4

The Ethical and Epistemic Impact of Artificial Intelligence in Education

Inmaculada Perdomo Reyes
Universidad de La Laguna, Spain
mperdomo@ull.edu.es
https://orcid.org/0000-0003-4838-7278

Fulgencio Sánchez-Vera
Universidad de la Laguna, Spain
fsanchev@ull.edu.es
https://orcid.org/0000-0002-0697-9120

Betty Estévez Cedeño
Universidad de la Laguna, Spain
bestevec@ull.edu.es
https://orcid.org/0000-0002-1590-952X

Abstract
This chapter examines the impact of Artificial Intelligence (AI) on education and creative work, highlighting the need for interdisciplinary analysis in the context of the interrelationship between technoscience and society. The ethical challenges of AI, such as algorithmic biases and epistemic injustice, and its influence on educational and decision-making processes are discussed. The text emphasizes the importance of regulation and accountability in the use of AI to preserve democratic values and fundamental rights, while also reflecting on the future of education in the age of AI.

Keywords: Artificial Intelligence, education, society, ethical challenges, epistemic injustice

4.1. Theoretical Approach: Technosciences and Society

In science, as in any other human cultural practice, values and interests are present in different forms and degrees at different stages of the scientific process. Moreover, scientific knowledge is the result of controlled research, thus experimentation and interpretation of data must satisfy the minimum level of empirical adequacy, although these are issues subject to contextual value-based decisions. Therefore, from the formulation of the questions that point out a research problem and the objectives set (cognitive and pragmatic or application), the selection of the most appropriate methodologies to address it, the processes of extraction, selection and interpretation of data, the delineation of hypotheses, the texture of the inferences, the results and even the public communication of the latter imply value-based judgments. The knowledge capitalization era (Echeverría, 2003) requires a more comprehensive interdisciplinary effort that enables the analysis of all the relevant aspects involved, focusing on the inherent and shaping values of technoscientific practice itself.

The term "technoscience", originally proposed by Bruno Latour in 1983, refers to the fusion of science, technology, industry and the military (Echeverría, 2003), as well as the hybridization of this complex and society. These developments are accompanied by radical changes in the ontological premises of the technosciences, as well as in some of their rhetorical and political strategies. From these multifaceted changes come new epistemologies and methodologies that emphasize the constructionist character of categories such as science, technology, and society. Science, Technology and Society (STS) theoretical approaches pay attention to these new challenges from strategies and concepts that capture the reality of this new organization of scientific-technological practice. These are "socio-technical systems, hybrid systems involving individual people, but also corporate actors such as companies and government bodies, as well as more abstract social entities such as institutions, laws and regulations, and other rules" (Franssen and Kroes 2009, p. 223). Proposals such as that of Jasanoff (2004, 2016) invites to modify the ap-

proach to capture the emerging characteristics of such hybridization: the configuration of *socio-technical imaginaries*, linked to the concept of co-production of science, technology and society. Unveiling these socio-technical imaginaries helps to explain why some scientific and social visions tend to gain more support and authority, and why some develop at the expense of others.

The analysis of the impacts of AI systems requires a critical analysis. Here, interdisciplinarity is absolutely necessary: technical and legal, philosophical, political, ethical and educational aspects are intertwined in the social phenomenon of AI. It is a major challenge for which the theoretical approach of STS Studies of co-production or mutual conformation of technoscience-society is proving to be fruitful (Wajcman, 2023). A diversity of actors or social agents transit and interact with the technoscientific system. They exercise their capacity to co-produce (unexpected or unforeseen uses of certain technologies, resignification, transformative symbolic uses, modification of values and behaviors, dissident narratives...) technology and its meanings. In this system, it is necessary to introduce the gender perspective in an intersectional key to address complexity in an effective and fruitful way and to develop research on this continuous dynamic.

4.2. From Expert Systems to Dataism and Epistemic Injustice of AI

AIs with autonomous learning based on neural networks are capable of making correlations and inferences from the millions of data they use (databases, images, results of human interactions with smartphones, etc.), reflecting, as if it were a mirror, the accumulation of human traits and characteristics. The Latin word *datum*, which comes from *dare* (to give), literally means *given*. Knowledge in the information regime strives to achieve total knowledge through algorithmic operation, substituting the narrative for the numerical. Dataism aims to calculate everything that is and will be (Byung-Chul Han, 2022, p. 21).

