International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 09 Sep. 2024, pp: 226-238 www.ijaem.net ISSN: 2395-5252

Risks of Generative Artificial Intelligence in Higher Education: A critical perspective

Rafael Mellado [1]

[1]Proffesor, Escuela de Comercio, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
Corresponding Author: Rafael Mellado

Date of Submission: 05-09-2024 Date of Acceptance: 15-09-2024

ABSTRACT: Generative artificial intelligence offers opportunities to enhance higher education, but it also presents risks and limitations that require attention. This article explores challenges such as the unreliability of generated information, biases in models, reduction of critical thinking, privacy and security concerns, and lack of transparency. It also proposes strategies for responsible use, highlighting the need for education-specific models, explainable systems, risk mitigation, data protection, and an appropriate regulatory framework. The article emphasizes a critical and ethical approach centered on the student for implementing generative AI in this context.
KEYWORDS: Generative Artificial Intelligence, Higher Education, AI Risks and Limitations.

I. INTRODUCTION

Generative artificial intelligence (GenAI) promises to transform higher education by generating content, answering questions, and adapting materials. While its potential to enrich the educational experience is significant, it is crucial to examine the risks and limitations associated with its adoption critically. This article analyzes the key challenges of implementing AI, focusing on ethical, responsible, and student-centered use. A review of the literature and case studies seeks to provide a reference framework for institutions to effectively leverage this technology, mitigating risks and addressing its limitations.

Generative artificial intelligence has emerged as a transformative technology across various fields, including education. This technology relies on models capable of generating new content, such as text and images, based on training data [1]. Advances in deep learning, such as generative adversarial networks (GANs) and autoregressive models, have enabled the development of increasingly sophisticated systems [2], [3].

Generative AI has garnered attention in education for its ability to personalize the learning experience. Language models have been used to generate explanations and responses tailored to students' knowledge levels [4], as well as to automatically create educational content such as summaries and exercises [5], [6]. It has also been employed to generate personalized feedback, provide specific guidance based on student performance [7], facilitate more individualized support, and reduce the workload for educators.

Another prominent application is the creation of immersive learning environments, where generative models develop adaptive simulations, enhancing educational experiences in fields like medicine and engineering [9].

However, challenges arise concerning the quality of generated content, the perpetuation of biases, data privacy, and the need for teacher training [10], [11]. These considerations underscore the importance of carefully assessing the risks and limitations to ensure this technology's ethical and effective use in education.

This article examines the risks and limitations associated with using generative artificial intelligence (AI) in higher education. A review of scientific literature and case studies identifies the main challenges universities face in adopting this technology. The risks include the reliability and accuracy of generated information, model biases, impact on critical thinking, and data privacy and security issues. Limitations such as the need for precise prompts, lack of validation in real-world contexts, and limited transparency in models are highlighted. The article offers recommendations for responsible use, including developing education-specific models, system explainability, and establishing an appropriate regulatory framework. In conclusion, it aims to promote an informed and responsible adoption of generative AI, leveraging its benefits while mitigating its challenges.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 09 Sep. 2024, pp: 226-238 www.ijaem.net ISSN: 2395-5252

II. APPLICATION OF GENERATIVE AI IN HIGHER EDUCATION

Generative Artificial Intelligence (IAGen) is a subset of AI that creates content such as images, music, and text by learning from large datasets and detecting patterns to generate results similar to those produced by humans. It is distinguished by its creative capabilities and the use of models such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) [12]. Figure 1 presents the key pillars of IAGen analyzed in this article.

Key Pillars of Generative AI

The design principles for generative artificial intelligence (IAGen) applications emphasize managing generative variability, ensuring multiple outcomes, fostering exploration with control, and providing clear mental models for users [13]. In medicine, IAGen promises improvements in diagnosis, treatment planning, and disease monitoring but raises concerns about trust, privacy, and regulatory issues [14], [15]. In the business domain, IAGen holds potential for positive use cases but also for malicious misuse, necessitating careful ethical scrutiny [16].

One key challenge is intellectual property, as IAGen raises questions about using training data and the potential displacement of original creative works [17]. Additionally, it impacts IT professionals, changing their roles and required skills as content creation becomes automated [18]. In terms of advances, IAGen has shown growth from unimodal to multimodal applications [19], and has progressed in brain imaging and network computing, presenting both challenges and future research directions [20].

In computational social sciences, IAGen has lowered barriers to entry, improving productivity and educational tools [21]. IAGen is being integrated into university environments in education, generating content and interactive experiences that can enhance learning. It can personalize experiences and facilitate adaptive assessments, improving knowledge retention [22]. However, it also presents challenges such as academic integrity and adapting curricula to digital literacy [23].

In medical education, IAGen supports self-directed learning and simulation scenarios, although it faces challenges regarding data accuracy [25]. It can also offer linguistic support for international students, but raises concerns about bias and academic integrity [26]. Its growing use in educational institutions by faculty and students underscores the importance of responsible integration [29].

Finally, IAGen in primary education has demonstrated potential for motivated learning, highlighting the need for informed implementation of these technologies [30], [31].

