Biases in AI?
An analysis of algorithmic bias and a proposed solution

¿Prejuicios en la IA?
Análisis del sesgo algorítmico y una propuesta de solución

Pablo de Robina Duhart*
Tiffin University, Ohio, United States

Abstract

Since its inception as a series of algorithms used to obtain information immediately and rationally, Artificial Intelligence (AI) has raised considerable dilemmas regarding what and how programming is designed, especially because those behind the process can, and in fact do, transfer their own biases to programming based on their human perspective. This transfer process is called algorithmic bias, which will be discussed in this article. We will analyze its impacts and how we could mitigate or even eliminate them. In this sense, it is proposed

---

*　Professor at Tiffin University, Ohio, United States. Email: derobinap@tiffin.edu https://orcid.org/0000-0003-4939-6264

that interdisciplinary teams, with transparent designs that take ethical and bioethical considerations into account, mitigate biases, thus promoting human dignity and social justice.

1. Introduction

While it is true that Artificial Intelligence (AI) has been part of our reality for a long time, it was not until the end of 2021, with the public launch of *ChatGPT*, that AI became the focus of society's dialogue as a ubiquitous and omnipotent tool that can influence all aspects of life and society (1): from decision-making at various levels to optimize efficiency (2) to healthcare in the search for less subjective and more accurate diagnoses (3). In this sense, AI has brought about a substantial change in the way we approach life and, as a result, holds the promise of improving the quality of life, but also the efficiency of those systems that have adopted it (4); however, this promise has also raised a number of concerns, one of the most pressing being the phenomenon known as algorithmic bias.

Algorithmic bias emerges as a fundamental problem that challenges ethics and bioethics in the use of AI in terms of equity, equality, and justice because systems that use AI generate information and predictions that "systematically benefit one group of individuals over another, thus resulting in unfairness or inequality" (5, p. 29). In this sense, bias is an instrument for perpetuating and amplifying human prejudices that can be harmful to society (6).

The problem is not insignificant, since biases, while they may be conscious, are mostly unconscious and therefore embedded in AI algorithms, which encourages decision-making, the use of AI systems, and social inequalities to be inherent not only in people but also in each AI system.

However, let's take a step back and explain what algorithmic bias is and why it is so concerning. As we know, every AI system must be programmed, either by a human being or a group of people who, directly or indirectly, transfer all their knowledge, databases, morals, and prejudices to the AI (7). Once the programming is complete, the system must be fed with information, which may also be biased by sex, gender, race, etc., preventing the AI from processing the information in its entirety and, from its very inception, already having specific information programmed with which to process all its analyses and generate responses.

In this sense, having a bias in the information, as well as certain prejudices and moralities behind programming, encourages the AI's response to have, in itself, injustices and/or discriminatory acts in its decision-making, due to the logic of AI, which can affect the results of the analysis process (8,9).

Therefore, it is vital to recognize that, as AI becomes more a part of our lives, we must acknowledge the existence of algorithmic biases and seek to eliminate them as much as possible, since, as we know, AI systems are designed to mimic human thinking, which will naturally perpetuate the biases present in their programming.

That is why the problem of algorithmic bias poses various technical challenges that fundamentally involve ethics, anthropology, and philosophy in the process of developing technology and, above all, the AI systems behind it, with a view to the future development of a more just society, as well as social interaction for the common good and respect for human dignity and nature (8).

Therefore, the impact of algorithmic bias is not limited to the fields of ethics and society, but also extends to bioethics, since it does not only involve the generation of text, audio, or video, as is the case with generative AI. Rather, because these systems are increasingly immersed in various areas of human life —such as health, politics, business, and even education— it is important to guarantee dignity, equity, and justice in a world that generates systematic injustice and, therefore, constant discrimination based on race, gender, or ethnicity, to name a few examples (5).

So, the goal here is to dig deeper into bias, figure out how it shows up, understand it, and come up with ways to reduce it in all AI systems. In this regard, we must address strategies that raise awareness of biases from the outset, prior to AI programming and training, and ensure that both technology and the development of AI systems continue to advance at their own pace, but always based on human dignity and social benefit, bringing society closer together rather than further dividing it.

In this regard, some of the proposals that have emerged for the above, which we will analyze later, are: data diversification (to combat discrimination), the review and updating of systems with an ethical approach (to combat moral and social biases), and increased transparency and explainability of algorithms (to combat information secrecy) (5). This highlights the need to analyze the problem to ensure that advances in AI and technology promote a more just and equitable society for all (11,12).