On the other hand, they also reflect the biases, prejudices and stereotypes that continue to structure our societies. The *algorithmic bias* acquired through data are multiple and varied. These are

related to the very nature of databases that do not actually represent the entire population (Criado, 2019). Generative AIs reify and circulate existing gaps and biases, but they give them a veneer of objectivity and neutrality despite the opacity of most of these processes. They can be true automated mechanisms of reproduction and generation of inequalities and exclusion (Eubanks, 2018). Technologies are both a reflection and a crystallization of social processes. Even so, there are still few studies focused on how power and gender relations end up integrated into technoscience, from the design to the setting of pragmatic goals, but it is widely recognized (Wajcman & Young, 2023) that both the workforce and the dominant cultures in technoscience are clearly representative of the groups of people that make up it, and that women, for example, do not represent more than 17-19% (Young, Wajcman, & Sprejer, 2023). In this way, AI as a disruptive technology represents a biased prolongation of the knowledge production model. Such a situation is of clear algorithmic and epistemic injustice and faces major challenges in education, at work and, especially, in public decision-making processes and in the generation and transfer of knowledge.

Algorithmic injustice is reflected by not considering the epistemic contribution of vulnerable voices or groups, underestimating the importance of contextual categories, both material and immaterial (Abdilla, 2021) and of corporeal entities or agential realism (Barad, 2007) of minorities, which are relegated to the products of the large AI technology industries (Catá, 2023), all of which are examples of the Western capitalist model. Therefore, the production of technoscientific knowledge is subject to algorithmization. In other words, the sociocultural processes, data, and institutions that, in natural language, we may -or may not- recognize as authoritative agencies of knowledge become algorithmic models of the globalized world.

New AI-mediated narratives construct hegemonic algorithmic cultures (Striphas, 2015 cited in Ricaurte, 2022). When the 'XCheck' program created by Facebook and initially designed a mechanism to review in more detail the measures taken against high-profile accounts, it ended up becoming a system that rather exonerated numerous celebrities, politicians or journalists from complying with the rules that are imposed on other users. This shows that, when an algorithmic sociocultural model is pro-

duced, there is a high probability of prolonging or generating a new version of capitalism, colonial and patriarchal (Ricaurte, 2022). Therefore, the role of ethics or the development of responsible AI (Torrones, 2020) becomes a core issue.

The work situation experienced in 2018 by Timnit Gebru and Margaret Mitchell,[1] both pioneering AI ethics researchers and former Google employees, suggests that, in a world imbued with bias, algorithmization enhances existing differences. Thus, an ethic code for the majority of the world (Ricaurte, 2022) in times of AI must be based on the recognition of deliberation as a cornerstone of technoscientific development and as a necessary dose of humanity in the construction of the new technology.

4.3. Generative AIs and Impacts in the World of Creative Works and Education

Nowadays, machines cross (or replace) the immaterial, cognitive and cultural work, the communicative activity of society and the deliberative capacity of citizenship. They learn by themselves, that is, it is no longer necessary to instruct them. They can find rules, correlations between x and y, not even foreseen by humans (which means significant advances in disciplines such as Medicine, when it comes to accurately diagnosing a disease from multiple and varied symptoms). However, machines do not know why this is the case. It is still a profoundly human task to give an account of the why, to clarify the frames of reference, the prior assumptions, the values and the objectives set. Although assisted by AI systems, human beings can increase the accuracy and speed of responses to decision-making but cannot be replaced by machines.