III. RISKS OF USING GENERATIVE AI IN HIGHER EDUCATION

The reliability and accuracy of the information generated by generative artificial intelligence (GenAI) systems are critical factors in the educational context. While these tools can provide quick and seemingly coherent responses, it is essential to critically assess the quality and truthfulness of the information they produce. A major challenge is the ability of GenAI models to "hallucinate" or generate information that seems plausible but is incorrect or lacks empirical basis [32]. These models rely on patterns learned from large volumes of data, which can lead to statements that sound convincing but are not supported by solid evidence [33].

For example, a study found that GPT-3 generated incorrect responses in 21% of cases, even though these responses appeared coherent and well-structured [33]. This finding underscores the need to carefully verify AI-generated information before relying on it. Additionally, accuracy may be compromised by biases present in the training data [34]. If datasets contain biases or gaps in information, these can be transferred to the generated responses, perpetuating inaccuracies or partial views.

Another concern is the lack of transparency regarding the sources of information used by GenAI models [35]. Unlike traditional academic work, where references to reliable sources must back claims, GenAI systems do not typically provide clear attribution of the sources used, making it difficult to verify the information and potentially leading to the spread of unverified or incorrect data.

To address these challenges, fostering critical thinking and strong digital literacy among

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 09 Sep. 2024, pp: 226-238 www.ijaem.net ISSN: 2395-5252

students and educators using GenAI tools is crucial [36]. This includes teaching strategies for evaluating the quality and reliability of information, cross-referencing multiple sources, and verifying data before accepting it as true. Additionally, GenAI developers must work to improve the accuracy and reliability of their models and increase transparency regarding the sources of data used and the limitations of these tools [37].

The biases in the data and models used by GenAI are also a significant concern, as they can have profound implications in the educational domain. These biases can perpetuate stereotypes, reinforce prejudices, and generate discriminatory or inaccurate content. Training data biases arise when the datasets used to train models are not representative of the real world's diversity [38]. For instance, if the data primarily comes from sources written by white men, the models may learn and reproduce biased perspectives related to gender, race, or ethnicity [34].

A notable case is GPT-3, which has been shown to generate text reflecting gender and racial biases in its training data [39]. For example, when asked to complete the sentence "The two genders are," GPT-3 generated stereotypical and binary continuations such as "male and female," ignoring the existence of non-binary identities [40]. Moreover, biases can also be introduced through the architecture and learning algorithms of GenAI models [41].

To mitigate these challenges, GenAI developers must adopt responsible practices in collecting, selecting, and preparing training data [42]. It is essential to obtain diverse and representative datasets and apply processing techniques to reduce biases, such as removing explicit stereotypes or the oversampling of underrepresented groups. Additionally, metrics and tools must be developed to detect and quantify biases in GenAI models [43]. This will enable researchers and educators to more effectively evaluate the fairness and impartiality of these tools before implementing them in educational settings.

Another important strategy is to promote transparency and explainability in GenAI models [44]. Developers should provide clear information about the data used to train the models, potential limitations and biases, and measures to mitigate them. This transparency will allow educators and students to make informed decisions about using these tools and critically evaluate their results. Moreover, from an educational perspective, it is essential to incorporate discussions about biases and equity into curricula related to GenAI [45].

The use of GenAI in education also poses the risk of over-reliance on these tools, which can reduce students' critical thinking skills. Excessive dependence on AI can lead students to accept the generated content without questioning its accuracy, assuming it is correct and reliable simply because it comes from an advanced technological source [11]. This over-reliance may cause students to forgo their critical judgment and verification of information, making them more vulnerable to misinformation and inaccuracies.

Additionally, the excessive use of GenAI can create a false sense of knowledge and competence [46]. By quickly obtaining generated responses, students may feel that they have learned and understood the material, when in reality, they have only accessed information without fully processing or internalizing the concepts. This illusion of knowledge can hinder deep learning and applying knowledge in different contexts.

Another risk is that dependence on GenAI may reduce students' motivation to engage in the learning process [47] actively. If students perceive that they can quickly and easily obtain answers and solutions through AI, they may lose interest in the cognitive effort required for authentic learning. This could lead to a decline in curiosity, initiative, and perseverance—essential skills for lifelong learning.

To address these challenges, it is important to foster a critical mindset toward GenAI in education [48]. Educators should emphasize the importance of questioning and verifying information, regardless of its origin. This includes teaching students how to assess the quality and credibility of sources, seek additional evidence, and cross-reference multiple perspectives.

Furthermore, it is crucial to design activities and assessments that require critical thinking and deep cognitive processing [49]. Instead of solely relying on GenAI, educators should promote active learning, problem-solving, and the application of knowledge in authentic contexts. This will help students develop reasoning, analysis, and evaluation skills rather than simply reproducing AI-generated information.

Another important strategy is to promote metacognition and self-regulation in students [50]. This involves helping students reflect on their learning, monitor their understanding, and adjust their strategies accordingly. Educators can help students become more critical and autonomous in their learning by fostering metacognitive awareness, even when using GenAI tools.

The use of GenAI in education also raises significant concerns regarding the privacy and security of students' and educators' data. These tools often require access to large amounts of personal information, such as demographic data, academic records, student work, and online communications [51]. If this data is not handled properly and securely, there is a risk of misuse, unauthorized sharing, or access by third parties [52], [53].

A major issue is the opacity and security risks associated with GenAI systems. On the one hand, the opaque nature of these systems makes it difficult for students and educators to see how their personal data is handled, raising concerns about a lack of control and informed consent [54]. On the other hand, GenAI systems are vulnerable to cyberattacks that could compromise the security of user data [55], [56].

