



Artificial intelligence as social technology in education: ethically rethinking contemporary paradigms

Inteligencia artificial como tecnología social en la educación: repensando éticamente los paradigmas contemporáneos

Efraín BÁMACA-LÓPEZ¹

Investigador asociado, Escuela de Periodismo, Facultad de Humanidades - Universidad de Santiago de Chile

<https://orcid.org/0000-0003-0518-2600>

edi.bamaca@usach.cl

Pablo DE LA VEGA²

Universidad Rafael Landívar - Guatemala, Guatemala

Red Internacional ALEC

<https://orcid.org/0009-0001-7547-170X>

padelavega@url.edu.gt

URL : <https://www.unilim.fr/trahs/7793>

DOI : 10.25965/trahs.7793

Licence : CC BY-NC-SA 4.0 International

Abstract: This paper evaluates the current paradigm of Artificial Intelligence and proposes an ethical approach to education that invites the outlining of guidelines. These are the principles of ethical practice: the principle of beneficence, which promotes the well-being; the principle of nonmaleficence, which seeks to avoid harm; the principle of autonomy, which drives self-determination and self-choice; and the principle of justice, which seeks harmony and equity. Added to this is the principle of explicability, which invites to understand the importance and how to act with AI in contemporary Education. However, if these principles are not applied with a broad vision of human needs, marginalization in technology will continue, preserving the social gaps in pedagogical models. Therefore, the possibilities of considering Artificial Intelligence as a social technology for educational purposes are required and respond to the objectives that go beyond the current pedagogical paradigm and promote the goals that all social technology seeks: a dignified future for human beings.

Keywords: education, Social technologies, Artificial Intelligence, Ethics, Human necessities

¹ Doctor en Ciencia, Tecnología y Sociedad, Máster en Estudios Latinoamericanos, Licenciado en Comunicación Social con mención en Radio y Televisión.

² Master of Arts en Filosofía por la Hochschule für Philosophie, Múnich, Alemania y Máster en Estudios Avanzados en Literatura Española e Hispanoamericana, Universitat de Barcelona, España. Profesor universitario de pregrado y posgrado en varias universidades de Guatemala.

Resumen: Este artículo evalúa el paradigma actual de la Inteligencia Artificial y propone un enfoque ético en la educación que invita a esbozar algunas directrices. Estos son los principios de la práctica ética: el principio de beneficencia, que promueve el bienestar; el principio de no maleficencia, que busca evitar el daño; el principio de autonomía, que impulsa la autodeterminación y la autoelección; y el principio de justicia, que busca la armonía y la equidad. A estos se suma el principio de explicabilidad, que invita a comprender la importancia y cómo actuar con la IA en la educación contemporánea. Sin embargo, si estos principios no se aplican con una visión amplia de las necesidades humanas, continuará la discriminación en la tecnología, preservando las brechas sociales en los modelos pedagógicos. Por lo tanto, se requiere la consideración de las posibilidades de la Inteligencia Artificial como una tecnología social con fines educativos, para responder a los objetivos que van más allá del paradigma pedagógico actual y promueven las metas que toda tecnología social busca: un futuro digno para los seres humanos.

Palabras clave: educación, tecnologías sociales, inteligencia artificial, ética, justicia social

Résumé : Cet article évalue le paradigme actuel de l'intelligence artificielle et propose une approche éthique de l'éducation, invitant à esquisser quelques lignes directrices. Il s'agit des principes de cette pratique éthique : le principe de bienfaisance, qui promeut le bien-être ; le principe de non-malfaisance, qui vise à éviter de nuire ; le principe d'autonomie, qui favorise l'autodétermination et le libre choix ; et le principe de justice, qui recherche l'harmonie et l'équité. À cela s'ajoute le principe d'explicabilité, qui invite à comprendre l'importance de l'IA dans l'éducation contemporaine et la manière d'agir avec elle. Toutefois, si ces principes ne sont pas appliqués dans une perspective globale des besoins humains, la marginalisation technologique persistera, perpétuant les inégalités sociales dans les modèles pédagogiques. Par conséquent, il est nécessaire d'envisager l'intelligence artificielle comme une technologie sociale au service de l'éducation, répondant ainsi à des objectifs qui dépassent le paradigme pédagogique actuel et promeuvent les buts que toute technologie sociale poursuit : un avenir digne pour l'humanité.

Mots clés : éducation, technologies sociales, intelligence artificielle, éthique, justice social

Resumo: Este artigo avalia o paradigma atual da Inteligência Artificial e propõe uma abordagem ética na Educação que convida a esboçar algumas diretrizes. São elas: os princípios da prática ética: o princípio da beneficência, que promove o bem-estar; o princípio da não maleficência, que busca evitar danos; o princípio da autonomia, que impulsiona a autodeterminação e a autoescolha; e o princípio da justiça, que busca a harmonia e a equidade. A estes se soma o princípio da explicabilidade, que convida a compreender a importância e a forma de atuação da IA na Educação contemporânea. Contudo, se esses princípios não forem aplicados com uma visão ampla das necessidades humanas, a marginalização na tecnologia persistirá, perpetuando as lacunas sociais nos modelos pedagógicos. Portanto, é necessário considerar a Inteligência Artificial como uma tecnologia social para fins educacionais, que responda a objetivos que transcendem o paradigma pedagógico atual e promovam os objetivos almejados por toda tecnologia social: um futuro digno para os seres humanos.

