

AI Ethics in the Fields of Education and Research: A Systematic Literature Review

Mauricio Rojas Contreras
Research Group CICOM
University of Pamplona
Pamplona, Colombia
mrojas@unipamplona.edu.co

Jorge Omar Portilla Jaimes
Research Group CICOM
University of Pamplona
Pamplona, Colombia
jorge.portilla2@unipamplona.edu.co

Abstract— The rapid emergence of artificial intelligence tools and their integration into various societal sectors have introduced significant ethical challenges. This research focuses on identifying the most frequently discussed ethical aspects of AI within the realms of education and research. A systematic literature review, supported by AI-driven tools, was conducted to address this objective. The study reveals that key ethical aspects recurrently examined include transparency and explainability, privacy and data protection, responsibility and accountability, and equity. Transparency and explainability are crucial for ensuring that AI systems in education operate in a comprehensible manner, fostering trust among educators and students. Privacy and data protection emphasize the importance of handling personal data with utmost care, adhering to relevant regulations. Responsibility and accountability highlight the need for clear mechanisms for overseeing AI deployment, while equity addresses the avoidance of bias and discrimination. The findings underscore the necessity of addressing a broad spectrum of ethical considerations throughout the lifecycle of AI systems to promote their responsible use in educational and research settings.

Keywords— *generative AI, AI Ethics, Education, Research, privacy*

I. INTRODUCTION

Artificial intelligence (AI) represents a powerful tool with the potential to significantly transform education and research, offering substantial benefits globally. However, its implementation raises fundamental ethical concerns. The inherent biases in AI algorithms can exacerbate discrimination, inequality, and digital divides, threatening cultural, social, and biological diversity. Additionally, the lack of transparency and intelligibility in the functioning of algorithms and the data used to train them poses serious ethical challenges. These issues can affect human dignity, human rights, and fundamental freedoms, as well as gender equality, democracy, and social, economic, political, and cultural processes. There is also an impact on scientific and engineering practices, animal welfare, and the environment [1].

In the context of UNESCO, particularly in the Recommendation on the Ethics of Artificial Intelligence, ethical principles are established to guide the development and use of artificial intelligence, including human dignity and well-being, inclusion and equity, stakeholder participation, equitable access, shared responsibility, and compliance with international norms. Regarding the field of education and research, essential aspects are specified, such as ethics training in AI, competence development, responsible research,

inclusion and diversity, interdisciplinary collaboration, impact assessment, and the promotion of open research [1].

On the other hand, in the context of the European Union's Artificial Intelligence Regulation, the objectives include promoting human-centered and reliable AI, protecting against harmful effects, and establishing standards to mitigate the risks associated with AI systems. To ensure that AI is human-centered, key provisions are established related to harmonized standards, a focus on fundamental rights, human oversight, risk assessment, and ethical principles [2].

Regarding the ethical aspects of AI, human action and oversight, technical robustness and safety, privacy and data management, transparency, diversity, non-discrimination and equity, social and environmental well-being, accountability, and the verification and validation of generated knowledge are considered.

Focusing on the ethical aspects of AI in the field of education and research, topics related to AI literacy, transparency and explainability, non-discrimination and equity, data protection and privacy, student well-being, stakeholder participation, responsible research, and the verification and review of knowledge generated in research should be considered.

Therefore, to minimize the gap in identifying the aspects to consider in AI ethics in the field of education and research, we aim to answer the following research question:

Q1: What are the essential aspects to consider in AI ethics in the field of education and research?

This article focuses on examining the ethical aspects of AI in the field of education and research using a systematic review method in accordance with PRISMA protocol recommendations. It provides a broad and well-founded perspective based on the UNESCO Recommendation on the Ethics of Artificial Intelligence [1].

II. RELATED WORKS

The dynamics of technology, particularly the emergence of generative artificial intelligence and AI-powered tools, have added value to processes integrating these technologies across different sectors of society. However, the integration of AI in the fields of education and research has created a set of gaps, notably in the management of AI's ethical aspects.

Greene et al. [3] identified the ethical aspects of AI as interdisciplinarity, ethics education, social impact, norms and regulations, responsibility and accountability, ethics in peer review, and the development of intellectual virtues.