To address these challenges, educational institutions and GenAI providers must implement robust privacy and security measures, including clear policies, protective technologies, and digital security education [57], [58]. Additionally, from a regulatory perspective, solid legal and ethical frameworks must be established to address privacy and security issues related to GenAI in education [59].

In summary, using GenAI in higher education offers significant opportunities but also presents considerable challenges related to information accuracy, biases, data privacy, and the reduction of critical thinking. Addressing these challenges requires collaboration between developers, educators, and AI experts to ensure these technologies' responsible and effective use.

IV. LIMITATIONS OF GENAI IN HIGHER EDUCATION

The effectiveness of generative AI (GenAI) in education largely depends on users' ability to provide clear and precise prompts. Since these systems rely on natural language processing, the way requests are phrased significantly influences the quality and relevance of the generated responses. A key limitation of GenAI is its dependence on user instructions [66]. Ambiguous, incomplete, or poorly formulated prompts may result in irrelevant or inadequate responses. For instance, if a teacher asks GenAI to create evaluation questions without specifying the difficulty level, key concepts, or format, the generated questions may not align with specific learning objectives or accurately assess student understanding.

A lack of prompt precision can also lead to inconsistent or contradictory responses from GenAI [67]. Vague instructions may produce incoherent content or conflicting information, confusing students and undermining trust in the tool. To address this, educators must develop the skills to formulate clear and complete instructions that align with learning objectives. Breaking complex requests into smaller, more specific components improves the likelihood of receiving high-quality responses [67]. For example, instead of asking for a general summary, a teacher should specify, "Generate a 300-word summary of Chapter X focusing on concepts A, B, and C, and provide relevant examples." This reduces ambiguity and improves the relevance and quality of the output.

The need for precise prompts is crucial to leverage GenAI's potential in education fully. Educational institutions should train teachers and students to interact effectively with these tools by crafting well-structured instructions. This will maximize the educational benefits of GenAI by producing more relevant, coherent, and high-quality content that supports the learning process.

Another significant limitation of GenAI in education is the lack of validation in authentic educational contexts. Most studies on GenAI have been conducted in controlled environments or with limited datasets, which do not reflect the complexity and diversity of real-world educational settings [11]. This gap between expected and actual performance can limit understanding GenAI's long-term effects on learning outcomes [69], [70]. To overcome this limitation, more field studies are needed to evaluate GenAI in diverse, authentic educational environments, considering factors such as student diversity, teaching styles, resource limitations, and specific educational goals [71], [72].

Collaboration between researchers, educators, and technology developers is essential to design studies that capture the complexity of real-world educational settings [73]. These studies should inform the continuous improvement of GenAI tools. From an ethical and equity perspective, validating GenAI in diverse contexts is crucial to ensure inclusivity and prevent exacerbating existing educational inequalities [10]. Addressing this limitation through comprehensive research will lead to more adaptable, effective, and equitable tools that benefit all students in various educational environments.

The lack of transparency in designing and developing GenAI systems presents significant ethical and practical challenges in education. Many GenAI systems function as "black boxes," where the internal processes, algorithms, and data used to generate content are not fully transparent or accessible to users, including educators and students [74]. This opacity raises concerns about the reliability, fairness, and accountability of GenAI in education, making it difficult to assess the generated content's quality, accuracy, and impartiality [75]. Hidden biases, such as gender, racial, or cultural, may also be embedded in AI-generated educational materials [76].

Furthermore, the lack of transparency impedes educators' ability to adapt and customize GenAI effectively to the specific needs of their students [77]. It also complicates identifying and correcting errors or inaccuracies in the generated content [58]. To address this limitation, developers must prioritize transparency and explainability in designing GenAI systems by providing clear and accessible information about the algorithms, training data, and decision-making processes [78]. Establishing ethical standards for GenAI development in education, focusing on transparency, fairness, and accountability, is also essential [10].

Encouraging educators and students to actively participate in the design and development of GenAI can improve transparency and increase the likelihood that these tools will be effectively adapted to specific educational contexts [72]. From a policy perspective, clear regulations and guidelines for using GenAI in education should be developed, requiring transparency, auditing, and accountability from developers and providers [79]. A strong regulatory framework will foster a more transparent and responsible ecosystem for GenAI in education.

V. RECOMMENDATIONS FOR RESPONSIBLE USE

Developing generative AI models specifically for education, known as "EdGPT," is a promising strategy to address the limitations of general-purpose models like GPT-3 in educational settings. Unlike general models, which are trained on vast, unstructured data, EdGPT models are built using carefully selected, verified educational resources, making them more effective and tailored for educational use [80], [81]. For instance, an EdGPT model designed for teaching high school mathematics would be trained using textbooks, curriculum guides, and problem-solving examples, allowing it to develop a deep and contextual understanding of mathematical concepts [81], [82].

Beyond specialized training data, EdGPT models benefit from integrating pedagogical strategies into their design by collaborating with education experts to incorporate sound educational principles into the model's architecture [83]. For example, an EdGPT model for writing instruction could include scaffolding strategies, formative feedback, and adaptation based on student skill levels, offering more personalized support aligned with best teaching practices. It is also essential to adapt EdGPT models to various educational contexts, including different educational levels, subjects, student populations, and learning environments [84], [85].