Palavras chave: educação, Tecnologias sociais, Inteligência artificial, Ética, justiça social

Introduction

In contemporary discussions of technology, Artificial Intelligence (AI) cannot be excluded. Owing to its wide range of applications, it is currently the most important technological innovation with massive repercussion on everyday human life. From areas of individual interaction to governmental and community-based organizations, contributing to health business, politics, entertainment, and every other field of human development, especially education. Simultaneously, it has influenced fields of knowledge that had not previously been contemplated and has made globalization progress with huge steps towards a future with many uncertainties. This is due to a disjunction: Is AI the solution to global problems or a tool that establishes the boundaries for an integral development of all human beings, affecting a great number of individuals and creating new gaps and necessities?

Since the implementation of AI involves the creation of digital worlds, augmented realities and a growing field of virtual possibilities, it is considered as an area of singular impact on every human being and becomes an imperative in any technological reflection. Especially in a world that is still divided and where every advance can set a route of benefit or prejudice. Therefore, various aspects can be seen in its application, such as in education, a vast field that highlights the need to improve educational quality in general, but with particular attention to less developed countries where technology can be ideal for educational progress. This improvement in educational quality can be aided by technological tools, such as Artificial Intelligence. Against this background, considering AI in education invites examination of how it can be used for pedagogical benefit, and not just as a technological development aimed at sustaining the educational gap, which have consequences in the global hierarchy and segregation of social benefits. This only makes more difficult access of basic needs and, therefore, improving the quality of life of every human being.

For the purposes of this manuscript, education will be understood as the formal and informal teaching-learning process, which, based on a curriculum established within a national framework, seeks to provide tools for personal development in a given social environment, motivating the full development of the human being within their context and providing them with the instruments to navigate life's situations. Thus, education can be enhanced using Artificial Intelligence.

Within this framework, the concept of social technology becomes relevant, especially in relation to the educational context. It constitutes a set of applications of technology with the purpose of solving problems of developing communities, primarily those that promote improvements in the quality of life through new knowledge and the application of more effective learning approaches. Social technology, through its practices and principles, seeks the inclusion, valorization and implementation of technology in such a way that solutions are managed to the problems of the most affected groups, especially those where the lack of resources and access to basic services constitutes the main threat. Ignoring ethical considerations encourages a thoughtless technological advance, favorable with the dynamics of a system that disregards human beings as ends in themselves and prioritize production objectives over environmental sustainability, thereby undermining the needs of the most vulnerable. These perspectives can be changed through an educational approach grounded in social and ethical focus.

For this reason, social technology in educational contexts seeks to foster meaningful change and achieve improvements for different social groups. If artificial intelligence is considered as the vanguard technological innovation of our time,

being part of every pedagogical and social field, it is imperative to ensure that its development and implementation are oriented toward its application and understanding as a form of social technology. This corresponds to reconceptualization and its appropriate application in educational institutions. The aspects that make artificial intelligence a social technology in Education and the minimum components that must be considered to achieve this objective are proposed below.

I- On technologies and social technologies

Every technological discovery and advance implies a dilemma: whether its application is beneficial or detrimental to human beings. Concurrently, this can be examined in other spheres, analyzing whether it affects nature and its components, having a direct or indirect impact on people. Critical reflection on the ethical implications did not become evident until a few decades ago, having been previously considered in an abstract way, emphasizing the technological influences on some social and cultural aspects (Verbeek, 2005). The need to talk about this influence is due to the progressive trust that humans place in technology (Nickel, et al., 2010), an aspect that has been cemented as the same technology increases its capabilities and areas of application, facilitating everyday practices in favor of comfort in a wide number of activities.

Given this scenario, the ways in which humans and technology coexist become increasingly complex, manifesting in forms and practices that were once regarded as utopias, bifurcating meanings and making a human being synonymous with using technology (Ede, 2019). While historical development presents a rudimentary use of technology during the first millennium of Western civilization, after the Renaissance and the knowledge revolutions, in the 16th and 17th centuries, certain *remundifying* reactions were applied (Castany, 2012), where the use of technology becomes the driving force shaping the trajectories humanity will follow in the coming centuries. Here, more specific and elaborate ways of teaching *techné* allow the creation of different sciences and knowledge about nature and the physical world with precision and veracity, transforming the ways in which human beings were educated.

