The ethics of Artificial Intelligence in education: towards responsible and inclusive use*[1]

Julio Cabero-Almenara[2]
Orcid: 0000-0002-1133-6031
Julio Barroso-Osuna[2]
Orcid: 0000-0003-0139-9140

Abstract

This article addresses the relationship between artificial intelligence (AI) and ethics in education, highlighting the need for responsible and inclusive use of this technology. It analyzes how AI, following the emergence of tools such as ChatGPT , is transforming educational processes, from personalized learning to the automation of teaching tasks. However, this technological advancement also entails significant ethical challenges, such as data privacy, equity of access, algorithmic bias, and environmental impact. The article emphasizes the importance of training both students and teachers not only in the technical use of AI, but also in its critical understanding and the adoption of ethical principles that guide its use. To this end, it calls for the establishment of international regulations that regulate this technology, protect human rights, and promote a fair distribution of its benefits. The article emphasizes that the real challenge lies in integrating AI into education in a way that fosters critical thinking, promotes equity, and responds to current and future social needs.

Keywords

Artificial Intelligence – Ethics – Digital Divide – Educational Transformation.

Introduction

To date, two major technologies have revolutionized educational institutions and the ways we access knowledge and information: books and the internet. These technologies have not only transformed educational paradigms but have also generated significant impacts on the access, production, and distribution of knowledge in society.

Regarding books, their incorporation represented a momentous change from both a social and educational perspective. As Harari (2024) points out, in 1424, the Cambridge University library had only 122 volumes, all of them manuscripts; this number indicates the limitations in access to knowledge existing at that time. However, the invention of the printing press, 30 years later, marked a turning point in human history, allowing the production of more than 12 million books in Europe 50 years later.

This technological advancement profoundly transformed education, altering the traditional roles played by teachers and students. Teachers were no longer the sole custodians of knowledge, as students could access information directly through books. Furthermore, the textbook emerged as a fundamental resource that standardized educational content, thus establishing a common foundation for all students in terms of learning and teaching processes.

The emergence of the Internet transformed the way we access information, delocalized it, and broke down the three units on which training was usually based: unity of time, unity of space, and unity of action.

But this situation has been transformed by Artificial Intelligence (AI) technology, which, although it has existed for some time, since the launch of ChatGPT on November 30, 2022, has acquired a leading role in all sectors of society, producing a revolution not only technological, but also social, cultural, economic, political, and of course educational. As Miao *et al.* (2021, p. 7) point out:

> In the last five years alone, thanks to some notable successes and its disruptive potential, artificial intelligence (AI) has moved from the backwaters of academic research to the forefront of public debate, including at the United Nations level.

Its incorporation has had an impact not only on the transformation of the processes followed in different actions and contexts, but also on the labor sector, where its impact on the market is estimated to be significant, since this generative technology could eliminate approximately 300 million full-time jobs. However, a positive impact is also expected, since its introduction in different sectors is thought to have an impact of 7% of global GDP, driven by its integration into multiple sectors (Hatzius *et al.*, 2023; Greenhouse, 2023).

But this strong and rapid penetration has also generated fears and anxieties in the education sector, quickly drawing attention to different aspects such as: updating student assessment methods, establishing policies and regulations in educational institutions for its use, and the need for teacher and student training for quality and ethical use (Lo, 2023). These fears quickly translated into the appearance of manuals, good practice guides, and restriction regulations. for their use prepared by the Universities, as can be clearly seen in the review carried out by González (2024).

Recommendations for use have been made not only by educational institutions, but also by social and political institutions, such as UNESCO (UNESCO, 2021a, 2021b) and the European Commission (European Commission, 2022). In the latter case, it should be noted that the European Union was the first region in the world to establish harmonized standards on artificial intelligence (EU Regulation 2024/1689 of 13 June 2024), which aim to improve the functioning of the internal market by establishing a legal framework, particularly for the development, market introduction, commissioning, and use of artificial intelligence systems.

What can be understood by AI?

When defining what AI is, the first thing is to assume that there is no single definition, although according to the World Commission on the Ethics of Scientific Knowledge and Technology of UNESCO (2019), it can be understood as machines that

> [...] are potentially capable of mimicking or even surpassing human cognitive abilities, including sensing, linguistic interaction, reasoning and analysis, problem-solving, and even creativity.

