness of AI systems applied in educational contexts. It highlights the importance of ensuring that these systems are designed to meet the specific learning objectives of each student, ensuring that the predictions, assessments, and classifications generated are accurate and relevant for their educational purpose. A thorough analysis of how AI interprets and responds to individual learning needs, with a focus on absolute transparency in AI processes and outcomes, must rest particularly in the hands of developers.

Indicators

> Use student indicator dashboards to guide them through their learning process (EU, 2022, p. 23).

There are various types of assessment, each serving a specific purpose. For example, summative assessment is conducted at the end of a learning period to measure a student's achievement and assign a grade. This type of assessment provides an overview of the mastery of content and skills acquired, and can be achieved through midterm exams, a final essay, or presentations (Cope & Kalantzis, 2019).

▼ p. 131

In addition to summative assessment, there are other forms of evaluation equally useful to educators, starting with diagnostic assessment, which should be conducted at the beginning of a course or learning unit to identify students' prior knowledge and detect potential difficulties and areas to address.

Formative assessment, on the other hand, takes place during the learning process and aims to

provide feedback to both the student and the teacher regarding progress and areas that need reinforcement. It is carried out through work samples, observation guides, checklists, rubrics, scales, etcetera. Formative assessment is divided into three phases:

Initial assessment: Conducted at the start of a cycle or unit, where the teacher gathers information about the student's competencies, knowledge, and challenges.

Ongoing assessment: Performed throughout the educational process to monitor student progress and adjust teaching strategies as needed.

Final assessment: Conducted at the end of the educational process to evaluate the achievement of objectives and reflect on learning (SEC, 2024).

Self-assessment is the process in which students reflect on their own performance, analyze their strengths and weaknesses, and identify areas for improvement. This type of assessment fosters metacognition and autonomy in learning, as students become more aware of their learning process and take responsibility for their own progress. It is relevant to conduct this at the end of the course so that, in addition to knowing the student's performance in the classroom, the teacher is also aware of the student's own expectations regarding their grade. Interestingly, contrary to what one might expect, many students underestimate their efforts and self-grade below their actual performance. Others, albeit a small percentage (less than 10%), may give themselves a higher grade than they deserve (Andrade, 2019).

▼ p. 132

Peer assessment involves students evaluating the work of their peers using pre-established criteria and rubrics. This type of assessment promotes collaboration, critical thinking, and the ability to give and receive constructive feedback. It also allows students to learn from their peers and develop evaluation skills. While objectivity can be a challenge if students know whose work they are assessing, whether due to friendship or antagonism, this risk can be mitigated by asking students to use their student number, for example, rather than their name (Dimitriadou & Lanitis, 2023).

In any type of assessment, whether summative, formative, diagnostic, or otherwise, indicators are required to measure and assess student learning or performance. Indicators are observable and measurable elements that allow educators to determine whether learning objectives have been met or expected skills have been developed. These indicators can be qualitative or quantitative, and their choice will depend on the type of assessment and the specific objectives to be evaluated. Without clear and well-defined indicators, assessment would lack objectivity, making it difficult to gauge student progress and make informed decisions about their learning (Cowley *et al.*, 2022).

Some schools, particularly private ones, have incorporated personalized online indicator dashboards into their systems. This tool provides students and parents with direct and clear feedback on the learning and assessment methods employed. By emphasizing the "how" rather than the "what" of learning, the goal is for students to develop crucial skills for success, such as critical thinking, adaptability, and self-awareness, beyond the academic realm. These platforms, which go beyond the traditional transcript (Kardex), allow for the evaluation of competencies acquired. This approach was already considered in Mexico during the evaluation phase of the common curricular framework (CONAEDU, 2022).

▼ p. 133

Special Needs Attention

> Provide personalized interventions to address special needs (EU, 2022, p. 23).

AI technology can enhance education by personalizing learning experiences and increasing accessibility for students with unique educational requirements. The use of artificial intelligence systems can analyze student data, allowing schools to quickly identify children and young people who may require personalized assistance. This proactive intervention enables the provision of tailored educational experiences to meet each student's individual needs.

It is important to emphasize the relevance of establishing procedures that allow teachers to monitor and proactively address situations requiring empathetic treatment of students and their guardians, underscoring the irreplaceable value of human intervention in the educational process. This highlights the need to preserve human agency and oversight within the educational environment when using AI tools to ensure that the emotional and psychological needs of students are adequately addressed (Chen *et al.*, 2022).