The 18th and 19th centuries were a period of significant improvement in quality of life, where the technological approach favored health, nutrition and access to basic commodities for humanity, and consequently, a wide access to education. All this thanks to the industrial pioneers and the rise of their innovative practices, including the use of new materials and the implementation of mass production. During the peak of this period, in the 20th century, this same development made it possible to work towards human needs in its different areas, such as building a better infrastructure for communication, commerce and social development, as well as the increase in the application of technology in resources from countless fields: society, economy, pedagogy, medicine, industry, architecture, etc. This shows how human knowledge and learning were translated into tangible benefits and ways to interact with the world. Also, this prompted reflection on environmental protection and the effective use of natural common goods, thus creating different epistemic currents that sought significant changes in the ways technologies are used.

Meanwhile, the computational revolution of the 21st century has driven advances in other areas, especially in digital technologies. The pinnacle was established with the spread of free AI in different human environments, resulting in the observation that “the ever increasing use of artificial intelligence in technical artifacts has resulted in a situation in which in many cases the behavior of technical artifacts comes close to human behavior” (Nickel, et al., 2010: 430). This has prompted the current

transformation in how technology is taught, understood, implemented and considered, while creating epistemic and ontological challenges, inviting to ask whether human constitution follows technological patterns and whether these new technologies will allow us to know ourselves in ways never imagined.

However, the dilemma expressed at the beginning should not be disregarded: whether its application is beneficial or detrimental to the human being and how to learn to manage this technology in education. This critique of technology has invited a more holistic perspective, but with broader social and pedagogical emphasis, considering the effect that any technological application has on individuals. By facilitating the forms of interaction with the environment, it also transforms the way in which human beings interact with each other. In this context, the objective must be positive not only for human beings but also, as far as possible, for all elements of their context.

Pursuing this purpose helps to balance benefits with responsible use, preventing the emergence of inappropriate applications or exclusive benefits, aspects that, unfortunately, have been put into practice without control in recent centuries and have had a major impact on humans. This perspective constitutes the basis of social technologies. Given the expansive nature and diversity of the educational field, the determination of which technologies may be considered beneficial must be related to the established educational objectives. Moreover, as a social technology, such innovations can contribute to the construction of a better social reality, especially in contexts characterized by the absence of specialized teachers, adequate pedagogical resources, or suitable environments for receiving education.

Building on these principles, the implementation of socio-technological systems becomes an imperative, which, “implies new ways to develop and implement heterogeneous socio-technical systems (for products, processes and organization) focused on the generation of the dynamics of social and economic inclusion, democratization and sustainable development” (Thomas, et al., 2005: 587). These aspects have a major impact on the way in which individuals relate to each other, while providing opportunities for inclusion, connection and interaction between human groups of different levels and backgrounds, enabling new paradigms that strengthen and foster social bonds.

Given such a perspective, technology applied from a human-centered paradigm will seek to positively influence and serve as a solution to the most urgent needs, especially those that become mandatory for a dignified life: health, education, gender equality, renewable energy, job opportunities, nutrition and access to basic services, among others. Following this parameter, they would become social technologies whose primary objective is human welfare. It is worth mentioning that “in using the term social technologies we refer to the tools and practices that constitute our increased capacity for personal communication, production, publication, distribution and sharing” (Hagen and Robinson, 2012: 31), practices that allow a better distribution and application of technologies and “that represent solutions for social inclusion and the improving of living conditions”³ (Instituto de Tecnologia Social, 2004: 130).

It must be acknowledged in this reflection that numerous theoretical and practical efforts have been undertaken to demonstrate the educational and social benefits of effective technological applications. A large number of these impulses are noticeable in contemporary institutions, with discourses on social technology in favor of the

³ Translation by the authors of this text.

inclusion of vulnerable sectors and practices that benefit the greatest number of individuals from the different angles of human requirement, in particular in the globalized areas and their dynamics, pedagogical needs and spaces which encourage new forms of participation (Rodrigues & Barbieri, 2008; Hagen & Robertson, 2010; Nord et al., 2014). Although social technologies have come to be understood quasi as social engineering, that is, mechanisms which guide human action (Derksen et al., 2012) this is a misunderstanding that does not correspond with the objectives currently understood for social technologies, with altruistic objectives that build common welfare. Notwithstanding that there is still a lot of work to be done, encouraging reflection, learning and the practice of social technologies will help create paths for the benefit of humanity.

II- Artificial Intelligence, characteristics and social problematization

With these guidelines in mind, a new element emerges into discussion in the pedagogical and social technology discourse, a phenomenon that has had an impact in recent years spanning political, scientific, epistemic, and ontological dimensions: artificial intelligence. It can be defined as “the field devoted to building artifacts capable of displaying, in controlled, well-understood environments, and over sustained periods of time, behaviours that we consider to be intelligent, or more generally, behaviours that we take to be at the heart of what it is to have a mind” (Arkoudas & Bringsjord, 2014: 34). Analyzing this concept allows us to see several notable features: first, its artificial, deliberately constructed, inorganic nature which operates according to mechanisms established by more basic technologies. This relationship is important because it facilitates the recognition of different levels of development and application of technologies. For example, at an elementary level there is the accumulation of technologies that convert natural raw materials into components for electrical devices which, at a higher level, form the frameworks enabling software functionality, thus constituting digital technologies.