In addition, delegating to AI systems decision-making processes such as the suitability of candidates for a job position, ac-

1. In 2020, Gebru was terminated from her job because she refused to retract the findings published in an academic paper after her superiors requested it. The referred paper explained the weaknesses of facial recognition and evidenced a 34% error margin in recognizing black women. More information in Pérez, 2020, and Hao, 2020. Mitchell was terminated months later. She is considered one of the leading experts in ethics applied to technology and one of the 100 most influential people of 2023 (Catá, 2023).

cessing life insurance or mortgages, the resolution of administrative applications, etc., offers public managers the necessary ethical distance to make decisions that are increasing the vulnerability, inequality and exclusion of the more disadvantaged people. Whether or not they belong to the statistically relevant, or reference, profiles identified by AI models, algorithms and systems, can mean the difference between being eligible or not to be a beneficiary of all the resources that a society puts into circulation to facilitate the lives of citizens. The line between classification based on algorithmic calculation (apparently objective) and the assessment of the specific conditions of cases that require human deliberation, reflection, rationality and communication is not so thin if it involves increasing the vulnerability and social exclusion of these people.

Of particular concern is the advance of generative AIs in those areas that reflect human skills, rationality and creativity: artistic and creative works, those requiring conceptual analysis and critical thinking, and those of knowledge transfer and shaping, as well as the teaching and learning process of new generations.

4.4. From Externalized Memory to Fractured Thinking

It would not be wrong to say that we have left behind the era of knowledge embodied in books with structured discourses and careful arguments. Now, we are in the fragmentation of thought. Rationality also requires time. That time no longer exists in the face of the continuous acceleration of processes. AI does not reason, it computes. Arguments can be improved and are the basis of continuous learning, especially in the formative stages; however, if we replace them with algorithms, even if they are optimized, this leads us to abandon the argumentative and reflective effort. This is, at the same time, an abandonment of thinking. We cannot process the enormous amount of information, nor fight against the speed of tweets, memes or the attractiveness of images that impact our brain, turning it into a true addiction to the continuous consumption of visual stimuli that impact our emotionality instantaneously.

> Big data and Artificial Intelligence enable the information regime to influence our behavior at a level that lies below the threshold of consciousness. The information regime takes hold of those pre-reflexive, instinctual, emotive layers of behavior that precede conscious actions. Its data-driven psychopolitics intervenes in our behavior without us being aware of it. (Byung-Chul Han, 2022, p. 15)

However, we are forced to think about and implement, in the teaching-learning process, new strategies that make use of the best of generative AI systems: especially those that process natural language and image creation, among others. We must incorporate generative AI systems and, at the same time, stimulate and enhance reflection, rational argumentation and meaningful learning. Rational-argumentative and meaningful learning are the basis for growth and development in the personal growth of the younger generations. This is essential for a mature and deliberative citizenship. Moreover, it is the basis for a well-functioning democratic society. Hannah Arendt was already pessimistic in the 1990s, but we are still obliged to make this effort.

> The effort of knowledge and perception is replaced with the business of distraction. The consequence is a rapid decline in human judgment. There is an unmistakable threat in it: it either makes the public immature or keeps them immature, and it touches on the social basis of democracy. We had fun until we died. (Hannah Arendt, 1996, p. 342, self-translation)

Education is constantly evolving, and the direct influence of technological advances at all levels has long been watched with concern. The widespread use of the Internet facilitated many tasks of searching for information, allowed communication in new formats and put into circulation all kinds of resources that have been changing the scenario and the patterns of the teaching-learning process. Adapting and incorporating new methodologies and evaluation systems has been a constant in recent years. Nevertheless, it is undeniable that the impact of AI systems on education poses a new challenge, especially Natural Language Generation models such as the one developed by Open AI (ChatGPT), and others like it. These systems can generate study materials, offer conceptual precisions when required,

bodies of structured information, explanations on a topic, solutions to mathematical problems, and so on. From this point of view, they can become learning partners, tutors who guide the process of self-learning or help in educational management.[2] At the same time, we easily delegate to this types of tools the effort of writing well-structured texts with personal involvement; we subrogate the act of formal teaching, which is considered an act of social interaction essential in the development of the human being and with a high load of emotional capital (Tarabini, 2020) provided by the presence of the teacher, particularly in the first years of life.