According to Baker, Smith, and Anissa (2019), AI refers to computer systems capable of performing cognitive tasks generally associated with the human mind, such as learning and problem-solving. This definition provides several key insights: first, AI is not limited to a single technology but encompasses a diverse set of tools, from algorithms to machine learning applications; second, its future evolution is uncertain, which raises, on the one hand, multiple possibilities in its development; and, on the other, a great lack of knowledge about its true capabilities and, therefore, the future consequences that its implementation may have.

But focusing on the educational sphere, what is truly significant is that it is reaching all sectors, functions, and people involved in educational institutions, from administrators to teachers and students, not to mention administrative and service staff and families. Furthermore, it is also having a significant impact on the way the curriculum is organized, teaching materials are generated, student assessments are performed, and the methods by which students are admitted to institutions.

Such is its relevance that, in a short time, research focused on the analysis of its possibilities and educational uses has increased. This is reflected in the increase in meta-analyses that are being carried out on its research in the search for solid principles for its incorporation into teaching (Gallent; Zapata; Ortego, 2023; García Peñalvo *et al.*, 2024; Mena-Guacas *et al.*, 2024; Bond *et al.*, 2024; López-Regalado *et al.*, 2024). Meta-analyses that indicate a series of issues: there is a strong concern about the ethical implications of the use of AI, the need for training for teachers and students, it facilitates the incorporation of innovative teaching practices, the need to carry out educational research to understand its real effect when it is incorporated into teaching and that its use improves academic performance.

Regarding the topic focused on in this article, the relationship between ethics and AI, it should be noted that several meta-analyses of research focused on this issue have already emerged (Vélez-Rivera *et al.*, 2024) that highlight the need for ethical regulations

and standards for the responsible use of AI in higher education, highlighting the importance of balancing its benefits with its ethical challenges and incorporating it from a critical perspective rather than an exclusively technological-instrumental one. Likewise, these meta-analyses underline the need to develop study programs that allow the training of professionals equipped with digital skills and ethical awareness for the design of these tools.

The integration of artificial intelligence into education can take various forms, depending on the educational level in which it is applied. However, in general terms, we can classify it into the following categories:

a) Guide students (teach students how to use AI and use it as a support for students).
b) Provide guidance to teachers (support for teachers in: research, teaching, evaluation, production of resources, construction of assessment instruments and tests, etc.).
c) Guide the institution (support the institution in diagnosis and planning).

More specifically and focused on teaching, they are specified in: use for student selection, promotion of personalized learning, virtual tutoring, sentiment analysis, automatic evaluation of the learning process followed by students and monitoring of the acquired learning process (Albarrán, 2023).

And in all these uses, AI offers a series of advantages for its use and incorporation into educational institutions, which, without intending to exhaust the topic, can be summarized as follows:

a) Providing information on students' backgrounds, allowing the training process to be tailored to their needs, characteristics, and training weaknesses.
b) Assist the teacher in making decisions about learning content.
c) Facilitate the organization of activity planning by the teacher.
d) Monitor the learning acquired by students.
e) Reduce teachers' workload.
f) Reduce the administrative burden on teachers.
g) Monitoring the learning process followed by the student.
h) Better prediction/evaluation of teacher performance/results.
i) Contribute to the achievement of the Sustainable Development Goals.
j) Powering a vast majority of apps, known as generative AI, through natural language. This makes them easier to use.
k) Carrying out automated assessments and evaluations. (Celik *et al.*, 2022; High, 2023; Curtain, 2024).

Faced with these possibilities, AI also presents a series of limitations, risks, and concerns that must be considered when incorporating it into educational institutions, such as:

a) Lack of precision and validity of the information generated, the appearance of so-called "hallucinations."

b) Issues surrounding responsibility and ethics regarding data acquisition by AI.

c) Curricular adaptation and role change.

d) Limitations and biases of the model.

e) Destruction of jobs, with the social implications that this entails.

f) Protection of intellectual property rights.

g) Limited mathematical reasoning.

h) Data privacy and security.

i) Problem of assuming it as the only truth.

j) Rethinking the ways in which students are assessed, due to the risks of plagiarism.

k) The lack of deep understanding of how it works.

l) Adoption and acceptance (Cooper, 2023; Jiménez *et al.*, 2023; Farrokhnia *et al.*, 2023; Mayol, 2023; Curtain, 2024).

Considering these limitations is of utmost importance, firstly, because, as will be analyzed throughout this work, many of them have ethical and legal implications; and, secondly, because both teachers and students tend to adopt very positive attitudes toward AI, without developing a critical vision regarding its possibilities and limitations, which can lead to an automatic and unconscious use of this tool. This is even more dangerous in the case of students, since AI has established itself as one of their preferred technological tools for carrying out academic activities, and they even use it when they are explicitly warned not to use it, as evidenced by Ramírez and López (2024) in their research.