The second element is the dynamics of interaction with the environment, which allows many tasks and processes performed through the measurement and identification of spatial properties. This allows the transfer of characteristics of the real world to the digital world and, thus, to interpret and transform aspects of the material world. Finally, there is the element that simulates human cognitive processes and the differentiation between the so-called ‘weak AI’ and ‘strong AI’ (Searle, 1980). This is the most challenging and striking feature of AI, as it transcends being a mere tool for understanding our cognitive processes, because “the appropriately programmed computer really *is* a mind, in the sense that computers given the right programs can be literally said to *understand* and have other cognitive states” (Searle, 1980: 417).

With this panorama, the possibilities for the use and application of AI become extensive, opening doors that previously did not exist and inviting us to explore and understand reality in different ways, many of which are novel and challenging, yet aimed at achieving more effective improvements. This is because AI is integrated in most human activities that use digital technologies, enabling the creation of supportive resources, such as recreational or leisure experiences in virtual worlds, spanning all fields of human knowledge. Concurrently, it supports daily tasks at industrial, professional and personal level, at home, in commerce, across all domains that involve human interaction with and understanding of the world.

However, its existence opens new horizons of possibility and inquiry that no other technology had previously offered, especially from its extraordinary capacity to

emulate human rational functioning. This poses several difficulties. Is AI an intelligence that elucidates the fundamental structures and components of the brain, including its cognitive and communicational processes? If this is the case, it could serve not only as a tool for understanding but also to influence and address challenges concerning cognitive functions and the best application of knowledge processes. This outlines its potential use for promoting inclusion across diverse domains, seeking improvements in the perception of the world, understanding of the environment, social dynamics and the prospective development of humanity. The example that is now becoming a benchmark par excellence and arguably the driving force behind contemporary reflection on AI, is the chatbot launched in November 2022 by Open AI: ChatGPT.

This application has revolutionized human-technology interaction. It is a Large Language Model capable of providing conversations on a wide range of topics, enabling the understanding of dialogues and the generation of contextually appropriate responses (Deng & Lin, 2023), while working “in ways that were strikingly similar to the left-hemisphere language regions of our brains, using the same computational principles” (Waldrop, 2024: 1). This similarity to human brain function sets a standard for incorporating ChatGPT into the study and application of human cognitive activities, enhancing language learning, skill development, memory enhancement, research support and its implementations as a study tool (Bai, et al., 2023). As can be seen in this enumeration, the appropriate application of a tool such as Chat GPT would greatly facilitate learning processes and human communication, especially in the educational and pedagogical field, facilitating more efficient knowledge acquisition, comprehension of complex content, and the selection of information to be learned.

If these tools were applied to educational contexts with structural limitations (that is, emerging economies, developing regions and other places with limited educational resources), improvements could be achieved in the way teaching is carried out, making it more effective and promoting personalized learning, thereby narrowing existing educational gaps through an AI that guides and provides students with tools *ad hoc* to their needs and contexts. This results in improved efficiency in the educational field (Mhlanga, 2023), particularly in those places where it is most necessary and where current technology does not allow equitable and dignified human development. Within this framework, AI would become a social technology, functioning as a solution that promotes the inclusion and improvement of living conditions of vulnerable communities.

It is worth mentioning that the improvement of living conditions means that every individual has access to services that allow them to develop fully, with sufficient resources to maintain their health and the corresponding job opportunities to obtain those resources. In this case, improvements in living conditions are closely linked to education, since it reflects a high quality of life, allowing for an education suited to the full development of every individual, providing learning opportunities that subsequently enable them to find employment opportunities to finance a sufficient quality of life.

However, the landscape can also entail disruptions and be used in the opposite way, not for the benefit, but to the detriment of human well-being. This occurs primarily, when applied exclusively, being accessible only to a small number of people and further exacerbating disparities between collective and individual benefit. At the same time, it becomes prejudice when applied against other human beings. Even its routine use can produce adverse conditions, for example in the context of large language models in education: “some individuals are concerned that ChatGPT may

produce a generation of students who need more originality, creativity, and autonomy. This is because this tool can provide learners with the answers to their queries without requiring them to think critically and independently” (Zayoud, et al., 2023: 80). These problematizations invite us to consider the dual potential of artificial intelligence: in a beneficial or negative way for human development. When properly defined, any technology can impact human development and create spaces for inclusion and improved quality of life, turning artificial intelligence into a social technology requires a careful approach to its application, with special emphasis on ethical analysis in the educational role. Reflecting on how AI should be conceptualized as a social technology in education becomes an imperative in these times of technological breakthrough. Accordingly, this reflection will be carried out, pointing out which elements should be considered to identify the benefits and responsible application of AI.