Rigorous evaluation and validation in real educational contexts are critical to ensuring the quality and effectiveness of EdGPT models. Pilot studies, user data collection, and learning outcome analysis should be conducted to identify areas for improvement and refine the models iteratively [76]. Involving teachers and students in this continuous improvement process ensures that EdGPT models meet the educational community's expectations and needs.

Ethical and equity considerations must also be integrated into developingEdGPT models, addressing issues such as data privacy, decision-making transparency, bias mitigation, and accessibility for all students [10]. Prioritizing these ethical principles in EdGPT design will promote responsible and equitable use of generative AI in education. In conclusion, developing EdGPT models tailored to educational needs is a promising strategy for harnessing the potential of generative AI while ensuring personalized and effective support. Addressing ethical and equity concerns is essential to achieving responsible use.

To leverage the potential of generative AI in education effectively and responsibly, it is crucial to implement strategies that mitigate risks, such as information reliability, bias, privacy, security, and equity [86]. One key strategy is establishing rigorous processes for verifying and validating AI-generated information and developing mechanisms to assess accuracy, relevance, and appropriateness before use in educational contexts. Peer review systems involving experts and educators could be implemented to review and validate the quality and relevance of generated materials [87], [88].

Bias detection and mitigation in AI models is another critical strategy. Tools and techniques should be employed to identify and quantify biases in training data and generated outputs, with methods in place to correct or compensate for these biases, such as data preprocessing techniques and post-processing methods [38], [89], [90]. Privacy and data security are critical concerns, requiring robust data protection policies and practices, including technical measures like encryption, access control, and anonymization. Clear informed consent policies and ethical data use practices should be established, alongside training for educators and students on secure data handling [91], [92], [93].

Another essential strategy is promoting transparency and explainability in generative AI models, providing mechanisms for users to understand how models work, what data they use, and how they produce results. User interfaces with clear explanations of internal processes, visualization tools to explore and analyze underlying data, and explanation techniques to interpret model outputs will help users trust and rely on AI-generated content [78], [94], [95].

Equity and inclusion are also fundamental, ensuring that the benefits of generative AI are accessible and equitable for all students. Models and applications must be designed to be inclusive and adaptable to diverse needs and contexts, addressing digital divides and access barriers. Engaging the educational community in developing and implementing risk mitigation strategies will help ensure fair access to AI's educational benefits [11], [75], [86].

Teacher training and professional development are essential for mitigating the risks of using generative AI in education. Educators must have the skills, knowledge, and resources to use AI technologies effectively and responsibly in their teaching practices. Training programs should cover AI fundamentals, educational applications, ethical considerations, and risk mitigation strategies. Collaboration and knowledge-sharing among educators will also help develop effective risk mitigation strategies tailored to specific educational contexts [81], [96].

Implementing these strategies comprehensively and adapting them to specific educational settings will allow the opportunities offered by generative AI to be realized while minimizing associated risks, ensuring a positive impact on learning and teaching.

Explainability is a critical aspect of using generative AI models responsibly in education. It allows users to understand how a model makes decisions and generates results, fostering trust and informed use [97]. In education, explainability is particularly important for teachers and students to evaluate the quality, relevance, and alignment of AI-generated content with learning objectives [98]. Teachers can maintain control over the teaching process by understanding how AI-generated content is produced, ensuring it aligns with educational goals and intervening as necessary [71], [72], [96].

Explainability also promotes student trust and acceptance of AI-generated content. When students understand how content is generated and what factors influence the model's responses, they are likelier to trust and use the information effectively for learning [99]. Conversely, if AI models are perceived as "black boxes," students may be skeptical and reluctant to accept the generated content [100]. Moreover, explainability is essential for identifying and addressing biases or errors in AI-generated content, as understanding how problematic outputs were produced is necessary to correct them [38], [101].

Achieving explainability requires developing techniques and approaches that enhance transparency in AI models, such as visualization tools to show relationships between inputs and outputs, natural language explanations of key decision factors, and sensitivity analyses to demonstrate how input changes affect outputs [94], [96], [102]. Involving teachers and students in designing explainable AI models ensures that the system is intuitive, transparent, and suited to educational contexts.

VI. CONCLUSIONS

This article has explored the risks and limitations associated with using generative AI in higher education, ranging from technical issues to ethical and social concerns. Key risks include the unreliability and inaccuracy of AI-generated information due to limitations in training data and inherent algorithmic biases. This can spread misinformation, confusing students and undermining the quality and integrity of the educational process. Additionally, the biases present in the data and models can amplify existing societal inequalities and stereotypes, perpetuating discrimination and exclusion in education if not adequately addressed. Over-reliance on generative AI is another significant risk, potentially reducing critical thinking among students and negatively affecting their ability to learn meaningfully and apply knowledge in real-world contexts.

Privacy and security concerns are also critical, as collecting, storing, and processing vast amounts of student data raise questions about privacy protection, informed consent, and ethical data use. Security breaches and unauthorized access to sensitive student information are also risks. The lack of transparency and explainability in AI model design presents another limitation, as it can undermine user trust, hinder error detection and correction, and complicate accountability for using these technologies in education. Additionally, the limited involvement of the educational community in the development and implementation of generative AI can lead to resistance and distrust. At the same time, the absence of appropriate regulatory frameworks may result in inconsistent practices, privacy violations, and other negative impacts on students and society.