Interdisciplinarity is conceived as the ability to integrate different disciplines to address the ethical issues of AI. Ethics education aims to provide training in ethical principles and their application to those who interact with AI. Social impact concerns evaluating the social implications of AI research. Norms and regulations seek to understand and comply with the ethical and legal standards affecting AI research. Responsibility and accountability aim to establish mechanisms for accountability in the development and use of AI technologies. Ethics in peer review seeks to include ethical impact statements in the peer review process during publication. The development of intellectual virtues aims to foster virtues such as curiosity, openness, and empathy among AI researchers.

Christoforaki and Beyan [4] specified the ethical aspects of AI as transparency and explainability, responsibility and accountability, privacy and data protection, fairness and non-discrimination, beneficence and non-maleficence, and education and awareness. Transparency and explainability are described as the characteristic of AI systems used in education being transparent in their operation and ensuring that their decisions are explainable to educators and students. Responsibility and accountability, from the authors' perspective, aim to establish who is responsible for the decisions made by AI. Privacy and data protection refer to handling the collection and use of students' personal data with the utmost care, ensuring the privacy and security of the information. Fairness and non-discrimination seek to ensure that AI systems do not perpetuate biases or discriminate against vulnerable groups. Beneficence and non-maleficence aim for AI systems to be designed and used to promote the well-being of students and avoid causing harm. Education and awareness aim to educate both educators and students about the ethical use of AI.

III. MATERIALS AND METHODS

The research strategy employed in this document is a systematic literature review, which allowed for the identification of the scope of studies on AI ethics in the field of education and research. The main objective of this study is to obtain a broad perspective on AI ethics in these fields, according to the UNESCO document titled Recommendation on the Ethics of Artificial Intelligence [1].

In this research, the following research question was examined:

Q1: What are the essential aspects to consider in AI ethics in the field of education and research?

To answer this question, the search equation was generated using a generative AI tool, as indicated in Table I.

TABLE I. SEARCH EQUATION

Prompt	generative AI	version	Search equation
Generate an optimized search equation for a systematic review article on 'AI ethics in the fields of education and research'	Chatgpt	4.0	("artificial intelligence" OR "AI") AND (ethics OR ethical) AND (education OR "educational research" OR academia OR "academic research") AND (guidelines OR policies OR "ethical

			considerations" OR "ethical issues" OR "moral implications")
--	--	--	--

Subsequently, relevant databases for the fields of social sciences, computer science, and engineering were identified. Databases that cover the most important conferences and journals in the field of educational technology were considered, such as Scopus and ScienceDirect.

Taking the search equation into account, a search was conducted in each of the electronic databases, applying filters for the years 2022-2024 and selecting only review articles. This search yielded the number of documents shown in Table II.

TABLE II. NUMBER OF DOCUMENTS BY DATABASE

Search equation	SCOPUS	SCIENCEDIRECT
("artificial intelligence" OR "AI") AND (ethics OR ethical) AND (education OR "educational research" OR academia OR "academic research") AND (guidelines OR policies OR "ethical considerations" OR "ethical issues" OR "moral implications")	75	521

Subsequently, inclusion (I) and exclusion (E) criteria were implemented to add or discard documents for review:

Inclusion criteria (I):

- I1: Articles published between 2022 and 2024.
- I2: Articles in the fields of social sciences, computer science, and engineering.
- I3: Articles written in English.
- I4: Review articles.

Exclusion criteria (E):

- E1: Articles whose title and abstract did not make a significant contribution to the study topic.
- E2: Articles that, upon full review, were not completely aligned with the study's objectives.

IV. RESULTS

The systematic literature review process was supported by the AI-powered tool Rayyan.ai, which allows for importing files from search results in the SCOPUS and SCIENCE DIRECT databases. Once the files were imported, Rayyan.ai integrated them and supported the screening process by analyzing each reference through the title and abstract. The result of this initial level of document filtering, based on the

title and abstract, identified 21 references for full-text analysis, as specified in Table III.