III- Ethical proposal of AI in Education

To begin with, one aspect must be considered: Considering the proposal of artificial intelligence as a social technology in education requires an ethical analysis framework that aligns with its intended purposes. Thus, it is essential to identify principles that govern its use and avoid indiscriminate application (whose development constitutes an element of prejudice towards other people) to understand AI as a social technology in education. Floridi, in his studies on digital technologies and their application to everyday life, proposes a reflection on the normative framework used in current ethical discussions and builds an ethical framework for AI through “the four core principles commonly used in bioethics: *beneficence, nonmaleficence, autonomy, and justice*” (2023: 60). Despite the novelty of AI, it is possible to apply previously established ethical frameworks in education when considering technologies. This is due to: a) the application of AI outlines similar problems that have been present in recent decades and allows us to identify the shortcomings that still exist in the use of technology and, b) these principles function as universal prescriptions that emphasize the desire to enhance human dignity in all technological practice. Bioethics has been a branch of moral reflection that arises in the face of the wide range of new discoveries in the medical field and where scientific advances must always have human well-being as their primary concern, thus avoiding giving more consideration to the discovery than to its effects. This can also be taken as an example of current digital technologies and their application in education.

Through an analysis of the four core principles, the ethical considerations of AI in education can be appreciated. First, in relation to the principle of beneficence, Floridi argues that this principle “firmly underlines the central importance of promoting the well-being of people and the planet with AI” (2023: 61). This dual focus is particularly significant. Not only human well-being must be prioritized, but also the environment, ensuring that natural resources are preserved and not exploited indiscriminately. Aiming for a sustainable planet, it encourages the teaching of ecological practices, where minimal resource consumption is required, with an emphasis on renewable resources. The impact of this extends beyond immediate outcomes, because having a sustainable environment the benefits will also be achieved in the quality of life of the people that inhabit them. Any human activity that is governed by the principle of beneficence will encourage the application of AI, seek the well-being of every individual, and influence areas that benefit and promote a dignified life.

Since ethical education is not only the practice that aims to achieve benefits, but also the reflection that prevents any harmful action, the second principle serves as the counterbalancing principle to the principle of beneficence: nonmaleficence. Many places where social technologies are applied correspond to areas with limited resources, or where the proliferation of exploitation means that the natural resources are not renewed, mainly affecting communities that depend on these resources. Simultaneously, a misteaches practice uses technologies to harm other human beings, which can be seen in organized crime or terrorism, who use technological advances for their actions.

Establishing this principle as the basis for the design and training of AI systems would imply that its use does not allow an action that harms human beings and the environment. At the same time, since artificial intelligence allows the analysis of actions, data handling and the projection of results with the use of this information, an algorithm can be established to carry out the principle of nonmaleficence, avoiding acting upon a request that violates these regulations, and functioning as a discussion and learning tool, which makes explicit the available courses of action to choose the one that is most ethically appropriate. From this perspective, the principle of non-maleficence allows the establishment of certain algorithms that regulate the implementation of AI in human activities, preventing it from being misused (Floridi, 2023).

The third postulated principle refers to the choice that AI generates, i.e. autonomy, by allowing it to make decisions on issues of great importance. This can be understood in two dimensions: first, in relation to the autonomy of AI and its impact on human beings, particularly when AI functions as a model, guide, or interlocutor for discussion; and second, in terms of AI functioning as a facilitator of human autonomy. Regarding the first dimension, contemporary reflections consider that “the risk is that the growth in artificial autonomy may undermine the flourishing of human autonomy” (Floridi, 2023: 62). This consideration is a problem to be analyzed in the application of AI in human dynamics, necessitating that its moral constraints remain operative, addressing the two previous principles (beneficence and nonmaleficence). Therefore, it is important that AI can be created with postulates and regulations that govern its decisions in environments that benefit humans, not allowing them (despite their generative nature) to influence decisions that affect people. In this way, AI can function as a discussion guide for the students, establishing cases and at the same time learning from possible solutions.

Regarding the second dimension, despite the high level of understanding and information processing, AI would have to be applied encouraging human autonomy. This can be implemented with great benefits for learning in vulnerable and need-based communities, where access to certain resources impedes the development of their members. To the extent that AI is available to serve these needs, human autonomy can always be promoted. Here, decision-making and the selection of those with the best consequences could be facilitated through AI, thus becoming a tool for pedagogical social benefit.

This AI reflection would propose the fourth principle: justice. By mediating and having an autonomous human decision on the AI's resolutions, it is necessary to correspond with fair elections, whose balance selects the most appropriate decision corresponding to human well-being. Thereby, AI would serve as a teaching tool “in correcting past wrongs, such as by eliminating unfair discrimination, promoting diversity, and preventing the reinforcement of biases or the rise of new threats to justice” (Floridi, 2023: 63).

However, Floridi adds one more principle to the already established by bioethical reflection: the principle of explicability. This approach seeks to integrate the epistemic character and the ethical dimension of AI, ensuring that its application embodies the four aforementioned principles. As Floridi says:

In order for AI to be beneficent and non-maleficent, we must be able to understand the good or harm that it is actually doing to society, and in what ways; for AI to promote and not constrain human autonomy, our 'decision about who should decide' must be informed by knowledge of how AI would act instead of us and how to improve its performance; and for AI to be just, we must know whom to hold ethically responsible (or actually legally liable) in the event of a serious, negative outcome, which would in turn require an adequate understanding of why this outcome arose, and how it could be prevented or minimized in the future (2023: 63).