In summary, synthesizing identified risks and limitations highlights the complexity and challenges of using generative AI in higher education. Active engagement from the educational community, the development of appropriate regulatory frameworks, and the promotion of ethical and transparent practices are essential to addressing these challenges and responsibly harnessing the potential of generative AI in education.

Given the numerous risks and limitations, all stakeholders must unite in a call to action to ensure these technologies' ethical and responsible use. Higher education institutions should establish clear ethical principles and institutional policies grounded in values like transparency, equity, privacy, security, and accountability, developed collaboratively with active input from the entire educational community. Investing in teacher training and professional development is critical to equip educators with the skills and knowledge necessary to use generative AI effectively and ethically in their teaching practices. Additionally, students should be empowered with resources and guidance to develop digital literacy and critical thinking skills for using AI responsibly in their learning.

Technology developers must also adhere to human-centered design principles and work closely with educators and researchers to create transparent, explainable systems tailored to educational needs and contexts. Multidisciplinary collaborations between academia, industry, and government are crucial for addressing the ethical and social challenges posed by generative AI in education, fostering research, and the development of best practices, standards, and regulations.

Furthermore, promoting transparency and accountability in the use of generative AI in higher education is essential. Institutions should be open about how these technologies are used, what data is being collected, and how AI-based decisions are made. Clear oversight, auditing, and remediation mechanisms of potential negative impacts must also be established.

Lastly, fostering public participation and open dialogue on the ethical and social implications of generative AI in education is vital, creating spaces for all stakeholders to voice concerns and engage in decision-making about its use. Developing and enforcing appropriate regulations and public policies, including laws on data privacy, algorithmic transparency, and accountability, is also crucial for ensuring ethical AI use in education.

In conclusion, this collective call to action requires continuous collaboration, reflection, and commitment from all stakeholders to ensure that generative AI is used to enhance education while minimizing risks and promoting the well-being of all involved.

REFERENCES

[1]. J. Gao, M. Galley, y L. Li, «Neural Approaches to Conversational AI», Found. Trends® Inf. Retr., vol. 13, n.o 2-3, pp. 127-298, feb. 2019, doi: 10.1561/1500000074.

[2]. I. J. Goodfellow et al., «Generative Adversarial Networks», 10 de junio de 2014, arXiv: arXiv:1406.2661. doi: 10.48550/arXiv.1406.2661.

[3]. A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, y I. Sutskever, «Language Models are Unsupervised Multitask Learners», 2019. Accedido: 12 de mayo de 2024. [En línea]. Disponible en: https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe

[4]. J. J. Williams et al., «AXIS: Generating Explanations at Scale with Learnersourcing and Machine Learning», en Proceedings of the Third (2016) ACM Conference on Learning @ Scale, en L@S '16. New York, NY, USA: Association for Computing Machinery, abr. 2016, pp. 379-388. doi: 10.1145/2876034.2876042.

[5]. Q. U. Ain, M. A. Chatti, K. G. C. Bakar, S. Joarder, y R. Alatrash, «Automatic Construction of Educational Knowledge Graphs: A Word Embedding-Based Approach», Information, vol. 14, n.o 10, Art. n.o 10, oct. 2023, doi: 10.3390/info14100526.

[6]. Q. Wang, S. Yu, y X. Wang, «Research on the Role of an Automatic Resource Generation System to Promote Chinese as a Second Language Learners' Learning in Colleges», J. Educ. Comput. Res., vol. 62, n.o 2, pp. 501-531, abr. 2024, doi: 10.1177/07356331231210940.

[7]. A. E. Elshafey, M. R. Anany, A. S. Mohamed, N. Sakr, y S. G. Aly, «Dr. Proctor: A Multi-modal AI-Based Platform for Remote Proctoring

in Education», en Artificial Intelligence in Education, I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, y V. Dimitrova, Eds., Cham: Springer International Publishing, 2021, pp. 145-150. doi: 10.1007/978-3-030-78270-2_26.

[8]. Y. Ehara, «Interpreting Neural CWI Classifiers' Weights as Vocabulary Size», en Proceedings of the Fifteenth Workshop on Innovative Use of NLP for Building Educational Applications, J. Burstein, E. Kochmar, C. Leacock, N. Madnani, I. Pilán, H. Yannakoudakis, y T. Zesch, Eds., Seattle, WA, USA → Online: Association for Computational Linguistics, jul. 2020, pp. 171-176. doi: 10.18653/v1/2020.bea-1.17.

[9]. N. Partarakis y X. Zabulis, «A Review of Immersive Technologies, Knowledge Representation, and AI for Human-Centered Digital Experiences», Electronics, vol. 13, n.o 2, Art. n.o 2, ene. 2024, doi: 10.3390/electronics13020269.

[10]. W. Holmes et al., «Ethics of AI in Education: Towards a Community-Wide Framework», Int. J. Artif. Intell. Educ., vol. 32, n.o 3, pp. 504-526, sep. 2022, doi: 10.1007/s40593-021-00239-1.