TABLE III. REFERENCES FOR SYSTEMATIC REVIEW

identifier	Reference	Title
R1	Greene, T., Dhurandhar, A., & Shmueli, G. (2022). Atomist or holist? A diagnosis and vision for more productive interdisciplinary AI ethics dialogue. <i>Patterns</i> , 4.[3]	Atomist or Holist? A Diagnosis and Vision for More Productive Interdisciplinary AI Ethics Dialogue
R2	Christoforaki, M., & Beyan, O.D. (2022). AI Ethics—A Bird’s Eye View. <i>Applied Sciences</i> [4]	AI Ethics—A Bird’s Eye View
R3	Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. <i>Expert Systems with Applications</i> [5]	Artificial intelligence in education: A systematic literature review
R4	Oforu-Ampong, K. (2024). Artificial Intelligence Research: A Review on Dominant Themes, Methods, Frameworks and Future Research Directions. <i>Telematics and Informatics Reports</i> . [6]	Artificial intelligence research: A review on dominant themes, methods, frameworks and future research directions
R5	Sharma, S. (2024). Benefits or Concerns of AI: A Multistakeholder Responsibility. <i>Futures</i> . [7]	Benefits or concerns of AI: A multistakeholder responsibility
R6	Memarian, B., & Doleck, T. (2023). ChatGPT in education: Methods, potentials and limitations. <i>Computers in Human Behavior: Artificial Humans</i> . [8]	ChatGPT in education: Methods, potentials, and limitations
R7	Rawas, S. (2024). AI: the future of humanity. <i>Discov. Artif. Intell.</i> , 4. [9]	AI: the future of humanity
R8	Mennella, C., Maniscalco, U., De Pietro, G., & Esposito, M. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. <i>Heliyon</i> , 10. [10]	Ethical and regulatory challenges of AI technologies in healthcare: A narrative review
R9	Ong, J.C., Chang, S.Y., William, W., Butte, A.J., Shah, N.H., Chew, L.S., Liu, N., Doshi-Velez, F., Lu, W., Savulescu, J., & Ting, D.S. (2024). Ethical and regulatory challenges of large language models in medicine. <i>The Lancet. Digital health</i> . [11]	Ethical and regulatory challenges of large language models in medicine
R10	Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. <i>J. Strateg. Inf. Syst.</i> , 32, 101772. [12]	Ethical management of human-AI interaction: Theory development review
R11	Bettayeb, A.M., Abu Talib, M., Sobhe Altayasinah, A.Z., & Dakalbab, F.M. (2024). Exploring the impact of ChatGPT: conversational AI in education. <i>Frontiers in Education</i> . [13]	Exploring the impact of ChatGPT: conversational AI in education
R12	Memarian, B., & Doleck, T. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher	Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence

	education: A systematic review. <i>Comput. Educ. Artif. Intell.</i> , 5, 100152. [14]	(AI) and higher education: A systematic review
R13	Qadhi, S.M., Alduais, A., Chaaban, Y., & Khraishah, M. (2024). Generative AI, Research Ethics, and Higher Education Research: Insights from a Scientometric Analysis. <i>Inf.</i> , 15, 325.[15]	Generative AI, Research Ethics, and Higher Education Research: Insights from a Scientometric Analysis
R14	Andrade-Girón, D.C., Marín-Rodríguez, W.J., Sandivar-Rosas, J., Carreño-Cisneros, E., Susanibar-Ramirez, E.T., Zúñiga-Rojas, M., Ángeles-Morales, J., & Villarreal-Torres, H. (2024). Generative artificial intelligence in higher education learning: A review based on academic databases. <i>Iberoamerican Journal of Science Measurement and Communication</i> . [16]	Generative artificial intelligence in higher education learning: A review based on academic databases
R15	Farrelly, T., & Baker, N. (2023). Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice. <i>Education Sciences</i> . [17]	Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice
R16	Kasani, P.H., Cho, K.H., Jang, J., & Yun, C. (2024). Influence of artificial intelligence and chatbots on research integrity and publication ethics. <i>Science Editing</i> . [18]	Influence of artificial intelligence and chatbots on research integrity and publication ethics
R17	Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. <i>International Journal of Educational Technology in Higher Education</i> , 20.[19]	Role of AI chatbots in education: systematic literature review
R18	Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. <i>Sustainability</i> . [20]	Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis
R19	Franco D'Souza, R., Mathew, M., Mishra, V., & Surapaneni, K.M. (2024). Twelve tips for addressing ethical concerns in the implementation of artificial intelligence in medical education. <i>Medical Education Online</i> , 29.[21]	Twelve tips for addressing ethical concerns in the implementation of artificial intelligence in medical education
R20	Maccaro, A., Stokes, K., Statham, L., He, L., Williams, A., Pecchia, L., & Piaggio, D. (2024). Clearing the Fog: A Scoping Literature Review on the Ethical Issues Surrounding Artificial Intelligence-Based Medical Devices. <i>Journal of Personalized Medicine</i> , 14.[22]	Clearing the Fog: A Scoping Literature Review on the Ethical Issues Surrounding Artificial Intelligence-Based Medical Devices
R21	Khalifa, M., Albadawy, M., Using artificial intelligence in academic writing and research: An essential productivity tool. <i>Computer Methods and Programs in Biomedicine Update</i> 5.[23]	Using artificial intelligence in academic writing and research: An essential productivity tool

With the references selected for full-text analysis, the AI-powered tool Chatpdf was used to analyze each article and address the research question. Once the results of all the article analyses were consolidated, the ethical aspects to be considered in the fields of education and engineering were identified, as shown in Figure 1.