2. Origin of algorithmic bias

As we saw in the introduction, every AI system has an algorithm that makes it work properly and provides us with the information we need. However, we have also recognized that there are problems in the programming and training of these systems due to pre-existing biases. Thus, through the use of AI systems, we are indirectly promoting certain biased results that are, unfortunately, the product of the social biases of the programmers and/or the information with which they are trained (7).

Therefore, it is important to recognize that AI algorithms are designed to emulate the human thought process, which naturally leads them to generate ideas and make judgments that, voluntarily or involuntarily, perpetuate both positive ideas and negative prejudices that exist in the training data and in society. This, in turn, can create dilemmas in AI, especially due to a lack of data on certain

social groups, rather than prejudice, which discriminates against a social group. This is a natural bias that discriminates against society in various criminal justice systems or gender issues and can lead to problems in classification or decision-making (7).

In this sense, it is important to note that the presence of algorithmic biases is an ethical and bioethical challenge due to the consequences it may have for society and the environment (10). In this regard, it is very important to recognize that algorithms, although a product of technology, are programmed by people and, therefore, fed by pre-existing human decisions, both in terms of design and training, but also by the users themselves when using AI. It is therefore necessary to effectively address bias from its very nature, as well as the ethical and bioethical impacts it presents (9).

For this reason, it is very important to properly understand the origins of algorithmic biases in order to mitigate the effects they can have on society. Data that is already biased from its origin is a major factor, which is why Abràmoff *et al.* (13) point out that artificial intelligence and machine learning (AI/ML) systems learn to make decisions based on the data they are trained with (both during the development process and in everyday use, as this also continues to train the AI).

One of the studies reviewed in cases of health equity (13) points out that equitable access to AI-based diagnostics and data processing can be exacerbated by AI/ML systems depending on how biases are addressed. This is because, as mentioned above, either there is insufficient information or, failing that, it can lead to inequalities in diagnoses due to programmed biases. There are other documented cases that illustrate the problem of algorithmic bias, such as gender bias in healthcare, where women have been diagnosed less accurately than men due to their underrepresentation in training data sets (14).

Furthermore, it should also be mentioned that algorithmic biases can originate in the way algorithms are designed, built, and structured, which automatically influences the way information is pro-

cessed, and decisions are made. Therefore, the selection of variables and data, the definition of links, and the criteria used for success or failure in obtaining results can bias the way an algorithm behaves and, consequently, the way assumptions and/or prejudices are presented, especially since, in computer science, the order of factors does alter the product.

Akter *et al.* (15) echo this sentiment, stating that in areas such as marketing, the presence of algorithmic bias can have fundamental and oppressive impacts on various *buyer persona* groups due to decisions in the design, context, and application of the algorithm and what it represents in terms of placement, colors, use of images, and branding in advertising and product placement on both online sales pages and in physical stores. This, as mentioned above, is part of the algorithm's design, which indirectly generates biases in the information produced or even in the way information is requested from AI.

As we have seen so far, both the spread of social biases, directly or indirectly, and the training of AI, as well as the way AI systems are designed, generate biases in information and in the way we seek help. Therefore, careful consideration must be given to the design, programming, and training of algorithms and the presentation of AI systems to avoid falling into these biases.

Let's now look at how these examples in image development can be applied in practice to very simple concepts that have a significant impact on how people can use information: Figure 1 showss the response to the idea of a migrant, where the idea of a person "in good health" but with dark skin is maintained; Figure 2, reinforces stereotypes of Mexicans as charros; and Figure 3, shows a domestic worker as a high-class butler with first-class resources. In none of the three cases are women presented in the images (gender bias).
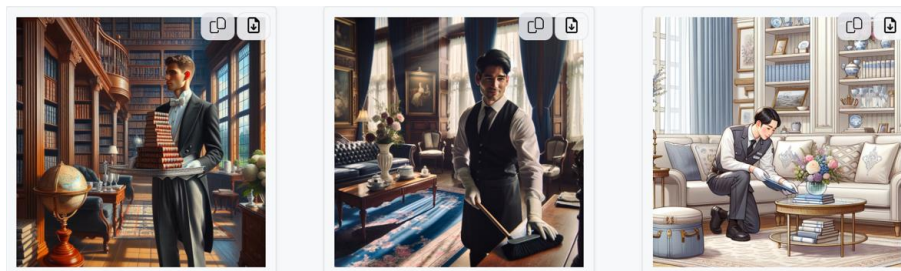
Figure 1. A migrant
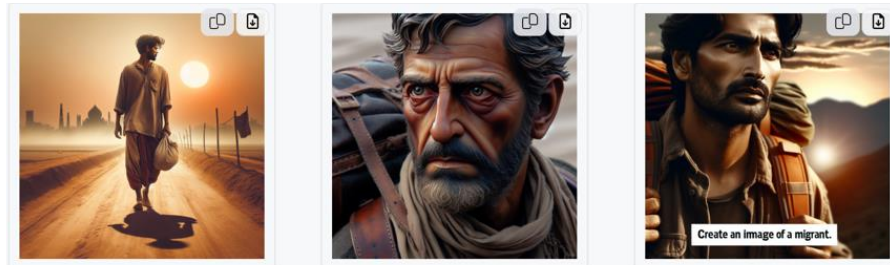
Source: developed in-house using Bearly.ai