The principle of explainability invites us to consider that the ethical use of AI also implies understanding of how it can be applied to human activities and assessing how its use should be regulated, promoting a more suitable path with the human goals of common development. This is the aim of Education. Taken together, the application of these principles, creating an AI ethics, is the first step towards valuing it as a social technology in pedagogical fields.

An example can help to explain the application of these principles in education. In Guatemala, a common phenomenon affecting children's education is that 60% of schools are multi-grade (OEI, 2023). This means that one person teaches different grade levels, which often share the same classroom. This phenomenon occurs primarily in rural areas, where the scarcity of educational resources, the lack of students, and the shortage of teachers force schools to combine primary school children in a single classroom. This is detrimental to their development, as teachers must cover shared subjects and content that is not tailored to the individual needs.

Given this phenomenon, the use of artificial intelligence in multi-grade classrooms is being considered, where teachers act as guides who focus exclusively on supporting students, while the students, using a tablet or personal computer, develop various topics with the assistance of artificial intelligence. However, caution must be exercised with the use of AI, especially in a context where the very needs of the students can lead to misuse. From this perspective, it is imperative that the application of artificial intelligence be guided by ethical principles, enabling its use as a social technology. In this example, the principle of beneficence is evident from the very application of AI, where each student can receive personalized support, while the teacher is assisted in developing the content. This demonstrates an initial benefit to the educational community. However, strict adherence to established guidelines for the use and application of artificial intelligence is essential to ensure the principle of non-maleficence. Establishing regulations that students understand can facilitate the effective implementation of this principle.

Since the very interaction with AI fosters in the student a capacity for independent learning, this connects to the principle of autonomy in the two dimensions mentioned: AI influences the process by presenting content, and the student develops the ability to relate to and discuss with their digital interlocutor. The principle of justice is part of the exercise that the teacher encourages in the classroom. This is done through discussion about the evaluation of the content and its application for the benefit of human beings in the classroom context. This serves as motivation for the student to learn that the information they obtain from AI

should always be examined and considered. Here, the final principle, explainability, is introduced. This entire exercise must be accompanied by reflection on why AI is used in a multi-grade school, what its benefits are, and that it is a technological tool used for social development. With this perspective in mind, students will be able to practice ethical reflection while learning and interacting with AI.

IV- Artificial intelligence as social technology in Education

The next step considers some aspects of Floridi's proposal, taking it as a basis but expanding on the perspectives that can be seen in these arguments. The principle of explicability lies at the core of Floridi's work, encouraging the development and use of knowledge about AI and motivating individuals to understand how AI systems operate. In addition, it is intended to improve its functioning, knowing how to act ethically against its misuse. However, this leaves aside an important reflection. One of the problems that technology has shown throughout history is that, often, its development does not follow a determined rational pattern, according to ethical objectives, but rather it provides new possibilities and opens paths, which were previously not considered. This constitutes the surprise factor that occurs in all technological and scientific advances (French & Murphy, 2023).

Given the surprise factor of AI, the possibilities of directing progress in new lines of research acquire an incomparable force. This novel status has repercussions on the ethical and pedagogical level, since novelty can become a threat to the principle of beneficence, ignoring the principle of human autonomy and making the avant-garde practice also contradictory to the principle of nonmaleficence. A rather alarming example today is that many jobs are starting to be automated in such a fashion that AI is introduced into many processes (Poba-Nzaou et al., 2021), which constitutes a critique of how functional human-skills are learned to develop these activities. This becomes a threat to educational institutions and its objectives, because humans learning skills and ability to work lose value when faced with an AI that performs the same tasks in a more effective way (less time and resources) and without the contractual obligations that every worker must have.

By applying Floridi's reflection, AI ethics, governed by the five principles mentioned and applied in Education, presents dilemmas that must be resolved with greater consideration for the human and not only for the benefit of technological advancement. AI, therefore, invites us to change the way technology was previously thought and considered, emphasizing that its interaction with humans is presented not only as a benefit, but as a threat. Preventing the latter, while maintaining ethical principles, will be achieved to the extent that the development and use of AI is not an exclusive response to its own thoughtless and arbitrary progress, but is primarily aimed at solving the most important needs of human beings. This promotes a change in the way AI is thought, no longer as a simple cutting-edge technology, but as a social technology.