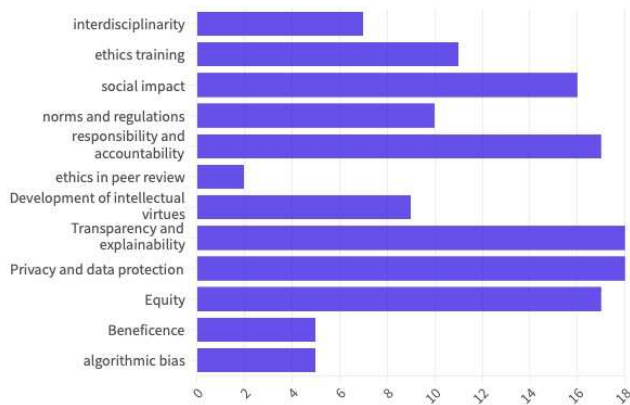


Fig. 1. Ethical aspects of AI

Figure 1 identifies transparency and explainability, as well as privacy and data protection, as the ethical aspects most frequently studied in the systematic literature review. With slightly lower frequency are fairness, responsibility, and accountability. On the other hand, ethics in peer review is identified as the ethical aspect analyzed with the least frequency in the systematic review.

At a disaggregated level, Figure 2 presents the ethical aspects of AI along with their occurrence in each of the references.

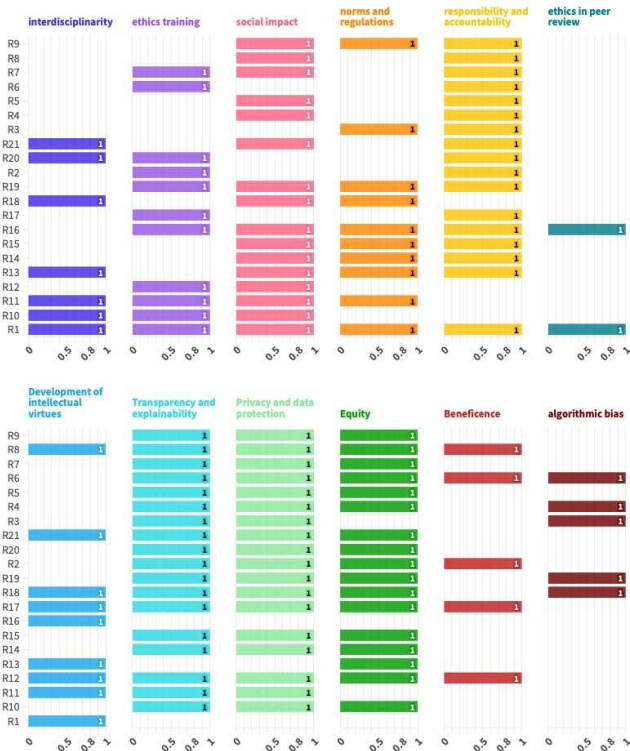


Fig. 2. Ethical aspects of AI with references

Figure 2 specifies the literature review references that identify each of the analyzed ethical aspects of AI. Specifically, it examines interdisciplinarity, ethics education, social impact, norms and regulations, responsibility and accountability, ethics in peer review, the development of intellectual virtues, transparency and explainability, privacy and data protection, fairness, beneficence, and algorithmic bias as relevant ethical aspects of AI.

An additional gap in the context of generative artificial intelligence is the control of hallucinations. Specifically, when a prompt is sent to a generative AI without specifying the temperature setting, it defaults to a value of 0.7. The temperature configuration variable can take values between 0 and 1. With temperature values close to 0, the response will be more deterministic and tends to minimize hallucinations. Conversely, when the temperature variable is set close to 1, the response will be more creative and hallucinations tend to be maximized. Some suggestions for configuring the temperature variable can be seen in Figure 3.