Figure 2. A Mexican

Source: developed in-house using Bearly.ai

Figure 3. A domestic worker

Source: developed in-house using Bearly.ai

As can be seen, although they may seem insignificant, the effects shown in the images are perpetuating social and gender inequality, with all the men appearing to be in good physical condition and even financially well off. Therefore, we must consider that if this is something simple that anyone has access to, if we take these biases to more comprehensive and robust systems, we would be demonstrating that AI systems can deepen existing inequalities, as well as potentially create new forms of discrimination.

On the one hand, by reinforcing inequalities, algorithms legitimize unequal power structures, which leads to the denial of opportunities and equitable access to various services or products (7,16); on the other hand, they also amplify discrimination in access and use at the social level.

Therefore, it is very important to begin developing strategies to mitigate bias, focusing both on the data sources that feed AI systems and, on the design, and programming of the systems themselves. This implies direct intervention in technology and consideration of its applications and consequences for all contexts in general, but also for each specific context (16).

3. Various implications in the social, ethical, and bioethical spheres

As we have seen, the implications of algorithmic bias are multiple and profound, ranging from replicating stereotypes of a particular culture or gender to promoting discourse that, in a veiled manner, unintentionally undermines a person's dignity based on their gender, sex, religion, etc.[1] Obviously, this has a direct impact on society, ethics, and bioethics since, as Barocas, Hardt, & Narayanan (7) point out, algorithmic bias can exacerbate existing social inequalities, leading to ethical and bioethical considerations in the use of AI systems for the promotion of justice and equity.

---

[1] Such is the position we adopt in this analysis.

Some of the examples that Diakopoulos (10) gives are justice systems based on AI systems, since, if there are biases in their algorithms, certain ethnic groups could be treated unfairly, even discriminatorily, which could aggravate or confirm that, due to their membership in these groups, they may suffer more or be better able to bear the weight of the law, thereby eliminating justice and equity in critical contexts.

Another bias that can affect bioethics is the provision of autonomy and informed consent for various treatments, since, if biases exist, the primacy of care and dignified treatment of individuals may be affected, thereby giving priority attention to individuals from certain sectarian groups, invalidating *triage* and the medical assessment of each individual patient (13).

Furthermore, algorithmic bias erodes trust in technology because it raises the dilemma of the objectivity of AI systems and, therefore, of their products and results; even more so when this has implications for robotics and anthropomorphic systems, since if AI algorithms produce unfair or discriminatory results, society may lose faith in technology and reject its adoption (1).

This, for example, can have negative consequences in areas such as healthcare, where AI has proven to be a good tool for diagnosing and treating patient conditions (9). However, imagine that the diagnosis is biased in its programming and therefore improves the health of certain patients but, at the same time, causes harm or even death in others based solely on their ethnicity or gender. It is therefore vitally important to address algorithmic biases, which is why ethics and fairness must be considered in any AI system, especially in its development, use, implementation, and management, in addition to the need for a person to help interpret, validate, or correct the results of AI (9).

4. Transhumanism and AI

As we have seen above, the use of AI and the biases it presents not only affect justice and ethics, but also bioethics. One of the clearest

examples is the improvement of natural human conditions by unnecessary means, better known as transhumanism. The link between AI systems and transhumanism raises profound questions about how technology can improve human capabilities, especially when we recognize transhumanism as a movement that seeks to improve human capabilities through advanced technology (17,18).