On the other hand, it should not be overlooked that its use is underpinned by a certain assumption of "technocentrism", where everything possible to achieve depends on the machine, as reflected in the following post written on a website in which technology is granted a magical power to transform education: "By combining artificial intelligence with approaches such as autonomous learning and peer collaboration, we can create more personalized and effective educational experiences. AI not only optimizes time and feedback, but also boosts participation and critical thinking. Power your classroom without borders!"

The educational debate surrounding AI has gained increasing relevance in recent years, and its analysis can be approached from at least two broad, complementary perspectives. The first, of a didactic nature, explores how AI tools can be integrated into teaching and learning processes, transforming pedagogical methodologies, educational resources, and the dynamics between teachers and students. The second, of a more philosophical and political nature, focuses on the social, cultural, and ethical implications of educating citizens for a world profoundly influenced by this technology.

From a didactic perspective, AI offers a wide range of possibilities for personalizing educational processes, improving accessibility, and enriching learning experiences. Machine learning- based systems allow content to be adapted to the needs and pace of each student, facilitating more inclusive and equitable education. However, this potential must be balanced with a critical reflection on the role of AI in developing autonomous individuals capable of critical thinking and not exclusively dependent on this technology.

The second perspective, which adopts a philosophical and political approach, requires a deeper analysis of the ethical, social, and political implications of living in a world transformed by AI. As highlighted by Ivanov (2023), Zawacki-Richter *et al.* (2019), and Cortina (2024), this reflection demands a critical questioning of both the challenges and risks posed by AI, as well as the values and principles that should guide its integration into society.

As mentioned in this introduction, the use of AI can entail various technical problems, as pointed out by different authors and institutions (UNESCO, 2021a; European Commission, 2022; Alonso-Rodríguez, 2024; Cortina, 2024), an issue that will be addressed below.

AI and ethics in its educational use

Since the emergence of AI, there has always been concern about the ethical aspects related to its use. UNESCO, at its Beijing meeting (UNESCO, 2018), made a series of recommendations (7, 12, 23, 24, 25, 26, 27, and 28) where it drew attention to the need for AI development to be human-controlled and people-centered. These recommendations, although suggesting its use in teaching, also drew attention to a number of aspects, such as: the need to maintain interaction between students and teachers; ensuring that AI promotes educational opportunities without distinction based on gender, disability, or social status; ensuring that students with disabilities or learning difficulties are not discriminated against; and ensuring that gender-based discrimination does not occur.

To these concerns we must add the correct vision shown by Cortina (2024, p. 205) when he states

> [...] it is no less true that techno-education has become a powerful economic weapon in the hands of companies and countries capable of leading it, so that those left behind in the competition for first place will have to pay dearly for the opportunity cost and will lose economic and political power.

For O'Neil (2017), the use of AI is generating various problems. On the one hand, the concentration of power in the hands of a small group of corporations and countries, primarily the United States and China, is increasing the global digital divide. On the other hand, the environmental impact of AI is significant, as the operation of data centers and model training require large amounts of energy and resources, contributing significantly to the climate crisis. Furthermore, gaps in ethics and transparency arise, given that AI systems often operate in an opaque manner, making it difficult to understand the reasons behind their decisions. Furthermore, its implementation has led to surveillance that threatens fundamental rights such as privacy and freedom of expression, while its algorithms can discriminate against certain groups. There are two models of data control in AI: in one case, it is controlled by companies, as in the United States, and in the other, by the state, as in China. Added to this is its impact on the workplace and its potential use for military purposes. Given this situation, the author emphasizes the need to develop international policies that regulate AI, protect human rights, and ensure the equitable distribution of its benefits.

This leads to a necessary reflection, since the real challenge does not simply lie in learning how to use AI applications, as this learning will be a natural and inevitable process that most people will acquire over time and practice, given that AI will be incorporated into a variety of technologies, and they work almost without us realizing it. Rather, the main challenge lies in understanding and managing the profound changes that AI will bring with it as it evolves and becomes increasingly powerful and sophisticated, tending towards the realization of distinctly human cognitive skills, such as: mental calculation, spatial navigation, routine problem solving, or basic analytical thinking (Tuomi, 2018; Loján *et al.*, 2024). These changes will affect not only the way we interact with technology, but also our social structures, the ethical values that guide our decisions, the prevailing economic models, and even our perception of the world and of ourselves as human beings in a digitalized society. The focus, therefore, must be on how to anticipate, adapt, and steer these impacts toward a more equitable, inclusive, and conscious future.