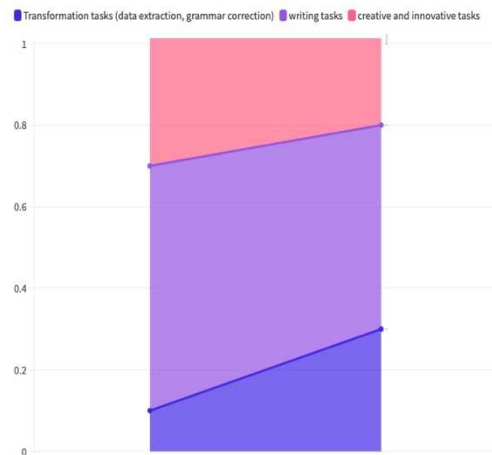


Fig. 3. Temperature variable configuration recommendation

Figure 3 provides some recommendations for configuring the temperature variable in the context of prompt engineering. Specifically, it is suggested that for research and educational activities, the temperature variable be set between 0.1 and 0.3, preferably closer to 0.1. For writing-related tasks, it is recommended to set the temperature variable in the range of 0.3 to 0.7. For creative tasks where the level of hallucinations is not important, it is suggested to configure the temperature variable with values between 0.7 and 0.9. These values are specified in the prompt by Configuration: Temperature: value.

In the context of this research, prompts with a temperature setting of 0.1 were used to find more deterministic responses and minimize hallucinations. Additionally, the triangulation method was used to verify and validate the knowledge generated by generative artificial intelligence [24],[25].

V. CONCLUSIONS

The systematic review in the article "AI Ethics in the Fields of Education and Research" reveals a wide range of essential ethical aspects for implementing AI in these fields.

The key aspects identified include interdisciplinarity, ethics education, social impact, norms and regulations, responsibility and accountability, ethics in peer review, the development of intellectual virtues, transparency and explainability, privacy and data protection, fairness, beneficence, and algorithmic bias. These aspects highlight the need for comprehensive and multidisciplinary approaches to effectively address ethical challenges. Additionally, they emphasize the importance of continuous ethics education, the adoption of clear standards, and a commitment to transparency and fairness. In summary, the ethical integration of AI in education and research requires a concerted effort to ensure these technologies benefit society in a fair and responsible manner.

Several critical ethical aspects of AI have been identified, with transparency and explainability, as well as privacy and data protection, standing out for their frequency and relevance. Transparency and explainability are essential for AI systems in education to function in a comprehensible manner and for their decisions to be clear to educators and students, thereby fostering trust. Privacy and data protection ensure that the collection and use of students' personal data are carried out with utmost care, respecting security and privacy, and complying with relevant regulations. These aspects are fundamental for the ethical implementation of AI in the educational and research fields, promoting the fair and responsible use of technology.

The ethical aspects of artificial intelligence in the field of education are essential to ensure the responsible and equitable use of these technologies. Transparency and explainability allow AI systems to be understandable and their decisions clear, fostering trust among educators and students. Additionally, privacy and data protection ensure that students' personal information is handled with the utmost care, complying with relevant regulations and safeguarding security. These ethical principles are crucial for the effective and fair integration of AI in education, promoting an inclusive and safe learning environment.

The ethical aspects of artificial intelligence in the field of research are crucial to ensuring integrity and responsibility in scientific development. Transparency and explainability of AI systems allow for greater understanding and replicability of results, strengthening trust within the scientific community. Privacy and data protection are vital for the ethical handling of sensitive information, ensuring compliance with regulations and data security. These ethical principles are essential to promoting responsible and sustainable research, facilitating technological advances that benefit society as a whole.

In the context of the ethical aspects of AI in education and research, the verification and validation of knowledge generated by generative AI are essential to ensure accuracy and reliability. It is recommended to minimize hallucinations by setting the Temperature variable to 0.1. Additionally, qualitative methodological triangulation allows for contrasting the information generated by AI with theoretical references and the researcher's perspective, promoting critical thinking skills. This combination of techniques ensures a rigorous evaluation and ethical use of AI in these fields.