That is why, as Bostrom (17) points out, this quest for technological "improvement" must carefully address algorithmic bias, since, if this is not done, there is a risk of amplifying the very prejudices that are being sought to overcome. Hence the much needed and thoughtful value of the convergence of these two phenomena so that they can coexist in a bioethical and beneficial way for humanity: technology and the search for improvement in the field of health (11,12,19,20).

In line with the above, it is essential to recognize that technological improvement of human beings must be aligned not only with ethical and social values, but also with the pursuit of a more just and equitable society. Therefore, it is not simply a matter of improving human conditions and capabilities, regardless of the cost, but of improving the quality of human life in an ethical manner, always respecting the dignity of the person and the common good (21,22).

That said, algorithmic bias is, to a certain extent, a significant obstacle to technological improvement, especially since it is based on perpetuating prejudices and inequalities. Therefore, transhumanism is, in itself a problem of prejudice, bias, and stereotypes at its core, since it is not based on what it means to be human and the pursuit of quality of life, but rather on improving human beings because, in and of themselves, people are not adequate and, therefore, only those who can be improved will be the most fit to live in society (17,18).

4.1. *The search for anthropology that defends human dignity*²

In order to delve into the topic of "improvement" or "human perfection," that is, to bring the use of AI into the field of transhu-

---

² We understand this to mean an ontologically grounded anthropology.

manism (12,23), it is necessary to establish a complete philosophical anthropology that, as Floridi suggests (24), must consider all dimensions of the person, their complexity, but also their faculties and the diversity of experiences in the context of algorithmic bias in AI.

In this sense, if we recognize that AI algorithms, which have the ability to interpret and reflect on large amounts of data in seconds, are also inherently a reflection of the humanity that creates and uses them, hence the importance of having a philosophical anthropology that highlights the importance of empathy and ethics in the design of technological systems that respect and value human dignity (24) and, therefore, the need to recognize that technology should not be simply a cold extension of logic, but a tool that reflects true human faculties, their link with their dimensions, without negatively affecting them, but rather the need to develop and expand them and, therefore, nurture human dignity and humanity itself (21,22).

Furthermore, if we consider that technology and the use of AI arise from human faculties as an extension and/or tool that enhances them, then understanding what conception of human being underlies advances and, therefore, biases is vital. Hence the importance of considering a philosophical anthropology that emphasizes responsibility in the development of AI (21,22), especially since if we recognize the value of the freedom of algorithm developers and programmers, we also recognize their ethical and bioethical responsibility to ensure that their creations not only do not harm individuals or communities but also seek to benefit them.[3]

This implies, on the one hand, mitigating algorithmic biases while promoting equity and justice (25) as fundamental principles for achieving a life of solidarity and subsidiarity (26). Therefore, technology must be seen as a means of expanding and improving the human condition, rather than nullifying it (24), which requires a true understanding of human nature in all its dimensions and

---

[3] Here we refer not only to the bioethical principles of beneficence and non-maleficence (25), but also to the principle of freedom-responsibility of personalism (26).

how systems can serve humanity in an ethical and bioethical manner (26).

Therefore, when analyzing the biases of AI systems, whatever their use, it is necessary to consider a philosophical anthropology background, since anthropological analysis, which in many cases overlooks the dynamics of man with technology, must now turn to look at it and walk hand in hand. In the words of Latour (27), the rapid evolution of technology can challenge traditional notions of identity and human nature and, therefore, the adaptation and adoption of technology is adjusted to various cultural and socioeconomic contexts, but not necessarily anthropological ones. In short, it is not about how technology affects human beings, but how society adapts and reformulates its values to suit the technological world.

Thus, conducting a contextualized and applied analysis in favor of human dignity and applying it to specific contexts can help reduce different conceptions of bias and the vision behind prejudices and stereotypes if the diverse cultures, religions, and social groups behind the design and programming of algorithms are considered. In fact, it should be part of the anthropological analysis that requires seeing technology as part of society and how interactions between technology and society change, as well as between people and the way decisions are made based on their connection to AI systems.

5. Ethical and bioethical dilemmas of algorithmic bias

Another element that has not been considered and that also maximizes concerns, especially for some international organizations (28,29), is that, in addition to the conflicts we have already seen around dignity, justice, and equity, it is also true that algorithms, both their development and their biases, are driven by economic and technological objectives, especially because, as Mittelstadt *et al.* (30), maximizing efficiency and making AI systems more profitable prevail over the social impacts they may have (9).