Entering directly into the topic of the relationship between AI and ethics, and following Cortina (2024), two fundamental perspectives can be identified in this relationship: on the one hand, establishing principles and, on the other, people acquiring ethical principles for its responsible use. Both perspectives will be addressed below.

Establish ethical principles in the design and construction of AI models

One of the criticisms of AI concerns the biases presented by the information it provides. Specifically, we can talk about what Sabzalieva & Valentini (2023, p. 11) call a cognitive bias, which suggests that AI

> [...] is not governed by ethical principles and cannot distinguish between right and wrong, true and false. This tool only collects information from databases and texts it processes on the Internet, so it also learns any cognitive biases found in that information.

To this perspective, the dimension of gender and diversity can be incorporated, as a consequence of the "lack of female participation in issues related to AI, in AI research and development and, on the other hand, the power of generative AI to produce and disseminate content that discriminates or reinforces gender and other stereotypes." (Sabzalieva; Valentini, 2023, p. 11). All of this is a result of the data with which these systems are trained and which can lead to biases in their responses, and thus affect the quality, veracity and precision of the answers offered. In a report prepared by IBM (2023), various examples are presented that show the existence of biases in artificial intelligence systems, which can lead to significant inequalities in different areas. In the healthcare sector, for example, computer-aided diagnosis (CAD) systems have been identified as having lower accuracy in assessing patients from minority groups, such as Black people, compared to white patients, due to the underrepresentation of certain data in training sets. Similarly, in digital advertising, it has been observed that targeting algorithms used

by platforms such as Google tend to more frequently show higher-paying job offers to men than to women, perpetuating gender gaps in access to economic opportunities. Meanwhile, in the security field, the use of AI-based surveillance and crime prediction tools has revealed biases stemming from training on historical arrest data, leading to an overrepresentation of racial profiling and a disproportionate targeting of marginalized communities.

These AI-induced biases are clearly observed in relation to a person's gender, age, and skin color in AI-powered image generation programs, where young age, men, and white people tend to be prioritized when generating images related to job positions (Thomson; Thomas, 2023). As Ortiz de Zárate (2023, p. 9) points out: "For years, various studies, some of the most recent from 2018, showed that while almost 100% of white men were successfully recognized by facial recognition systems, the success rate dropped to 35% in the case of racialized women."

This situation is having repercussions in various areas, which has led to calls (Yu; Guo, 2023) to establish appropriate measures to develop explainable and fair algorithms, update encryption technology and formulate relevant laws and regulations to protect data, as well as improve the quality and quantity of the data sets with which the models are trained. At the same time, new AI training models are suggested, and in this sense, UNESCO (2023) proposes changing the models used for AI training, and adopting new formats and proposals:

> [...] it is essential that efforts be made to refine the basic models, not only by adding subject matter knowledge and removing biases, but also by adding knowledge about relevant learning methods, and how this can be reflected in the design of algorithms and models. The challenge is to determine to what extent EdGPT models can go beyond subject knowledge to also address learner-centered pedagogy and positive interactions between teachers (UNESCO, 2023, p. 13).

In addition to the risks arising from bias, another ethical issue arises, referring to the difficulty we may encounter in protecting citizens' privacy. Every time we interact with AI, we are providing it with information about ourselves, our likes and dislikes, and therefore providing information that allows it to understand us in depth.

This aspect of privacy is particularly concerning when analyzing one of the possibilities that AI offers to teaching with so-called "personalized learning"; that is, the adaptation of training to the cognitive characteristics and needs of the student. To implement this model, the machine must have certain information about the student, and therefore a series of questions arise: Who guarantees that this information cannot be used in the future for purposes other than those for which it was initially requested? Or who ensures that the data collected about the student will not be used in the future in selection processes, such as for job selection, or to offer commercial products to the student or his or her family?

In this last aspect, Alonso-Rodríguez (2024, p. 84) draws attention to the fact that the

> [...] AI developments in education may interfere with people's autonomy and responsibility and hinder universal rights (UN General Assembly, 1948) such as privacy (Art. 12), equality (Art. 1), and non-discrimination (Art. 2). This has broad social and ethical implications.