REFERENCES

- [1] UNESCO, "Recomendación sobre la ética de la inteligencia artificial," 2022. [En línea]. Disponible: <https://www.unesco.org/open-access/terms-use-ccbysncsa-sp>.
- [2] Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 12 de julio de 2024, sobre la inteligencia artificial. [En línea]. Disponible: <http://data.europa.eu/eli/reg/2024/1689/oj>
- [3] T. Greene, A. Dhurandhar, G. Shmueli, "Atomist or holist? A diagnosis and vision for more productive interdisciplinary AI ethics dialogue". *Patterns*, 4, 2022.
- [4] M. Christoforaki, O. Beyan, "AI Ethics—A Bird's Eye View", *Applied Sciences*, 2022.
- [5] S. Wang, F. Wang, Z. Zhu, J. Wang, T. Tran, Z. Du, "Artificial intelligence in education: A systematic literature review", *Expert Systems with Applications*, 2024.
- [6] K. Ofori-Ampong, "Artificial Intelligence Research: A Review on Dominant Themes, Methods, Frameworks and Future Research Directions". *Telematics and Informatics Reports*, 2024.
- [7] S. Sharma, "Benefits or Concerns of AI: A Multistakeholder Responsibility". *Futures*, 2024.
- [8] B. Memarian, T. Doleck, "ChatGPT in education: Methods, potentials and limitations". *Computers in Human Behavior: Artificial Humans*, 2023.
- [9] S. Rawas, "AI: the future of humanity". *Discov. Artif. Intell.*, 4., 2024.
- [10] C. Mennella, U. Maniscalco, G. De Pietro, M. Esposito, "Ethical and regulatory challenges of AI technologies in healthcare: A narrative review". *Heliyon*, 10, 2024.
- [11] J.C. Ong, S. Y. Chang, W. William, A. J. Butte, N. H. Shah, L. S. Chew, N. Liu, F. Doshi-Velez, W. Lu, J. Savulescu, D. S. Ting, "Ethical and regulatory challenges of large language models in medicine". *The Lancet. Digital health*, 2024.
- [12] T. Heyder, N. Passlack, O. Posegga, "Ethical management of human-AI interaction: Theory development review". *J. Strateg. Inf. Syst.*, 32, 101772, 2023.
- [13] A. M. Bettayeb, M. Abu Talib, A. Z. Sobhe Altayasinah, F. M. Dakalbab, "Exploring the impact of ChatGPT: conversational AI in education". *Frontiers in Education*, 2024.
- [14] B. Memarian, T. Doleck, "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review". *Comput. Educ. Artif. Intell.*, 5, 100152, 2023.
- [15] S. M. Qadhi, A. Alduais, Y. Chaaban, M. Khraisheh, "Generative AI, Research Ethics, and Higher Education Research: Insights from a Scientometric Analysis". *Inf.*, 15, 325, 2024.
- [16] D. C. Andrade-Girón, W. J. Marín-Rodríguez, J. Sandivar-Rosas, E. Carreño-Cisneros, E. T. Susanibar-Ramirez, M. Zúñiga-Rojas, J. Ángeles-Morales, H. Villarreal-Torres, (2024). "Generative artificial intelligence in higher education learning: A review based on academic databases". *Iberoamerican Journal of Science Measurement and Communication*, 2024.
- [17] T. Farrelly, N. Baker, "Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice". *Education Sciences*, 2023.
- [18] P. H. Kasani, K. H. Cho, J. Jang, C. Yun, "Influence of artificial intelligence and chatbots on research integrity and publication ethics". *Science Editing*, 2024.
- [19] L. Labadze, M. Grigolia, L. Machaidze, "Role of AI chatbots in education: systematic literature review". *International Journal of Educational Technology in Higher Education*, 20, 2023.
- [20] Z. Bahroun, C. Anane, V. Ahmed, A. Zacca, "Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis", *Sustainability*, 2023.
- [21] R. Franco D'Souza, M. Mathew, V. Mishra, K.M. Surapaneni, "Twelve tips for addressing ethical concerns in the implementation of artificial intelligence in medical education", *Medical Education Online*, 29, 2024.
- [22] A. Maccaro, K. Stokes, L. Statham, L. He, A. Williams, L. Pecchia, D. Piaggio, "Clearing the Fog: A Scoping Literature Review on the Ethical Issues Surrounding Artificial Intelligence-Based Medical Devices". *Journal of Personalized Medicine*, 14, 2024.

- [23] M. Khalifa, M. Albadawy, "Using artificial intelligence in academic writing and research: An essential productivity tool", *Computer Methods and Programs in Biomedicine Update* 5, 2024.
- [24] R. Solórzano, "La triangulación metodológica como herramienta para el análisis de las estrategias de comunicación en las webs universitarias latinoamericanas", *Communication & Methods*, Vol. 4, no2, ORCID: <https://orcid.org/0000-0003-0081-9123>, 2022.
- [25] A. Niloy, A. Bari, J. Sultana, R. Chowdhury, et al., "Why do students use ChatGPT? Answering through a triangulation approach", *Computers and Education: Artificial Intelligence* 6 100208, <https://doi.org/10.1016/j.caeai.2024.100208>, 2024.