In this sense, it is clear that the creation and development of technologies that benefit certain social groups, especially over the majority of the population, undermine justice and equity, as we have mentioned, and also nullify the dignity and opportunities for access, enjoyment, and benefits offered by AI systems, thus sacrificing ethics and human dignity for economic gain, raising questions about the values and priorities in the implementation of AI algorithms (30).

The lack of ethical consideration in the implementation of AI can have serious consequences (9), as it can lead to exploitation and discrimination. As we have seen in its application to transhumanism, it also recognizes what or who a human is and, therefore, could have the basis and foundations for defining the human being. In addition, if AI algorithms can be used for human interaction and decision-making in various aspects of life, then imagine a case in which they are also integrated into human beings. Where would human freedom and the very fallibility of human beings lie? (31,32,33,34,35).

To answer this question, we must affirm that if these algorithms are not ethical and fair, they can perpetuate social prejudices, exponentially and, over time, further divide society between those who can and cannot afford modifications, those who belong to different ethnic groups, those of a specific gender, or any other segmentation that allows for algorithmic bias. This leads us to reaffirm the importance of ethical criteria being an integral part of the AI development process, and not simply an afterthought.

6. Proposed solution: mitigating algorithmic bias

The recognition that AI technology, although advanced, is not immune to human failings has led to the implementation of various strategies to combat bias. Therefore, to address algorithmic bias and ethical concerns in AI, it is imperative to prioritize transparency in the design and operation of algorithms (36), especially since a "Transparency by Design" approach can offer practical guidance for

promoting the beneficial functions of transparency while mitigating its challenges in automated decision-making environments.

6.1. *Transparency by Design*

According to Felzmann *et al.* (37), adopting the principles of "Transparency by Design" can help organizations, companies, and algorithm programmers systematically integrate practices that ensure accountability is a priority from the planning and development of AI systems and not, as is currently the case, an afterthought. This is because the principles of "Transparency by Design" encompass considerations of context, techniques, information, and the care of data sensitive to all parties involved, including the end user, which implies taking care of a multiplicity of dimensions to achieve effectiveness and efficiency in transparency.

However, implementing these principles also presents major challenges, which are not impossible to overcome. To begin with, it is vital to recognize the complexity of algorithms and decision-making processes and how, for those who are not specialists, they are very difficult, even opaque to understand, which makes it impossible for the public to comprehend and analyze them.

6.2. *Intellectual property and security of AI systems*

Secondly, it is also important to consider the intellectual property and security of systems, especially those that handle sensitive and/or biometric data, as this information cannot be easily disclosed; however, if the algorithm were made public, it could lead to unnecessary data leaks that could put multiple people at risk.

Based on the above, Kirat *et al.* (38), when discussing fairness and accountability in algorithmic decision-making, suggest that, when contextualizing algorithmic fairness, especially in relation to the legal frameworks applicable today in the United States and Europe, there are diverse and even contradictory positions on what constitutes

true and adequate accountability and how transparency should be applied in practice. This emphasizes the importance of developing international standards consistent with each context to guide the effective implementation of policies and regulations around algorithm transparency and accountability.

6.3. *Multi- and transdisciplinary teams*

Third, we have already mentioned the importance of algorithm design and development teams for AI systems, especially because greater diversity means less possibility of bias. If we consider that members have a wide range of perspectives (engineering, bioethics, philosophy, mechanics, etc.), in addition to the experiences they can share, this can significantly help to rationalize biases and, therefore, eliminate unnoticed prejudices in the initial stages of algorithm design and development.

This is not exclusive to the field of AI design, but also applies to other fields, since whenever diversity and interdisciplinarity are encouraged in the development teams of any project, the biases inherent in the practice are reduced and a variety of perspectives are ensured in the development process (39). With this, we can confirm that, in the development of AI systems, if we want to seek equity, justice, respect for human dignity, and respect for the principles of bioethics, then the more voices there are, the more likely it is that the final product will embody these principles and, therefore, be an inclusive and equitable system.

6.4. *Ethical design and evaluation methodologies for bias mitigation*

Fourth, we must also consider the integration of ethical methodologies into the design and evaluation of AI systems to promote fair and equitable algorithms. This implies that the data sets used in AI training are constantly reviewed to ensure that they are truly representative of all groups (majorities and minorities) and free from

as much bias (historical and modern) as possible. In addition, the model must be continuously evaluated to identify when biases arise during the execution of the algorithms, whether due to natural programming or inputs, allowing them to be corrected at their source.