Without forgetting that in the field of education, a series of reflections must be made regarding the future consequences that some of the decisions currently being made regarding the use of AI may have for people:

> The potential harms arising from diagnosing and predicting students' learning outcomes that may affect their future development. 2) The problems generated by the decisions that AI systems could make due to their influence on the educational decisions of teachers, families, and other stakeholders (including legislators). 3) The impacts on the development and maturity of individuals, especially in the early stages of education, as a consequence of the change in roles that alters the relationship between teachers and students (Alonso-Rodríguez, 2024, p. 85).

The words pointed out by Cortina (2024, p. 31) can serve to clarify what is being discussed, for this author: "between making use of intelligent systems (be they machines, algorithms or robots) when making decisions and delegating to these intelligent systems significant decisions for the lives of people and nature."

But talking about establishing ethical principles for the use of AI implies not forgetting that its incorporation into the education system is widening the digital divide between countries and groups of people. Those with greater resources are incorporating it more easily and en masse, and consequently are more likely to benefit from the advantages it offers. This will widen the already existing digital divide, as evidenced during Covid-19, directly affecting social and educational equity. This situation is exacerbated by the fact that not all services or versions of AI apps are free. Consequently, the type of use a person can make of it and the quality of the product the tool offers will be conditioned by their social and economic characteristics, and this will become an element of widening the digital divide between people. For example, in the case of ChatGPT, the free version 3.5, compared to the paid version 4.0, has a limited capacity to hold long conversations, only works with text, makes more errors and biased information, and reviews the information up to a specific date.

Acquisition of ethical values by the person when using AI

It's not just about adopting a series of measures to ensure that models are designed with ethical considerations in mind. It's also about training people to use AI ethically and responsibly.

It is no surprise to us that, since the appearance of ChatGPT, the main concerns of teachers and educational administrators have focused on the possibility it offers to facilitate plagiarism of student work, and therefore encourage dishonesty among certain students when completing academic work.

This situation has been exacerbated by the limited effectiveness of traditional plagiarism detection tools available on the market, or by the ineffectiveness of tools for detecting work generated by ChatGPT or similar content generation tools in ensuring 100% accuracy of how the text was generated (GPTZero, Plagscan, etc.), as well as by the lack of more advanced options. Furthermore, the idea that teachers could identify misuse of these technologies based on the language used has not been supported by research.

However, the situation should not only be approached from what has already been discussed, but from aspects much more transcendental for the future of individuals and society. Authors such as Desmurget (2020) or Luri (2020) have been drawing attention, on the one hand, to the conceptual and scientific impoverishment of the people being trained in our educational institutions, and on the other, to the decrease in the levels of demand and requirements requested of students to be able to pass the courses and training cycles. Demands that are progressively decreasing, due, on the one hand, to the pressure from institutions, which in their standards must achieve successful levels of students who pass the course, and on the other, by a certain enlightened "pedagogical progressivism" of many "YouTube" pedagogues, where the effort to be made by students is in seeing, and in "goodness" and not in understanding and the realization of cognitive effort.

There should be no doubt that this situation will progressively lead not only to a conceptual impoverishment of students, but also to profound consequences for the future and for the social advancement prospects of students from working-class backgrounds. We have always maintained that public educational institutions should require a high level of effort, as this is the best social policy to facilitate the advancement and social progress of the working classes. In this sense, Cortina's reasoning is clear and irrefutable:

> [...] people with purchasing power take care to give their children a personalized education, which demands effort, linguistic and cultural skills of all kinds so that they are well prepared, while in public education, often, instead of providing students with a similar preparation, all kinds of facilities are provided so that they can obtain diplomas without any effort. This is a clearly demagogic approach, which reinforces what might be called "the cultural poverty trap," because it is a dead end. It does not help to educate in excellence, but in mediocrity (Cortina, 2024, p. 218).

However, faced with this aspect, a lax use of AI in training processes, which is limited to facilitating the completion of work and activities assigned by teachers, there is a more complex problem and that is the lack of vision and critical reflection by the student of the information with which they work, a key aspect for the formation of a professional personality that can carry out a transformation of social reality. Precisely, this is one of the great limitations that AI can bring to education, since, by offering a comfortable and simple tool for completing schoolwork, it can lead students to relegate the cognitive effort necessary to construct critical thinking about the reality they analyze, which ends up simplifying and impoverishing their intellectual maturity.

To some extent, and in relation to what was discussed above, the problem arises from the possible lack of autonomy that students may acquire, as they become overly

dependent on the tool for their educational development and delegate the implementation of different training activities to it, while casually accepting the results it offers.