The EU document underscores the importance of inclusivity and accessibility in the design of

intelligent systems, particularly in their ability to interact effectively with students who have disabilities or special educational needs. For example, AI can provide tools such as screen readers, voice recognition, automatic subtitles, and sign language translation, enabling students with visual, auditory, or speech impairments to access and participate in the learning process.

▼ p. 134

JAWS (Job Access With Speech) is a screen reader designed for computer users whose vision loss prevents them from seeing screen content or navigating with a mouse. JAWS provides voice output and a Braille system for the most commonly used computer applications. NVDA (NonVisual Desktop Access) is another solution, with the advantage of being free software that offers access to blind individuals (NVDA, 2024).

AI can also help identify students' learning needs early and continuously track their progress, allowing for quicker and more effective interventions when detecting any anomalies through ad hoc tests that AI can develop depending on the level, age, and context, among other factors.

Essay Grading

> Grading essays using automated tools (EU, 2022, p. 24).

There is a significant debate around whether essay-based evaluations are coming to an end due to the use of generative artificial intelligence, as tools like ChatGPT and other large language models (LLMs) can produce coherent and seemingly original texts in response to prompts and rubrics. This raises concerns about academic integrity and the effectiveness of traditional essay-based assessments in measuring students' understanding and critical thinking skills. Some argue that AI-generated essays lack the depth and nuance of human-written work, while others view this as a challenge to adapt assessment methods to this new reality (Fütterer *et al.*, 2023).

▼ p. 135

The potential transformation of the educational evaluation process through the adoption of AI technology, using an automated grading system based on natural language models, aims not only to streamline and improve the accuracy of written task assessments but also to enrich the pedagogical feedback provided to students. The system's ability to analyze texts in detail, identify errors, and assign grades automatically presents an opportunity to reduce the workload of teachers, allowing them to focus on more qualitative aspects of written work (Herbold *et al.*, 2023).

Furthermore, the use of artificial neural networks trained on historical cases points towards increasingly refined evaluations, adapted to common grammatical error patterns among students, promising not only precision but also personalized learning. Additionally, the inclusion of plagiarism detection functionality reinforces academic integrity, ensuring that the evaluated work meets required standards.

School Services Management

> Management of student enrollment and resource planning (EU, 2022, p. 25).

Implementing AI in schools offers great potential to assist administrators with their workload. There never seem to be enough hours to keep up with paperwork, scheduling, record-keeping, and planning needed to manage a school. But with AI, given the amount of data it can analyze, it can help ease this burden. Above all, the goal is to give time back to administrators and teachers alike. Technology is just a tool, but if used correctly, it can help schools.

▼ p. 136

Additionally, schools can use historical academic data to set personalized goals for student performance and anticipate their success in various subjects. This approach reflects an innovative and proactive use of AI to improve school planning and student learning, signaling a shift towards more data-driven educational practices (Popenici, 2022).

AI can also be a valuable tool in helping at-risk students by identifying patterns and warning signs in their performance and behavior. Predictive analysis can be implemented, as AI systems can anticipate which students may face academic difficulties and offer personalized interventions. For example, AI can recommend additional learning resources, adjust the teaching pace, or suggest specific tutoring sessions. Additionally, AI allows teachers to monitor student progress in real-time, facilitating a quick and effective response, which can be crucial in improving academic outcomes and preventing failure (Okado *et al.*, 2023).

Moreover, specific responsibilities should be assigned to school staff and tutors for the monitoring and analysis of AI-generated results. This monitoring aims not only to ensure the system's proper functioning but also to use the collected data to continuously improve teaching, learning, and assessment processes. This approach involves diligent supervision and constant adaptation based on results, ensuring that AI systems effectively and ethically contribute to educational monitoring from a school administration perspective (Babu & Ontario, 2023).

▼ p. 137

Use of Chatbots

> Use chatbots to guide students and their parents through administrative tasks (EU, 2022, p. 25).

The EU proposes the use of chatbots or virtual assistants to facilitate and improve the educational experience for both students and their families. Technology can be a powerful tool to ensure that all students, regardless of their individual challenges, have access to the same opportunities and support within the educational system.

Using chatbots to guide students can be an effective tool for providing academic follow-up to those who need it most, as they can offer personalized support and feedback tailored to individual learning needs. These AI-powered tutors can engage students in interactive dialogues, answer questions, provide explanations, and offer practice exercises, making the learning process more engaging and effective (Bruno *et al.*, 2023).

Additionally, chatbots can be available 24/7, providing students with immediate support whenever they need it, regardless of location or time. This can be particularly beneficial for students who struggle in traditional classroom environments or need extra help outside school hours.