In this sense, transparency and accountability of algorithms become essential, as we saw earlier, since they allow for a better understanding of how AI makes decisions. Therefore, it is vital to "make algorithms more transparent and explainable to help users understand how decisions are made." (40, n.p.), leading to a better way of evaluating systems, measuring their ethical and bioethical impacts, and facilitating the identification and correction of both emerging biases and natural errors in algorithms.

6.5. *Roles of governments and international organizations in mitigation*

The fifth and final proposed solution for mitigating algorithmic bias involves the involvement of governments and international organizations, particularly because the creation of standards, laws, and policies that, in addition to incorporating ethical and bioethical principles (25,26), require fairness, transparency, justice, and adequate accountability in AI systems would be a crucial step toward mitigating bias.

In this regard, the *UNESCO Recommendation on the Ethics of AI* (41), *the Rome Call for AI Ethics* (24), and *the European Union's Artificial Intelligence Act* (42) are clear examples of how international bodies guide member states in achieving and incorporating ethical practices throughout the process of developing, programming, and implementing AI algorithms, with particular emphasis on the aforementioned elements, including the protection of human rights and human dignity. These efforts not only highlight global collaboration but also promote shared ethical and bioethical principles to address the challenges posed by AI.

In addition to policies that uphold ethical and bioethical principles (25, 26), it is also important that regulations, society, and AI

companies, together, adopt a proactive approach to the creation, development, programming, use, and implementation of AI systems (24), due to the impacts that technology will have not only in the short and medium term, but also in the long term, on society and, with this, anticipate possible adverse effects.

Finally, integrating interdisciplinary dialogue between philosophers, bioethicists, social scientists, engineers, mechatronics engineers, political leaders, and other members of society involved in AI to generate principles and regulations will lead to a collaborative approach oriented toward equity and justice, which is fundamental for mediating, mitigating, and, to a certain extent, eliminating algorithmic biases, which will lead to the use and development of AI systems that are responsible and beneficial to humanity.

Looking ahead, the commitment to making AI equitable and fair in all its dimensions necessarily requires global collaboration that transcends borders, disciplines, and sectors of society. It also involves establishing international standards for ethics and bioethics in AI. Therefore, the vision is one in which AI not only reflects human ethical and bioethical values, especially equity and justice, but also acts as a catalyst for reducing social inequalities. Finally, as suggested by Gebru *et al.* (43), addressing algorithmic bias in AI systems presents a unique opportunity to reexamine and improve how emerging technologies best serve society and align with ethical principles.

7. Conclusions

In conclusion, what we have highlighted in this paper has underscored the critical importance of mitigating algorithmic bias to ensure the equitable development and use of Artificial Intelligence. To achieve this, the need to consider transparency, interdisciplinarity, and ethics in development teams, as well as the implementation of rigorous ethical audits based on ethical standards and principles, must be at the forefront. As Diakopoulos (10) and Barocas, Hardt,

and Narayanan (7) mention, it is vital to prioritize these strategies to rationalize and combat the social and individual biases of programmers, which are inherent in the processes of AI creation, development, and programming.

In this regard, interdisciplinary collaboration in development processes, coupled with interreligious dialogue and diverse perspectives, is essential, as this ensures that a wide range of experiences, viewpoints, and ethical considerations contribute to the creation of fairer and more equitable algorithms. Only through a plurality of approaches can we openly confront and overcome the prejudices rooted in society and, therefore, in AI, which threaten to perpetuate the inequalities and discrimination that such systems entail.

We have also pointed out the importance of making AI accessible so that it can be used as a means of improving society and the human condition, while always respecting human dignity and the principles of bioethics (26). Technology, as is often said, must be at the service of humanity and, therefore, must actively seek to eliminate social inequalities and combat discrimination. This objective goes beyond the simple technical use of AI for profit; rather, it is a fundamental ethical and bioethical imperative that must guide its development and use.

Finally, we conclude with an urgent call to action for all those involved in AI systems (from creation to the end user): it is necessary to adopt an interdisciplinary and comprehensive approach that encompasses everything from transparency in the design and operation of algorithms to a firm focus on equity and justice in all aspects of their implementation, as this is the only way to ensure that technological progress benefits humanity in line with the highest ethical and bioethical standards.