On the other hand, teachers can also generate materials and obtain answers to numerous questions about the best way to organize their educational and assessment practice, among other tasks. There is the option to perform an automated assessment and to know in real time the performance and progress of the students. The question immediately arises as to the authenticity of the learning and assessment process: does this develop an adequate level of critical, argumentative and communicative skills? Moreover, do we learn to differentiate correct, truthful or relevant information from that which is not? And do they develop the capacity for independent thinking or do they get used to quick and limited answers to save the delivery of a paper or the preparation of an exam or assignment? In addition, the use of these tools produces a loss of the value of authorship, as they do not consider the citation process relevant and the concept of plagiarism is relativized.[3]

As teachers, we are also concerned about the loss of the ability to concentrate, loss of the ability to argue and to make well-founded judgments, loss of creative, deliberative and communicative skills. We observe with concern the fragmentation of attention and thought, which are fundamental values for becoming

2. In the field of management, the Lola (University of Murcia, 2028) or ADA cases of (University of Jaén, 2021) are pioneering examples that have implemented AI-based virtual assistants. With a chatbot, they have provided information for new students and have helped to resolve doubts about degrees, credits and procedures.

3. In a recent study on dishonest practices among students in the first year of the Bachelor's Degree in Primary Education, 40% of students perceive that the practice of plagiarism in academic tasks is common despite the fact that they also recognize that if it is proven has serious consequences for their academic performance (Sánchez-Vera et al., 2023).

good professionals in the future and having an adequate personal development. Acting ethically means being able to take responsibility for our judgments and actions and their consequences, something that machines cannot do, no matter how sophisticated they become.

4.5. Regulation of Artificial Intelligence and the Future of Democracies

The conflicts of values and the problem of many things (*numerous interconnected scientific and technical elements*) and too many hands (*many* agents, with different objectives and values, at different levels), such as AI systems (Coeckelbergh, 2021, p. 98-99), challenge the issue of responsibility and reliability in a techno-science whose problems of opacity, lack of transparency, and explainability demand the necessary limitation through regulation. The EU has recognized this for years when it formed the specific commission that developed the AI White Paper published in 2020. It states: "Given the major impact that AI can have on our society and the need to build trust, it is vital that European AI is grounded in our values and fundamental rights such as human dignity and privacy protection" (p.2). Moreover, the requirements for trustworthy AI are seven: human action and oversight; technical soundness and safety; privacy and data management; transparency; diversity, non-discrimination and equity; social and environmental well-being and accountability. With a risk-based approach that grads these systems from unacceptable risk, high risk, limited risk and minimal risk, on December 9, 2023, an agreement was reached between the Member States and the European Parliament and, although it must be ratified, the text defines the obligations and rules by which this technology must be governed. Europe will only allow the use of facial recognition and biometric control systems in special cases and with judicial authorization. Regarding generative or foundational models such as ChatGPT, these will have to meet transparency criteria and it will be necessary to clarify whether a text, an image, or a song has been generated by AI. In addition, it must be ensured that the data that have been used respects copyright. The

law will not be available for at least three years, but the European initiative undoubtedly sets limits to these developments so as not to jeopardize the rights and values that we consider central to democratic societies.

4.6. Artificial Intelligence and the Future of Education

The expansion of digital technologies in the last decades and recent advances in generative Artificial Intelligence are shaking traditional pedagogical models. The traditional educational approach, rooted in books and structured arguments, is compromised in the face of new scenarios of fragmented and accelerated thinking due to the efficiency of digital technologies and generative AI.

We should carefully consider the role of algorithms in educational processes, as AI computes and optimizes, but humans deliberate and reason. AI systems can give logical arguments, but they do not possess the faculty of reasoning in the human sense (Larson, 2022). In education, the opportunity to integrate AI is yet to be determined, that is, it is necessary to establish what, how and when to use technology, otherwise we can fall into dependence on algorithms and affect the critical, argumentative and communicative skills of students. The ability to reason and argue, which is essential for meaningful learning, cannot be relegated to the background by being replaced with technological assistants that, in pursuit of efficiency, strip us of fundamental skills.

In this scenario, we must be aware of the risks of over-relying on automated technology in educational contexts. The promise that technology will provide a more accessible and personalized education, with fair and efficient assessments, and data analysis for informed decisions, is undoubtedly attractive. However, the challenges posed by AI in terms of digital divide, data privacy and security, lack of transparency, and the incorporation of biases into algorithms, require careful thought and action. In addition, excessive automation could lead to a loss of human interaction essential for social and emotional learning (Sánchez-Vera,

2022). At this point, the role and training of educators becomes essential to manage integration and change.