[11]. O. Zawacki-Richter, V. I. Marín, M. Bond, y F. Gouverneur, «Systematic review of research on artificial intelligence applications in higher education – where are the educators?», Int. J. Educ. Technol. High. Educ., vol. 16, n.o 1, p. 39, oct. 2019, doi: 10.1186/s41239-019-0171-0.

[12]. Sanhita Kar, Chayantika Roy, Meghna Das, SoumitaMullick, y Rupa Saha, «AI Horizons: Unveiling the Future of Generative Intelligence», Int. J. Adv. Res. Sci. Commun. Technol., pp. 387-391, sep. 2023, doi: 10.48175/IJARSCT-12969.

[13]. J. D. Weisz, M. Muller, J. He, y S. Houde, «Toward General Design Principles for Generative AI Applications», 13 de enero de 2023, arXiv: arXiv:2301.05578. doi: 10.48550/arXiv.2301.05578.

[14]. T. S. Musalamadugu y H. Kannan, «Generative AI for medical imaging analysis and applications», Future Med. AI, vol. 1, n.o 2, p. FMAI5, dic. 2023, doi: 10.2217/fmai-2023-0004.

[15]. P. Zhang y M. N. Kamel Boulos, «Generative AI in Medicine and Healthcare: Promises, Opportunities and Challenges», Future Internet, vol. 15, n.o 9, Art. n.o 9, sep. 2023, doi: 10.3390/fi15090286.

[16]. S. Houde et al., «Business (mis)Use Cases of Generative AI», ArXiv, mar. 2020, Accedido: 12 de mayo de 2024. [En línea]. Disponible en: https://www.semanticscholar.org/paper/Business-(mis)Use-Cases-of-Generative-AI-Houde-Liao/274a7839ca00ed0291d904b9ca6ee9bca9d8ad6d

[17]. P. Samuelson, «Generative AI meets copyright», Science, vol. 381, n.o 6654, pp. 158-161, jul. 2023, doi: 10.1126/science.adi0656.

[18]. V. K. Nhavkar, «Impact of Generative AI on IT Professionals», Int. J. Res. Appl. Sci. Eng. Technol., vol. 11, n.o 7, pp. 15-18, jul. 2023, doi: 10.22214/ijraset.2023.54515.

[19]. R. Gozalo-Brizuela y E. C. Garrido-Merchán, «A survey of Generative AI Applications», 14 de junio de 2023, arXiv: arXiv:2306.02781. doi: 10.48550/arXiv.2306.02781.C. Gong et al., «Generative AI for brain image

[20]. computing and brain network computing: a review», Front. Neurosci., vol. 17, jun. 2023, doi: 10.3389/fnins.2023.1203104.

[21]. Y. Zhang, «Generative AI has lowered the barriers to computational social sciences», 17 de noviembre de 2023, arXiv: arXiv:2311.10833. doi: 10.48550/arXiv.2311.10833.

[22]. K. Kadaruddin, «Empowering Education through Generative AI: Innovative Instructional Strategies for Tomorrow's Learners», Int. J. Bus. Law Educ., vol. 4, n.o 2, Art. n.o 2, ago. 2023, doi: 10.56442/ijble.v4i2.215.

[23]. K. Walczak y W. Cellary, «Challenges for higher education in the era of widespread access to Generative AI», Econ. Bus. Rev., vol. 9, pp. 71-100, 2023, doi: 10.18559/ebr.2023.2.743.

[24]. C. Stone, «Artificial intelligence in social work practice education. The potential use of Generative AI for learning», J. Pract. Teach. Learn., vol. 20, n.o 3, Art. n.o 3, oct. 2023, doi: 10.1921/jpts.v20i3.2192.

[25]. C. Preiksaitis y C. Rose, «Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review», JMIR Med. Educ., vol. 9, n.o 1, p. e48785, oct. 2023, doi: 10.2196/48785.

[26]. T. Farrelly y N. Baker, «Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice», Educ. Sci., vol. 13, n.o 11, Art. n.o 11, nov. 2023, doi: 10.3390/educsci13111109.

[27]. A. Olga et al., «Generative AI: Implications and Applications for Education», 22 de mayo de 2023, arXiv: arXiv:2305.07605. doi: 10.48550/arXiv.2305.07605.

[28]. M. R. Morris, «Scientists' Perspectives on the Potential for Generative AI in their Fields», 3 de abril de 2023, arXiv: arXiv:2304.01420. doi: 10.48550/arXiv.2304.01420.

[29]. S. Yeralan y L. A. Lee, «Generative AI: Challenges to higher education», Sustain. Eng. Innov., vol. 5, n.o 2, Art. n.o 2, sep. 2023, doi: 10.37868/sei.v5i2.id196.

[30]. J. S. Jauhiainen y A. G. Guerra, «Generative AI and ChatGPT in School Children's Education: Evidence from a School Lesson», Sustainability, vol. 15, n.o 18, Art. n.o 18, ene. 2023, doi: 10.3390/su151814025.

[31]. M. AL-Smadi, «ChatGPT and Beyond: The Generative AI Revolution in Education», 26 de noviembre de 2023, arXiv: arXiv:2311.15198. doi: 10.48550/arXiv.2311.15198.

[32]. L. De Angelis et al., «ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health», Front. Public Health, vol. 11, p. 1166120, abr. 2023, doi: 10.3389/fpubh.2023.1166120.