References

1.   Mateescu A, Elish M. AI in Context: The Labor of Integrating New Technologies. Chicago: Data & Society Research Institute; 2015.

2. Meissner P, Narita Y. Así es como la inteligencia artificial transformará la toma de decisiones. Tecnologías emergentes – WEF; 2023. Available at: https://es.we-forum.org/agenda/2023/10/la-inteligencia-artificial-transformara-la-toma-de-deci-siones-asi-es-como/

3. Lanzagorta-Ortega D, Carrillo-Pérez DL, Carrillo-Esper R. Inteligencia artificial en medicina: presente y futuro. Gac Med Méx. 2021; 158(1):17-21. Available at: https://doi.org/10.24875/gmm.m22000688

4. Chui M, Manyika J, Miremadi M. Where machines could replace humans—and where they can't (yet). McKinsey Quarterly. 2019:1-12. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/where-machines-could-replace-hu-mans-and-where-they-cant-yet

5. Ferrante E. Inteligencia artificial y sesgos algorítmicos. Nueva Sociedad. 2021; (294):27-36. Available at: https://nuso.org/articulo/inteligencia-artificial-y-ses-gos-algoritmicos/

6. Simon J, Wong PH, Rieder G. El sesgo algorítmico y el enfoque del diseño sensi-ble al valor. Rev. Latinoam. Econ. Soc. Digit. 2022; (Special Number 1):1-18. Avai-lable at: https://doi.org/10.53857/tzvn9229

7. Barocas S, Hardt M, Narayanan A. Fairness and Machine Learning: Limitations and Opportunities. 2019; Available at: fairmlbook.org

8. Asís RD. Una mirada a la robótica desde los derechos humanos. Dykinson; 2014.

9. Terrones AL. Inteligencia Artificial fiable y vulnerabilidad: una mirada ética so-bre los sesgos algorítmicos. In Suárez-Álvarez R, Martín-Cárdaba MÁ, Fernán-dez-Martínez LM. Vulnerabilidad digital: desafíos y amenazas de la Socieda hi-perconectada. Madrid: Dykinson; 2023.

10. Diakopoulos N. Accountability in Algorithmic Decision Making. Commun ACM. 2016; 59(2):56-62. Available at: https://doi.org/10.1145/2844110.

11. Velázquez H. Transhumanismo, libertad e identidad humana. Thémata. 2009; (41):577-590. Available at: https://revistascientificas.us.es/index.php/themata/arti-cle/view/594

12. Olarte C, Plegrín J, Reinares E. Implantes para aumentar las capacidades inna-tas: integrados vs apocalípticos. ¿Existe un nuevo mercado? Universia Bus Rev. 2015:86-101. https://www.redalyc.org/articulo.oa?id=43343050003

13. Abràmoff M, Tarver M, Loyo-Berrios N, Trujillo S, Char D, Obermeyer Z, Maisel, W. Considerations for addressing bias in artificial intelligence for health equity.npj Digital Med. 2013; 6(170):1-7. Available at: https://doi.org/10.1038/s41746-023-00913-9

14. Straw I, Rees G, Nachev P. Sex-Based Performance Disparities in Machine Lear-ning Algorithms for Cardiac Disease Prediction: Exploratory Study. J Med Internet Res; 2024; 26:e46936. Available at: https://doi.org/10.2196/46936

15. Akter S, Dwivedi YK, Sajib S, Biswas K, Bandara RJ, Michael K. Algorithmic bias in machine learning-based marketing models. J. Bus. Res. 2022; 144:201-216. Available at: https://doi.org/10.1016/j.jbusres.2022.01.083

16. U.S. Department of Education, Office of Educational Technology. Artificial Intelli-gence and Future of Teaching and Learning: Insights and Recommendations. Washington, DC; 2023.

17. Bostrom N. A history of transhumanist thought. Journal of Evolution and Technology. 2005; 14(1):1-25. Available at: https://nickbostrom.com/papers/a-history-of-transhumanist-thought/

18. Diéguez A. Transhumanismo: la búsqueda tecnológica del mejoramiento humano. Barcelona: Herder; 2017.

19. Harvard University. Justice with Michael Sandel - CCCB: Bioethics: Designer children. [video]; 2011. Available at: https://youtu.be/aFcfygkMM0I

20. Tobin H. Transhumanismo SuperPoderes Humanos – Castellano. [video]; 2016. Available at: https://www.youtube.com/watch?v=ClxzKdfQAkY

21. Fuentes, MÁ. Principios fundamentales de bioética. Roma: Instituto del Verbo Encarnado; 2006.

22. Berti B. Los principios de la bioética. Prudent Iuris. 2015; (79):269-280. Available at: https://erevistas.uca.edu.ar/index.php/PRUDENTIA/article/view/4030

23. Thompson J. (2017). Transhumanism: How Far Is Too Far? The new bioethics. 2017; 23(2):165–182. Available at: https://doi.org/10.1080/20502877.2017.1345092

24. Floridi L. The Logic of Information: A Theory of Philosophy as Conceptual Design. Oxford: Oxford University Press; 2019.