As Morozov (2024) recently pointed out, we can incorporate AI from two perspectives, one would be to use it to augment the person and the other to improve it.

> The difference between these two paradigms is subtle but crucial. Augmentation is when we use our cell phone's GPS to navigate an unfamiliar place: it allows us to reach our destination faster and more easily. The benefit is fleeting. If this technological aid were taken away from us, we would feel even more helpless. Enhancement involves using technology to develop new skills–in this case, refining our innate sense of direction by using advanced memorization techniques or learning to understand the signs of nature. In short, augmentation deprives us of certain capabilities in the name of efficiency, while enhancement allows us to acquire others and enriches our interactions with the world (Morozov, 2024, p. 26).

Furthermore, it is essential to adopt a critical view of AI, and this requires recognizing that technologies, including AI, are not neutral; they are imbued with the ideologies, interests, and values of those who design and control them. This fact highlights the need to train citizens with advanced digital skills, enabling them not only to use these technologies but also to understand how they work, how they are created, and the interests underlying their development and application.

This suggests that the real educational challenge lies not only in teaching how to use AI tools, but in preparing students for a world in which this technology plays a central role. This involves fostering critical and ethical education that transcends mere technical mastery and promotes a deep understanding of the social, economic, and cultural implications of AI.

Educating critical and conscious citizens in a world with AI involves challenging narratives that present this technology as inevitable or inherently beneficial, instead promoting a dialogue that leads to an understanding of its limitations, biases, and risks. Only through this critical reflection will it be possible to build a society in which AI is used responsibly and equitably, respecting the rights and dignity of all people.

This approach not only redefines educational priorities, but also raises an urgent question: how to educate in an era where AI is not just a tool, but a transformative actor in the social fabric? Or how to create a critical awareness in citizens that allows them to understand and adapt to the chaotic situations that will arise in the early stages of AI penetration in all sectors? As Ferrarelli (2024) recently pointed out in a report for the OEI regarding how to approach AI from Latin America, its use will not only present a series of ethical dilemmas, some of which have been analyzed in this article, but also environmental, labor, and pedagogical dilemmas. Furthermore, as the author of the report warns, there are a series of misunderstandings that should be unraveled, such as that AI is not: "a search engine," "neutral," "reliable," and "creative."

References

ALBARRÁN, Eloy. **Hacia una educación personalizada y adaptativa**: la disrupción de la inteligencia artificial. Caracas: Centro Internacional de Educación Continua: Universidad Pedagógica Experimental Libertador, 2023.

ALONSO-RODRÍGUEZ, Ana María. Hacia un marco ético de la inteligencia artificial en la educación. **Teoría de la Educación: Revista Interuniversitaria**, v. 36, n. 2, p. 79-98, 2024. https://doi.org/10.14201/teri.31821

ALTO, Valentina. **Inteligencia artificial generativa como modelos de ChatGPT y OpenAI**. Madrid: Anaya, 2023.

BAKER, Toby; SMITH, Laurie; ANISSA, Nandra. Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. **Nesta**, 2019. Disponible en: https://www.nesta.org.uk/report/education-rebooted/. Acesso en: 20 Mayo 2025.

BOND, Melissa; KHOSRAVI, Hassan; DE LAAT, Maarten; BERGDAHL, Nina; NEGREA, Violeta; OXLEY, Emily; PHAM, Phuong; CHONG, Sin Wang; SIEMENS, George. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. **International Journal of Educational Technology in Higher**, v. 21, 2024. https://doi.org/10.1186/s41239-023-00436-z

CELIK, Ismail; DINDAR, Muhterem; MUUKKONEN, Hanni; JÄRVELÄ, Sanna The promises and challenges of artificial intelligence for teachers: a systematic review of research. **TechTrends**, v. 66, p. 616-630, 2022. https://doi.org/10.1007/s11528-022-00715-y

COMISIÓN EUROPEA. **Directrices éticas sobre el uso de la inteligencia artificial (IA) y los datos en la educación y formación para los educadores**. Luxemburgo: Unión Europea, 2022.

COOPER, Grant. Examining science education in ChatGPT: an exploratory study of generative artificial intelligence. **Journal of Science Education and Technology**, v. 32, p. 444-452, 2023. https://doi.org/10.1007/s10956-023-10039-y

CORTINA, Adela. **¿Ética o ideología de la inteligencia artificial?** Barcelona: Paidós, 2024.