[33]. O. Evans et al., «Truthful AI: Developing and governing AI that does not lie», 13 de octubre de 2021, arXiv: arXiv:2110.06674. doi: 10.48550/arXiv.2110.06674.

[34]. E. M. Bender, T. Gebru, A. McMillan-Major, y S. Shmitchell, «On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? □», en Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, enFAccT '21. New York, NY, USA: Association for Computing Machinery, mar. 2021, pp. 610-623. doi: 10.1145/3442188.3445922.

[35]. E. Dimitriadou y A. Lanitis, «A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms», Smart Learn. Environ., vol. 10, n.o 1, p. 12, feb. 2023, doi: 10.1186/s40561-023-00231-3.

[36]. K. H. D. Tang, «Implications of Artificial Intelligence for Teaching and Learning», Acta Pedagog. Asiana, vol. 3, n.o 2, Art. n.o 2, feb. 2024, doi: 10.53623/apga.v3i2.404.

[37]. D. Dillion, N. Tandon, Y. Gu, y K. Gray, «Can AI language models replace human participants?», Trends Cogn. Sci., vol. 27, n.o 7, pp. 597-600, jul. 2023, doi: 10.1016/j.tics.2023.04.008.

[38]. N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, y A. Galstyan, «A Survey on Bias and Fairness in Machine Learning», ACM Comput. Surv., vol. 54, n.o 6, p. 115:1-115:35, jul. 2021, doi: 10.1145/3457607.

[39]. T. B. Brown et al., «Language Models are Few-Shot Learners», 22 de julio de 2020, arXiv: arXiv:2005.14165. doi: 10.48550/arXiv.2005.14165.

[40]. M. Nadeem, A. Bethke, y S. Reddy, «StereoSet: Measuring stereotypical bias in pretrained language models», en Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), C. Zong, F. Xia, W. Li, y R. Navigli, Eds., Online: Association for Computational Linguistics, ago. 2021, pp. 5356-5371. doi: 10.18653/v1/2021.acl-long.416.

[41]. H. R. Saeidnia, «Ethical artificial intelligence (AI): confronting bias and discrimination in the library and information industry», Libr. Hi Tech News, vol. ahead-of-print, n.o ahead-of-print, ene. 2023, doi: 10.1108/LHTN-10-2023-0182.

[42]. T. Gebru et al., «Datasheets for datasets», Commun. ACM, vol. 64, n.o 12, pp. 86-92, nov. 2021, doi: 10.1145/3458723.

[43]. C. Bartneck, C. Lütge, A. Wagner, y S. Welsh, «Trust and Fairness in AI Systems», en An Introduction to Ethics in Robotics and AI, C. Bartneck, C. Lütge, A. Wagner, y S. Welsh, Eds., Cham: Springer International Publishing, 2021, pp. 27-38. doi: 10.1007/978-3-030-51110-4_4.

[44]. D. Carneiro y P. Veloso, «Ethics, Transparency, Fairness and the Responsibility of Artificial Intelligence»,

en New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence, J. F. de Paz Santana, D. H. de la Iglesia, y A. J. López Rivero, Eds., Cham: Springer International Publishing, 2022, pp. 109-120. doi: 10.1007/978-3-030-87687-6_12.

[45]. L. N. Onyejegbu, «Challenges of Integrating AI Ethics into Higher Education Curricula in West Africa: Nigerian Universities Narrative», en AI Ethics in Higher Education: Insights from Africa and Beyond, C. C. Corrigan, S. A. Asakipaam, J. J. Kponyo, y C. Luetge, Eds., Cham: Springer International Publishing, 2023, pp. 57-66. doi: 10.1007/978-3-031-23035-6_5.

[46]. T. Li, E. Reigh, P. He, y E. Adah Miller, «Can we and should we use artificial intelligence for formative assessment in science?», J. Res. Sci. Teach., vol. 60, n.o 6, pp. 1385-1389, 2023, doi: 10.1002/tea.21867.

[47]. P. Kaledio, A. Robert, y L. Frank, «The Impact of Artificial Intelligence on Students' Learning Experience», 1 de febrero de 2024, Rochester, NY: 4716747. doi: 10.2139/ssrn.4716747.

[48]. J. C. Sánchez-Prieto, F. Huang, S. Olmos-Migueláñez, F. J. García-Peñalvo, y T. Teo, «Exploring the unknown: The effect of resistance to change and attachment on mobile adoption among secondary pre-service teachers», Br. J. Educ. Technol., vol. 50, n.o 5, pp. 2433-2449, 2019, doi: 10.1111/bjet.12822.

[49]. L. Theodoulides, M. Niklová, K. Liptáková, G. Nafoussi, y M. Haviar, Fostering critical thinking in higher education through a coaching approach: theory and practical applications. Belianum. Matej Bel University Press, 2020. Accedido: 13 de mayo de 2024. [En línea]. Disponible en: https://repo.umb.sk/xmlui/handle/123456789/97

[50]. L. Anthonysamy, A. C. Koo, y S. H. Hew, «Self-regulated learning strategies in higher education: Fostering digital literacy for sustainable lifelong learning», Educ. Inf. Technol., vol. 25, n.o 4, pp. 2393-2414, jul. 2020, doi: 10.1007/s10639-020-10201-8.