However, there is another aspect to contemplate in this reflection: With AI, to whom do ethical principles apply? As mentioned, AI becomes a privilege. Its effects are partially given, especially in large economies and groups with easier access to new technologies. Social gaps are a notorious sign that beneficence, nonmaleficence, autonomy and justice are not equitable. This must be changed from the basic level of education and the teaching of AI. If appropriate measures are not taken, the technological advance of AI will become a huge tool against human dignity that will affect, mainly, the most vulnerable. To do this, it is important to consider AI as a social technology and work to strengthen this conceptualization, practice and its constitution. Three steps are proposed below to build the paradigm of AI as a social

technology in Education: a) universal access to AI, b) the identification and resolution of needs in vulnerable communities, and c) AI becoming a product of the same groups.

Throughout history, not all technologies have always been accessible to all people. While in the past there were external elements that affected access, such as cultural context and distance, the ease of globalization today allows these possibilities to be narrowed and technologies to be made accessible. This invites several considerations related to teaching AI: first, access to computerized tools, from micro levels (such as personal cell phones) to macro levels (such as specialized industrial machines with global impact). A second consideration involves building a digital infrastructure that enables communication on a global scale, seeking to reach remote locations and places of urgent need. It is worth mentioning that, many times, the lack of access is not due to geographical or resource difficulties, but to political will. An aspect that, given this change of perspective and consideration of AI as a social technology, must also be transformed from school institutions, making it a priority for every State to provide the facilities for its implementation and teaching.

By facilitating universal access to AI, many people would be given the application tools that improve their quality of life and the ways in which they interact, both at the individual level, and as community or state organizations, thus helping in education, health, commerce, leisure and the various areas of daily life of groups with more material needs. But this is where the use of AI must be based on education that teaches how to use this tool ethically. Therefore, considering AI as a social technology implies that its application must be governed by morality and simultaneously put the aforementioned principles into practice. The use of AI should not be indiscriminate but accompanied by analysis and criticism at its points of application. This can even be carried out with the same AI tools, teaching how to use AI, from the outset, to enable viable solutions in the necessary contexts.

The application of AI as an educational social transformer, bringing proposals for solutions to community needs, will strengthen the links of interaction with human beings through the implementation of good practices, which can be taken as an example and be replicated. This allows us to proceed to the third stage, in which AI becomes a product controlled by the same groups. This will be a massive step in education, encouraging the creation of new AI tools according to community needs. It is worth mentioning that “social technology involves the construction of solutions collectively by those who will benefit from these solutions and who act autonomously, that is, they are not just users of imported solutions or produced by specialized teams, as is the case with many proposals from different branches of appropriate technology” (Rodrigues & Barbieri, 2008: 1075). This allows us to see how individuals involved in the dynamics of applying AI as a social technology will also become primary actors in community transformation, directly influencing social change and responses to their most urgent needs.

This demonstrates how incorporating AI into everyday life dynamics implies more effective learning and interaction that recognizes the value of both the benefit and the threat that AI poses. By highlighting this, more appropriate methods for its learning and application can be proposed, especially those whose objectives seek social transformation. Thus, AI proves to be a suitable technology for the purposes and tasks of social technologies, and, insofar as it is thought, functions as a tool for identifying needs while becoming a product of disadvantaged groups, thereby making it another pedagogical element for the positive transformation of society.

Final considerations

Since social technology seeks inclusion, implementation of benefits and valorization of human beings, especially those most in need, the dynamical teaching and appropriate learning of AI can act as a social technology that manages solutions to global problems in education, emphasizing that its application in disadvantaged contexts can bring many benefits to the extent that it is implemented well. Although AI closely resembles human behavior and thinking, this resemblance should not be misleading, leading to the assignment of tasks that require deeper ethical reflection and that must always place human beings at the center. If this is not taken into consideration, its application could be detrimental to human beings from the most basic levels.

Social technologies seek to develop new systems that favor social dynamics and when technology is applied within a paradigm that primarily teaches consideration of the human being, it will not only influence positive practices in the interaction of individuals with themselves and their environment but also serve as a pedagogical solution to the most urgent needs. All this outlines the role of teaching AI as a tool for social transformations and to achieve a dignified life in every human sphere, especially in classroom dynamics, allowing new possibilities for learning and teaching, benefiting both students in the apprehension of content and teachers in giving their classes.

When considering AI as a social technology in education, it is necessary to remember that the approach of its application emphasizes the framework of ethical analysis needed, in this case, the principles of beneficence, not maleficence, equity and justice. Principles that promote paths to maintain human dignity from nursery levels. If these become universal prescriptions, they will succeed in establishing human dignity in every AI technological learning, and in its future development, ensuring that human learning at the school level can be reflected in social dynamics.

For this purpose, human reflection must always be included in accordance with ethical objectives, avoiding that progress, with its wide range of possibilities, leads the development and application of AI into arbitrary and harmful fields for human beings. The three steps outlined above are therefore elements that make AI a social technology in education: a) universal access to AI, b) identification and resolution of needs in vulnerable communities, and c) AI becoming a product of these same groups. Due to its artificial and inorganically constructed nature, which functions as a consequence of more basic technologies, AI can be analyzed from the same community level, with greater emphasis on the identification and resolution of social needs. It is now necessary to promote this change of paradigm and emphasize that, prior to any technological progress, the paramount importance of solving the most elementary problems of the human being must not be overlooked, thereby enabling the promotion of a dignified life for humanity in all levels and spheres.