DESMURGET, Michel. **La fábrica de cretinos digitales**. Barcelona: Paidós, 2020.

FARROKHNIA, Mohammadreza; BANIHASHEM, Seyyed Kazem; NOROOZI, Omid; WALS, Arjen. A SWOT analysis of ChatGPT: implications for educational practice and research. **Innovations in Education and Teaching International**, v. 61, n. 3, p. 460-474, 2023. https://doi.org/10.1080/14703297.2023.2195846

FERRARELLI, Mariana. **Inteligencia artificial y educación**: insumos para su abordaje desde Iberoamérica. Buenos Aires: OEI, 2024.

GALLENT, Cinta; ZAPATA, Alfredo; ORTEGO, José Luis. El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. **Relieve**: Revista Electrónica De Investigación y Evaluación Educativa, v. 29, n. 2, 2023. https://doi.org/10.30827/relieve.v29i2.29134

GARCÍA PEÑALVO, Francisco; LLORENS-LARGO, Faraón; VIDAL, Javier. La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. **Ried:** Revista Iberoamericana de Educación a Distancia, v. 27, n. 1, p. 9-39, 2024. https://doi.org/10.5944/ried.27.1.37716

GONZÁLEZ, Gregorio. **1 d.C. (Después de ChatGPT):** inteligencia artificial en la educación superior. Valencia: PUV Universitat de València, 2024.

GREENHOUSE. US experts warn AI likely to kill off jobs and widen wealth inequality. **The Guardian**, 2023. Disponible en: https://is.gd/n38xqn. Acceso en: 12 ene. 2025.

GRUPO DE TRABAJO DE INTELIGENCIA ARTIFICIAL GENERATIVA DE LA UNAM. **Recomendaciones para el uso de inteligencia artificial generativa en la docencia**. México, DC: UNAM, 2023.

HARARI, Yuvak. **Nexus**: una breve historia de las redes de información desde la Edad de Piedra hasta la IA. Barcelona: Debate, 2024.

HATZIUS, Jan; BRIGGS, Joseph; KODNANI, Devest; PIERDOMENICO, Giovanni. **The potentially large effects of artificial intelligence on economic growth**. Goldman Sachs, 2023. Diponible en: https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html. Acesso en: 20 Mayo 2025.

IVANOV, Stanislav. The dark side of artificial intelligence in higher education. **The Service Industries Journal**, v. 43, p. 1055-1082, 2023. https://doi.org/10.1080/02642069.2023.2258799

INTERNATIONAL BUSINESS MACHINES. El sesgo de la IA, a la luz de ejemplos reales. **IBM**, 2023. Disponible en: https://www.ibm.com/es-es/think/topics/shedding-light-on-ai-bias-with-real-world-examples. Acesso em: 19 maio 2025.

JIMÉNEZ, Luis; LÓPEZ-GÓMEZ, Julio; MARTÍN-BAOS, José Angel; ROMERO, Francisco; SERRANO-GUERRERO, Jesús. ChatGPT: reflexiones sobre la irrupción de la inteligencia artificial generativa en la docencia universitaria. **Actas de las Jenui**, n. 8, p. 113-120, 2023.

LO, Chung. What is the impact of ChatGPT on education? A rapid review of the literature. **Education Sciences**, v. 13, n. 4, p. 410, 2023. https://doi.org/10.3390/educsci13040410

LOJÁN, María del Cisne.; ROMERO, J. Antonio; SANCHO AGUILERA, David; ROMERO, A. Yajaira. Consecuencias de la dependencia de la inteligencia artificial en habilidades críticas y aprendizaje autónomo en los estudiantes. **Ciencia Latina Revista Científica Multidisciplinar,** v. 8, n. 2, p. 2368-2382, 2024. https://doi.org/10.37811/cl_rcm.v8i2.10678

LÓPEZ REGALADO, Oscar; NÚÑEZ-ROJAS, Nemecio; RAFAEL LÓPEZ GIL, Oscar; SÁNCHEZ-RODRÍGUEZ, José. El análisis del uso de la inteligencia artificial en la educación universitaria: una revisión sistemática. **Pixel-Bit**: Revista de Medios y Educación, n. 70, p. 97-122, 2024. https://doi.org/10.12795/pixelbit.106336

LURI, Gregori. **La escuela no es un parque de atracciones**. Barcelona: Paidós, 2020.