However, it is essential to ensure that these chatbots are carefully designed and implemented. They must be transparent about their AI nature, prioritize student privacy, and be integrated into a comprehensive educational approach that includes human interaction and support (Lena, 2021; Von Braun *et al.*, 2021).

▼ p. 138

Although some experts believe that chatbots could become common tools in schools, a recent study reveals significant issues with the current generation of social robots. The research (Guggemos *et al.*, 2020) conducted in a U.S. primary and secondary school with a humanoid robot called "Pepper" found that these robots are not yet ready for widespread use in classrooms. For example, they struggle to function in noisy environments, cannot follow complex conversations or interpret subtle facial expressions, and are unable to recognize individuals without identification devices.

While intelligent systems like chatbots can simulate social interactions, their inability to genuinely feel or empathize with users may limit the depth and quality of these interactions. It is important to acknowledge and communicate these limitations to avoid misunderstandings about AI's capacity to replace authentic human interactions (Dettmers *et al.*, 2023). These issues raise important considerations for the successful integration of social robots in education, as seen in schools in Kumamoto, Japan, where robotic teachers are already interacting with students, with hopes for greater success (McCurry, 2023).

This approach reflects a forward-looking vision in which the integration of technological solutions in education extends beyond mere automation, aiming for a meaningful transformation in how teaching and learning processes are managed and supported, making education more accessible, efficient, and personalized.

It is crucial to emphasize that AI systems must explicitly communicate to users that their social interaction is simulated, clarifying that they lack the ability to feel or empathize. This point highlights the importance of maintaining a clear distinction between authentic human interactions and computer-generated simulations, acknowledging the inherent limitations of AI in terms of emotional understanding and empathy, which are essential for social and environmental well-being (Bearman *et al.*, 2023).

▼ p. 139

Additionally, there is a need to implement effective strategies for monitoring and verifying that AI systems meet the goals, purposes, and applications for which they were designed, thus ensuring their technical robustness and security. This involves establishing regular review and evaluation procedures to confirm that the system is functioning properly according to the established

expectations and requirements (Educación3.0, 2024).

Discussion and Analysis

The *Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators* within the framework of the European Commission's Digital Education Action Plan 2021-2027 is a document intended to serve as a guide for classroom teachers, school administrators, policymakers, students, and parents to understand both the potential benefits and the ethical challenges in implementing AI systems in education.

On the positive side, the adaptability and personalization enabled by AI-driven technologies stand out for their potential to revolutionize the learning experience and better meet the diverse needs of students. The ability to provide personalized interventions and assessments promises to improve learning outcomes and accessibility. Moreover, the use of AI for administrative tasks such as managing enrollments and resource planning could increase efficiency and equity in school operations.

However, the document rightly emphasizes the critical importance of addressing key ethical considerations regarding the use of AI in schools. It underscores the need for diversity, non-discrimination, and equity in the development and deployment of AI systems.

▼ p. 140

AI models can easily perpetuate and amplify the social biases present in their training data. Rigorous testing and audits are necessary to detect biases and ensure that these technologies do not inadvertently exclude certain student populations or widen divisions between different demographic groups. For instance, in Mexico, where disparities between the northern and southern regions of the country have been a persistent issue, it is essential to avoid deepening this gap through the adoption or lack of adoption of technologies that may carry inherent biases.

For this reason, the guidelines stress the importance of transparency regarding how AI systems make decisions, as well as clear measures of accountability. The “black box” nature of many AI algorithms makes it difficult for educators and administrators to fully understand and oversee their internal workings. It is crucial to establish robust processes for explainability, auditing, and assigning responsibilities.

The immense data-gathering capabilities of AI-driven educational technologies raise significant privacy concerns. Strict data protection policies, clear usage guidelines, and parental and student consent procedures must be implemented. Excessive reliance on these systems also risks eroding the personal connection between students and teachers.

While AI can augment and enhance educational practices, the document appropriately emphasizes the need to preserve human agency, empathy, and oversight. Over-automation of basic teaching and assessment functions could undermine the irreplaceable role of human teachers and their pedagogical expertise.

In conclusion, the ethical guidelines presented provide a solid foundation for ensuring the responsible and beneficial use of AI in education. However, the true challenge lies in translating these principles into consistent and applicable practices across various institutions and educational contexts. Ongoing research, stakeholder collaboration, and a willingness to adjust course as issues arise will be crucial to navigating this technological transformation while preserving the core mission and values of education.

▼ p. 141

Findings

The European Union's ethical guidelines for the use of AI in education represent a commendable effort to proactively address the complex issues that arise at the intersection of transformative technology and the noble mission of teaching and learning. The emphasis on fundamental principles such as transparency, accountability, equity, and human agency provides a solid framework for navigating this transition.