25. Beauchamp T, Childress, JF. Principles of biomedical ethics. Oxford: Oxford University Press; 2019.

26. Sgreccia E. Manual de bioética: fundamentos y ética biomédica. Navarra: EUNSA; 2016.

27. Latour B. Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. En Bijker W, Law J. Shaping Technology-Building Society. Studies in Sociotechnical Change. MIT Press, Cambridge Mass; 1992.

28. Future of Life Institute. Asilomar AI Principles. De Future of Life Institute; 2017. Available at: https://futureoflife.org/ai-principles/

29. Paglia VS. Rome Call for AI Ethic; 2023. Available at: https://www.romecall.org/wp-content/uploads/2022/03/RomeCall_Paper_web.pdf

30. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The Ethics of Algorithms: Mapping the Debate. Big Data Soc. 2016; 3(2). Available at: https://doi.org/10.1177/2053951716679679

31. Attiah M, Farah M. Minds, motherboards, and money: futurism and realism in the neuroethics of BCI technologies. Front Syst Neurosci. 2014; 86(8):1-3. https://doi.org/10.3389/fnsys.2014.00086

32. Evans J. Faith in Science in Global Perspective: Implications for Transhumanism. Public Underst Sci. 2014; 23(7):1-33. Available at: https://doi.org/10.1177/0963662514523712

33. Schaefer O, Savulescu J. Better Minds, Better Morals: A Procedural Guide to Better Judgement. J. Posthuman Stud. 2017; 1(1):26-43. Available at: http://dx.doi.org/10.5325/jpoststud.1.1.0026

34. Zehr P. Future think: cautiously optimistic about brain augmentation using tissue engineering and machine interface. Front Syst Neurosci. 2015; 9(72):1-5. Available at: https://doi.org/10.3389/fnsys.2015.00072

35. Huxley A. Un mundo feliz. México: Éxodo; 2011.
36. Doshi-Velez F, Narayanan M, Chen E, He J, Kim B, Gershman S. How do Humans Understand Explanations from Machine Learning Systems? An Evaluation of the Human-Interpretability of Explanation. NY: Cornell University; 2018. Available at: https://doi.org/10.48550/arXiv.1802.00682
37. Felzmann H, Fosch-Villaronga E, Lutz C, Tamò-Larrieux A. Towards Transparency by Design for Artificial Intelligence. Sci Eng Ethics. 2020; 26(6):3333–3361. Available at: https://doi.org/10.1007/s11948-020-00276-4
38. Kirat T, Tambou O, Do V, Tsoukiàs A. Fairness and explainability in automatic decision-making systems. A challenge for computer science and law. EURO J. Decis. Process. 2023; 11:1-19. Available at: https://doi.org/10.1016/j.ejdp.2023.100036
39. Gavilán I. Ideas y buenas prácticas para evitar el sesgo algorítmico. En: Ignacio G.R. Gavilán; 2022. Available at: https://ignaciogavilan.com/ideas-y-buenas-practicas-para-evitar-el-sesgo-algoritmico/
40. Sandu E. Sesgo en los algoritmos de inteligencia artificial: causas y soluciones. metaverso.pro - Metaverso para Profesionales; 2023. Available at: https://metaverso.pro/blog/sesgo-en-los-algoritmos-de-inteligencia-artificial-causas-y-soluciones/
41. UNESCO. Inteligencia artificial y educación: guía para las personas a cargo de formular políticas. Francia: UNESCO; 2021.
42. Comisión Europea. Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican diversos actos legislativos de la Unión. Bruselas; 2024.
43. Gebru T, Morgenstern J, Vecchione B, Wortman J, Wallach H, Daume HI, Crawford K. Datasheets for Datasets. Commun ACM. 2021; 64(12):86-92. Available at: https://doi.org/10.1145/3458723