MAYOL, Julio. Inteligencia artificial generativa y educación médica. **Educación Médica**, v. 24, n. 4, p. 1-3, 2023. https://doi.org/10.1016/j.edumed.2023.100851

MENA-GUACAS, Andrés; VÁZQUEZ-CANO, Estevan; FERNÁNDEZ-MÁRQUEZ, Esther; LÓPEZ-MENESES, Eloy. La inteligencia artificial y su producción científica en el campo de la educación. **Formación Universitaria**, v. 17, n. 1, p. 155-164, 2024. http://dx.doi.org/10.4067/s0718-50062024000100155

MIAO, Fengchun; HOLMES, Wayne; HUANG, Ronghuai; HUI, Zhang. **Inteligencia artificial y educación:** promesas e implicancias. París: Unesco, 2021.

MOROZOV, Evgeny. Otra inteligencia artificial es posible. **Le Monde Diplomatique en Español**, n. 346, p. 25-28, 2024.

O'NEIL, Cathy. **Armas de destrucción matemática**. Madrid: Capitán Swing, 2017.

ORTIZ DE ZÁRATE, Lucía. Sesgos de género en la inteligencia artificial. **Revista de Occidente**, n. 502, p. 5-20, 2023.

RAMÍREZ, Gina; LÓPEZ, Alison. Reflexiones éticas sobre la evaluación y la inteligencia. *In*: MUÑOZ, Juan Miguel; LORENZO, Neus; SUÑÉ, Xavier (org.). **Inteligencia artificial en la microeducación**: transformando el aula del futuro. [*S. l.*]: Odite, 2024. p. 123-134.
SABZALIEVA, Emma; VALENTINI, Arianna. **ChatGPT e inteligencia artificial en la educación superior**. Santiago de Chile: Unesco-Iesalc, 2023.

THOMSON, Travis Jason; THOMAS, Ryan. Ageism, sexism, classism and more: 7 examples of bias in AI-generated images. **The Conversation**, 2023. Disponible en: https://theconversation.com/ageism-sexism-classism-and-more-7-examples-of-bias-in-ai-generated-images-208748. Acesso em: 19 maio 2025.

TUOMI, Ilkka. The impact of artificial intelligence on learning, teaching, and education. Policies for the future. *In*: CABRERA, Marcelino.; VUORIKARI, Riina.; PUNIE, Yves. (ed.). **EUR 29442 EN**. Luxembourg: European Union, 2018. p. 1-48. Disponível em: 10.2760/12297. JRC113226. Acesso em: 19 maio 2025.

UNESCO. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. **Beijing consensus on artificial intelligence and education**. Beijing: Unesco, 2019. Disponívle em: https://unesdoc.unesco.org/ark:/48223/pf0000368303. Acesso em: 19 maio 2025.

UNESCO. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. **Guidance for generative AI in education and research**. París: Unesco, 2023.

UNESCO. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. **Inteligencia artificial y educación**: guía para las personas a cargo de formular políticas. París: Unesco, 2021a.

UNESCO. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. **Recomendación sobre la ética de la inteligencia artificial**. París: Unesco, 2021b.

VÉLEZ, Rocio; MUÑOZ ÁLVAREZ, Diana; LEAL ORELLANA, Priscila; RUIZ GARRIDO, Alejandra. Uso de Inteligencia Artificial en educación superior y sus implicancias éticas: mapeo sistemático de literatura. **Hachetetepé: Revista científica de Educación y Comunicación**, v. 28, p. 1105, 2024. https://doi.org/10.25267/Hachetetepe.2024.i28.1105

YU, Hao; GUO, Yunyun. Generative artificial intelligence empowers educational reform: current status, issues, and prospects. **Frontiers in Education**, v. 8, 1183162, 2023. https://doi.org/10.3389/feduc.2023.1183162

ZAWACKI-RICHTER, Olaf; MARÍN, Victoria; BOND, Melissa; GOUVERNEUR, Franziska. Systematic review of research on artificial intelligence applications in higher education – where are the educators? International **Journal of Educational Technology in Higher Education**, v. 16, n. 1, p. 1-27, 2019. https://doi.org/10.1186/s41239-019-0171-0

Editor: Prof. Dr. Hugo Heredia Ponce

Julio Cabero Almenara is a professor at the University of Seville. He is the Director of the Didactic Research Group (GID-HUM-0390) and President of EDUTEC. His area of expertise is Applied Technologies in Education.

Julio Barroso Osuna is a professor at the University of Seville. He is a member of the Didactics Research Group (GID-HUM-0390) and EDUTEC. His area of expertise is related to Technologies Applied to Education.