As Latin American observers, we are acutely aware that digital and media literacy gaps in our region pose significant challenges to realizing this vision of responsible AI integration in education. The widespread disparities in access to technology, discussed in chapter two, along with digital skills and the quality of education, mean that the benefits of these advanced tools may disproportionately reach already privileged populations, further entrenching systemic inequalities.

Nevertheless, we can remain cautiously optimistic. Through concerted and collaborative

efforts among policymakers, educators, technologists, and civil society, we believe it is possible to chart a path toward the inclusive and ethical application of AI in the service of all students. Localized and contextualized approaches that empower teachers and communities can ensure that these technologies truly serve the most marginalized, rather than leaving them behind.

Without a doubt, the road ahead will be long and arduous, requiring sustained commitment to closing digital divides, strengthening media literacy, and cultivating the human skills that AI will never replace. But if we remain true to the spirit of these ethical guidelines—valuing transparency, accountability, and the primacy of human action—we can trust that the transformative potential of AI can be harnessed to create a more equitable and robust educational landscape, ultimately translating into quality education for all. It is an ambitious yet worthy aspiration, one we must pursue with rigor, creativity, and an unwavering moral compass.

REFERENCES

Andrade, H. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education, 4*, 87. https://doi.org/10.3389/feduc.2019.00087

Babu, G. y Ontario, S. W. (2023). Managing the Strategic Transformation of Higher Education through Artificial Intelligence. Administrative Sciences. https://doi.org/10.3390/admsci13090196

Bearman, M., Ryan, J. y Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: a critical literature review. *Higher Education*, *86*(2), 369-385. https://doi.org/10.1007/s10734-022-00937-2

Bruno, B., Amirova, A., Sandygulova, A., Lugrin, B. y Johal, W. (2023). Culture in Social Robots for Education. In: Dunstan, B.J., Koh, J.T.K.V., Turnbull Tillman, D., Brown, S.A. (eds). *Cultural Robotics: Social Robots and Their Emergent Cultural Ecologies. Springer Series on Cultural Computing*. Springer. https://doi.org/10.1007/978-3-031-28138-9_9

Chen, X., Zou, D., Xie, H., Cheng, G. y Liu, C. (2022). Two decades of artificial intelligence in education. *Educational Technology & Society*, *25*(1), 28-47. https://scholars.ln.edu.hk/ws/portalfiles/portal/41220986/25_1_03.pdf

CONAEDU. (2022). *Marco Curricular Común, EMS 2022* (PDF). DGEMS. https://educacionmediasuperior.sep.gob.mx/work/models/sems/Resource/13516/1/images/MarcoCurricularComunEMS2022.pdf

Cope, B. y Kalantzis, M. (2019). Education 2.0: Artificial intelligence and the end of the test. *Beijing International Review of Education, 1*, 528–543. https://doi.org/10.1163/25902539-00102009

Cowley, H. P., Natter, M., Gray-Roncal, K., Rhodes, R. E., Johnson, E. C., Drenkow, N., . . . Gray-Roncal, W. (2022). A framework for rigorous evaluation of human performance in human and machine learning comparison studies. *Scientific Reports*, *12*(1), 5444. https://doi.org/10.1038/s41598-022-08078-3

Curry, R. (2023). *Hackers see wealth of information to steal in children's school records* [Web]. CNBC. https://www.cnbc.com/2023/12/27/hackers-see-wealth-of-information-to-steal-in-kids-school-records.html

Dettmers, T., Pagnoni, A., Holtzman, A. y Zettlemoyer, L. (2023). Qlora: Efficient finetuning of quantized LLMs. *arXiv*. https://doi.org/10.48550/arXiv.2305.14314

Dimitriadou, E. & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments*, 10(12). https://doi.org/10.1186/s40561-023-00231-3

Educación3.0. (2024). *La importancia de garantizar la seguridad digital en las aulas* [Web]. Educación 3.0. https://www.educaciontrespuntocero.com/empresas/seguridad-digital-en-las-aulas/

Faraz, A., Mounsef, J., Raza, A. y Willis, S. (2022). Child safety and protection in the online gaming ecosystem. *IEEE Access, 10*, 115895-115913. https://www.doi.org/10.1109/ACCESS.2022.3218415

Floridi, L. (2023). AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models. *Philosophy & Technology*, *36*(1), 15. https://doi.org/10.1007/s13347-023-00621-y

Floridi, L., Cowls, J., Beltrametti, M. *et al.* AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds & Machines* 28, 689–707 (2018). https://doi.org/10.1007/s11023-018-9482-5