In a world where AI is redefining the educational landscape, we must ensure that the democratization and personalization of learning, facilitated by technology, are balanced with the quality and relevance of the education provided. In this sense, following Amartya Sen's view of development as freedom (Sen, 1999), we must underline the importance of empowering people for active and meaningful participation in society, and this implies fostering critical thinking and ethical understanding in students. It is about offering a comprehensive and holistic education that not only focuses on technical skills, but also on the development of broader human capabilities.

Furthermore, considering the impact of technology on the public sphere (Habermas, 1984), it is essential to reflect on the responsibility of algorithmic systems in social and political values. AI has the potential not only to form opinions and wills, but also to shape interactions in a democratic environment. Habermas, while not specifically focusing on AI, provides a useful framework for considering how technology affects public discourse and opinion formation. The emphasis of this author on rational and deliberative communication in the public sphere can be a valuable reference point for evaluating and guiding the development of educational technologies that respect and promote democracy and citizen participation.

Looking to the future, it must be ensured that technology not only advances in terms of efficiency, but also contributes to the development of an informed, critical and ethically engaged citizenry. The incorporation of AI in the field of education represents not only a technological evolution, but also a challenge of political, philosophical and social nature, which demands a multidisciplinary approach to ensure that its application reinforces democratic values and promotes and ensures equitable social development.

4.7. Conclusion

The challenge we face with AI, particularly in education, is served. However, we must not lose sight of the fact that this is

only a tool at the mercy of decision-makers in this or any other area. As at other times in history, when technological evolution is generated, our duty is to remember that adaptation and its use must be the product of deep public deliberation.

Knowledge of reality should not be subrogated to the prevailing algorithm. However, we must humbly acknowledge the efficiency, speed and ability to manage huge amounts of data possessed by digital technologies, intelligent systems, chatbots or any other tool based on Artificial Intelligence. In the words of Duede (2023), we may be dealing with an instrumental epistemology, simply different from the epistemology of experts, but reality should not be reduced to one or the other. In this sense, these are two different categories and perhaps valid for obtaining in-depth knowledge, a key aspect of the teaching and learning process. But there are more categories that ethics reminds us not to forget: minority groups, vulnerable people, and contexts are also fundamental categories of knowledge that cannot and should not be forgotten in the complex reality that permeates AI.

Acknowledgement

This work is fully collaborative and is part of the research project "Vulnerabilidad, precariedad y brechas sociales. ¿Hacia una redefinición de los derechos fundamentales?", PID2020-114718RB-I00, funded by MICIU/AEI/10.13039/501100011033

References

Arendt, H. (1996). *Verdad y política. Entre el pasado y el futuro. Ocho ejercicios sobre la reflexión política*. Península.

Barad, K. (2007). *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University.

Byung-Chul Han (2022). *Infocracy*. Polity.

Catà Figuls, J. (2023, 27 November). Margaret Mitchell: "Las personas a las que más perjudica la inteligencia artificial no deciden sobre su regulación". *El País*, https://elpais.com/tecnologia/2023-11-27/margaret-mitchell-las-personas-a-las-que-mas-perjudica-la-inteligencia-artificial-no-deciden-sobre-su-regulacion.html?ssm=whatsapp_CC