References

- Arkoudas, K. and Bringsjord, S. (2014) Philosophical foundations. In: Frankish, K. and Ramsey, W. M. (Eds.). *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, pp. 34-63.
- Bai, L., Liu, X. and Su, J. (2023). ChatGPT: The cognitive effects on learning and memory, *Brain-X*, 1(3). DOI:10.1002/brx2.30.

- Castany Prado, B. (2012). “Perdida toda coherencia”: el descubrimiento de América en la “crisis de la conciencia europea”, *Anales de Literatura Hispanoamericana*, 41, pp. 19-44. DOI:10.5209/revALHI.2012.v41.40290.
- Deng, J. and Lin, Y. (2023). The Benefits and Challenges of ChatGPT: An Overview, *Frontiers in Computing and Intelligent Systems*, 2(2), pp. 81-83. DOI:10.54097/fcis.v2i2.4465.
- Derksen, M., Vikkelsø, S. and Beaulieu, A. (2012). Social technologies: Cross-disciplinary reflections on technologies in and from the social sciences. *Theory & Psychology*, 22(2), 139-147. DOI:10.1177/0959354311427593.
- Ede, A. (2019). *Technology and Society. A World History*. Cambridge: Cambridge University Press.
- Floridi, L. (2023). *The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities*. Oxford: Oxford University Press.
- French, S. and Murphy, A. (2021). “The Value of Surprise in Science”, *Erkenntnis*, 88(4), pp. 1447-1466. DOI:10.1007/s10670-021-00410-z.
- Hagen, P. and Robertson, T. (2010). Social technologies: challenges and opportunities for participation. In: Association for Computing Machinery *Proceedings of the 11th Biennial Participatory Design Conference (PDC '10)*. New York, NY, pp. 31-40. DOI:10.1145/1900441.1900447.
- Instituto de Tecnologia Social. (2004). Reflexões sobre a construção do conceito de tecnologia social. In: Lassance Jr AE, Claiton JM, Siqueira Barbosa EJ, et al. (eds) *Tecnologia social: uma estratégia para o desenvolvimento*. Rio de Janeiro: Fundação Banco do Brasil, pp. 117-134.
- Mhlanga, D. (2023). ChatGPT in Education: Exploring Opportunities for Emerging Economies to Improve Education with ChatGPT, *SSRN Electronic Journal* [Preprint]. DOI:10.2139/ssrn.4355758.
- Nickel, P.J., Franssen, M. and Kroes, P. (2010). Can we make sense of the notion of trustworthy technology?, *Knowledge, Technology & Policy*, 23(3-4), pp. 429-444. DOI:10.1007/s12130-010-9124-6.
- Nord, J. H., Paliszkievicz, J. and Koohang, A. (2014). Using Social Technologies for Competitive Advantage: Impact on Organizations and Higher Education, *Journal of Computer Information Systems*, 55(1), pp. 92-104. DOI: 10.1080/08874417.2014.11645744.
- OEI. (2024). *Memoria 2023-2024*. Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura.
- Poba-Nzaou, P., Galani, M., Uwizeyemungu, S. and Ceric, A. (2021). The impacts of artificial intelligence (AI) on jobs: an industry perspective, *Strategic HR Review*, 20(2), pp. 60-65. DOI:10.1108/shr-01-2021-0003.
- Rodrigues, I. and Barbieri, J. C. (2008). A emergência da tecnologia social: revisitando o movimento da tecnologia apropriada como estratégia de desenvolvimento sustentável, *Revista De Administração Pública*, 42(6), pp. 1069-1094. DOI:10.1590/s0034-76122008000600003.
- Searle, J. R. (1980). Minds, brains, and programs, *Behavioral and Brain Sciences*, 3(3), pp. 417-424. DOI:10.1017/S0140525X00005756.
- Thomas, H., Fressoli, M. and Becerra, L. (2012). Science and technology policy and social ex/inclusion: Analyzing opportunities and constraints in Brazil and

Argentina, *Science and Public Policy*, 39(5), pp. 579-591.
DOI:10.1093/scipol/scso65.

Verbeek, P. P. (2008). Morality in Design: Design Ethics and the Morality of Technological Artifacts. In: Kroes P, Vermaas PE, Light, A et al. (eds) *Philosophy and Design*. Springer: Dordrecht, pp. 91-103. DOI:10.1007/978-1-4020-6591-0_7.

Waldrop, M. M. (2024). Can ChatGPT help researchers understand how the human brain handles language?, *Proceedings of the National Academy of Sciences*, 121(25). DOI:10.1073/pnas.2410196121.

Zayoud, M., Oueida, S., Awad, P. and Ionescu, S. (2023). Impact of Chat GPT on Education: Challenges and Opportunities, *International Conference of Management and Industrial Engineering*, 11, pp. 75-82. DOI: 10.56177/11icmie2023.18.