Fütterer, T., Fischer, C., Alekseeva, A., Chen, X., Tate, T., Warschauer, M., & Gerjets, P. (2023). ChatGPT in education: Global reactions to AI innovations. *Scientific Reports*, 13, Article 15310. https://doi.org/10.1038/s41598-023-42227-6

Guggemos, J., Seufert, S. y Sonderegger, S. (2020). Humanoid robots in higher education: Evaluating the acceptance of Pepper in the context of an academic writing course using the UTAUT. *British Journal of Educational Technology*, *51*(5), 1864-1883. https://doi.org/10.1111/bjet.13006

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z. y Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific reports*, *13*(1), 18617. https://doi.org/10.1038/s41598-023-45644-9

Léna, P. (2021). Robotics in the Classroom: Hopes or Threats? In J. von Braun, M. S. Archer, G. M. Reichberg, & M. Sánchez Sorondo (Eds.), *Robotics, AI, and Humanity: Science, Ethics, and Policy* (pp. 109-117). Springer International Publishing. https://doi.org/10.1007/978-3-030-54173-6_9

Morandín-Ahuerma, F., Romero-Fernández, A., Villanueva-Méndez, L., y Santos-Cabañas, E. (2023). Hacia una fundamentación ético-normativa del sujeto de derecho. *Revista Jurídica Crítica y Derecho*, *4*(6), 1-12. https://doi.org/10.29166/cyd.v4i6.4242

McCurry, J. (2023). *Japanese city to use robots to tackle rise in truancy* (Web). The Guardian. https://www.theguardian.com/world/2023/sep/06/japanese-city-robots-tackle-rise-in-school-truancy

NVDA. (2024). *NVDA en español* [Web]. https://nvda.es/

Okado, Y., Nye, B. D., Aguirre, A., & Swartout, W. (2023). Can Virtual Agents Scale Up Mentoring?: Insights from College Students Experiences Using the CareerFair.ai Platform at an American Hispanic-Serving Institution. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova, *Artificial Intelligence in Education*, Springer Cham.

Popenici, S. (2022). *Artificial intelligence and learning futures: Critical narratives of technology and imagination in higher education*. Routledge. https://t.ly/6MRkk

Satariano, A., Mozur, P., Conger, K. y Frenkel, S. (2024, 19 de julio). *Chaos and Confusion: Tech Outage Causes Disruptions Worldwide* [Web]. New York Times. https://www.nytimes.com/2024/07/19/business/microsoft-outage-cause-azure-crowdstrike.html

SEC. (2024). *Evaluación formativa en el MCCEMS* [PDF]. Subsecretaria de Educación Media Superior. https://bsu.buap.mx/etj

Shah, P. (2023). *AI and the Future of Education: Teaching in the Age of Artificial Intelligence*. John Wiley & Sons. https://bsu.buap.mx/etf

UE GDPR. (2016). *General Data Protection Regulation* [PDF]. Parlamento Europeo y el Consejo de la Unión Europea. https://gdpr-info.eu/

UK Gov. (2023, 3 de febrero). *Data protection in schools*. Department for Education. https://www.gov.uk/guidance/data-protection-in-schools/managing-breaches-of-data

Unión Europea. (2021). *Plan de Acción de Educación Digital (2021-2027)*. European Commission, Directorate-General for Education, Youth, Sport and Culture. https://education.ec.europa.eu/es/focus-topics/digital-education/action-plan

Unión Europea. (2022). *Directrices éticas sobre el uso de la inteligencia artificial (IA) y los datos en la educación y formación para los educadores* [PDF]. European Commission, Directorate-General for Education, Youth, Sport and Culture. https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_es

Véliz, C. (2020). *Privacy Is Power: Why and How You Should Take Back Control of Your Data*. Bantam. https://bsu.buap.mx/etg

Von Braun, J., Archer, M. S., Reichberg, G. M. y Sánchez Sorondo, M. (2021). AI, Robotics, and Humanity: Opportunities, Risks, and Implications for Ethics and Policy. En J. von Braun, M. S. Archer, G. M. Reichberg, & M. Sánchez Sorondo (Eds.), *Robotics, AI, and Humanity: Science, Ethics, and Policy* (pp. 1-

13). Springer International Publishing. https://doi.org/10.1007/978-3-030-54173-6_1

Zekos, G.I. (2022). Digital Politics, GDPR, and AI. In *Political, Economic and Legal Effects of Artificial Intelligence. Contributions to Political Science*. Springer, Cham. https://doi.org/10.1007/978-3-030-94736-1_11