Coeckelbergh, M. (2021). *Ética de la inteligencia artificial*. Cátedra.
Criado Pérez, C. (2019). *Invisible Women: Exposing Data Bias in a World Designed for Men*. Chatto & Windus.
Duede, E. (2022). Instruments, agents, and Artificial Intelligence: novel epistemic categories of reliability. *Synthese*, *200*, 491. doi.org/10.1007/s11229-022-03975-6
Echeverría, J. (2003). *La revolución tecnocientífica*. FCE.
Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. St. Martin's.
European Commission (2020). *White Paper on Artificial Intelligence. A European approach to excellence and trust*. https://op.europa.eu/es/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1
Franssen, M. y Kroes, P. (2009). Sociotechnical Systems. In J. Olsen, S. Pedersen, & V. Hendricks (Eds.). *A Companion to the Philosophy of Technology*. Blackwell.
Habermas, J. (1984). *The Theory of Communicative Action*. Beacon.
Hao, K. (2020). Claves del artículo por el que Google despidió a la líder de ética de IA. *MIT Technology Review*. https://www.technologyreview.es//s/12958/claves-del-articulo-por-el-que-google-despidio-la-lider-de-etica-de-ia
Jasanoff, S. (Ed.) (2004). *States of Knowledge. The Co-production of Science and Social Order*. Routledge.
Jasanoff, S. (2016). *The Ethic of Invention. Technology and The Human Future*. W.W. Norton & Company.
Krishnan, A., Abdilla, A., Jung Moon, A., Affoso Souza, C., Adamson, C., Lach, E., Ghazal, F., Fjel, J., Taylor, J., Havens, J. C., Jayaram, M., Morrow, M., Rizk, N., Ricaurte Quijano, P., Cetin, B., Chatila, R., Dotan, R., Mhlambi, S., Jordan, S., &Rosenstock, S. (2021). *AI Decolonial Manyfesto*. https://many-festo.ai/
Larson, E. J. (2022). El mito de la Inteligencia Artificial. Por qué las máquinas no pueden pensar como nosotros lo hacemos. Shackleton.
Pérez Colomé, J. (2020, 13 December). Por qué el despido de una investigadora negra de Google se ha convertido en un escándalo global. *El País*. https://elpais.com/tecnologia/2020-12-12/por-que-el-despido-de-una-investigadora-negra-de-google-se-ha-convertido-en-un-escandalo-global.html?event_log=go
Ricaurte, P. (2022). Ethics for the majority world: AI and the question of violence at scale. *Media, Culture & Society*, *44*, 726-745. https://doi.org/10.1177/01634437221099612

Sánchez Vera, F., Estévez Cedeño, B., Perdomo Reyes, I., & Lozano Galván, E. (2023). Prácticas académicas deshonestas y desarrollo profesional de los estudiantes de magisterio: análisis de los aspectos emocionales y motivacionales implicados. In Santos Villegas, M. J., Alcalá del Olmo Fernández, M. J., Fernández Cerero, J., & Montenegro Rueda, M. *Desafíos educativos a través de la interdisciplinariedad en la investigación y la innovación.* Dykinson.

Sen, A. (1999). *Development as Freedom.* Oxford University.

Tarabini, A. (2020). ¿Para qué sirve la escuela? Reflexiones sociológicas en tiempos de pandemia global. *Revista de Sociología de la Educación*. https://doi.org/10.7203/RASE.13.2.17135

Terrones R, A. (2021). *Inteligencia artificial responsable. Humanismo tecnológico y ciencia cívica* (doctoral thesis). Universitat de València

Universidad de Jaén (2021, 3 June). La Universidad de Jaén, pionera en poner en marcha asistentes conversacionales de atención al usuario. *Diario Digital*. https://diariodigital.ujaen.es/institucional/la-universidad-de-jaen-pionera-en-poner-en-marcha-asistentes-conversacionales-de

Universidad de Murcia (2018, 5 July). La Universidad de Murcia presenta a Lola, un asistente de inteligencia artificial para ayudar a los nuevos alumnos. *Gabinete de Prensa de la Universidad de Murcia*. https://www.um.es/web/sala-prensa/-/la-universidad-de-murcia-presenta-a-lola-un-asistente-de-inteligencia-artificial-para-ayudar-a-los-nuevos-alumnos

Wacjman, J., & Young, E. (2023). Feminism confronts AI. In J. Browne, S. Cave, E. Drage, & Kerry McInerney (Eds.). *Feminism and AI* (pp. 47-64). Oxford University.

Young, E., Wajcman, J., & Sprejer, L. (2023). Mind the gender gap: inequalities in the emergent professions of Artificial Intelligence (AI) and data science. *New Technology, Work and Employment*, 1-24. https://doi.org/10.1111/ntwe.12278