[51]. L. Huang, «Ethics of Artificial Intelligence in Education: Student Privacy and Data Protection», Sci. Insights Educ. Front., vol. 16, n.o 2, pp. 2577-2587, jun. 2023, doi: 10.15354/sief.23.re202.

[52]. D. Amo Filvá et al., «Local technology to enhance data privacy and security in educational technology», Int. J. Interact. Multimed. Artif. Intell., vol. 7, n.o 2, pp. 262-273, 2021, doi: 10.9781/ijimai.2021.11.006.

[53]. A. Sattikar y R. Kulkarni, «A Role of Artificial Intelligence Techniques in Security and Privacy Issues of Social Networking», vol. 2, n.o 1, 2012.

[54]. X. Liu et al., «Privacy and Security Issues in Deep Learning: A Survey», IEEE Access, vol. 9, pp. 4566-4593, 2021, doi: 10.1109/ACCESS.2020.3045078.

[55]. T. Mazhar et al., «Analysis of IoT Security Challenges and Its Solutions Using Artificial Intelligence», Brain Sci., vol. 13, n.o 4, Art. n.o 4, abr. 2023, doi: 10.3390/brainsci13040683.

[56]. B. K. Mohanta, D. Jena, U. Satapathy, y S. Patnaik, «Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology», Internet Things, vol. 11, p. 100227, sep. 2020, doi: 10.1016/j.iot.2020.100227.

[57]. B. Karan y G. R. Angadi, «Potential Risks of Artificial Intelligence Integration into School Education: A Systematic Review», Bull. Sci. Technol. Soc., vol. 43, n.o 3-4, pp. 67-85, dic. 2023, doi: 10.1177/02704676231224705.

[58]. B. Memarian y T. Doleck, «Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review», Comput. Educ. Artif. Intell., vol. 5, p. 100152, ene. 2023, doi: 10.1016/j.caeai.2023.100152.

[59]. L. Floridi, The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford University Press, 2023. doi: 10.1093/oso/9780198883098.001.0001.

[60]. M. A. Al-Asadi y S. Tasdemir, «Using Artificial Intelligence Against the Phenomenon of Fake News: A Systematic Literature Review», en Combating Fake News with Computational Intelligence Techniques, M. Lahby, A.-S. K. Pathan, Y. Maleh, y W. M. S. Yafooz, Eds., Cham: Springer International Publishing, 2022, pp. 39-54. doi: 10.1007/978-3-030-90087-8_2.

[61]. A. D'Ulizia, M. C. Caschera, F. Ferri, y P. Grifoni, «Fake news detection: a survey of evaluation datasets», PeerJComput. Sci., vol. 7, p. e518, jun. 2021, doi: 10.7717/peerj-cs.518.

[62]. Z. Cai, Q. Fan, R. S. Feris, y N. Vasconcelos, «A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection», en Computer Vision – ECCV 2016, B. Leibe, J. Matas, N. Sebe, y M. Welling, Eds., Cham: Springer International Publishing, 2016, pp. 354-370. doi: 10.1007/978-3-319-46493-0_22.

[63]. T. Raha et al., «Neural Models for Factual Inconsistency Classification with Explanations», en Machine Learning and Knowledge Discovery in Databases: Research Track, D. Koutra, C. Plant, M. Gomez Rodriguez, E. Baralis, y F. Bonchi, Eds., Cham: Springer Nature Switzerland, 2023, pp. 410-427. doi: 10.1007/978-3-031-43418-1_25.

[64]. H. Luan et al., «Challenges and Future Directions of Big Data and Artificial Intelligence in Education», Front. Psychol., vol. 11, oct. 2020, doi: 10.3389/fpsyg.2020.580820.

[65]. F. Pedro, M. Subosa, A. Rivas, y P. Valverde, «Artificial intelligence in education : challenges and opportunities for sustainable development», Minist. Educ., 2019, Accedido: 13 de mayo de 2024. [En línea]. Disponible en: https://repositorio.minedu.gob.pe/handle/20.500.12799/6533

[66]. M. Yasunaga, J. Leskovec, y P. Liang, «LM-Critic: Language Models for Unsupervised Grammatical Error Correction», 7 de octubrede 2021, arXiv: arXiv:2109.06822. doi: 10.48550/arXiv.2109.06822.

[67]. W. Wang, P. Li, y H.-T. Zheng, «Consistency and Coherency Enhanced Story Generation», en Advances in Information Retrieval, D. Hiemstra, M.-F. Moens, J. Mothe, R. Perego, M. Potthast, y F. Sebastiani, Eds., Cham: Springer International Publishing, 2021, pp. 694-709. doi: 10.1007/978-3-030-72113-8_46.

[68]. C. Wang, Z. Li, T. Chen, R. Wang, y Z. Ju, «Research on the Application of Prompt Learning Pretrained Language Model in Machine Translation Task with Reinforcement Learning», Electronics, vol. 12, n.o 16, Art. n.o 16, ene. 2023, doi: 10.3390/electronics12163391.

[69]. B. Cope, M. Kalantzis, y D. Searsmith, «Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies», Educ. Philos. Theory, vol. 53, n.o 12, pp. 1229-1245, oct. 2021, doi: 10.1080/00131857.2020.1728732.

[70]. I. Roll y R. Wylie, «Evolution and Revolution in Artificial Intelligence in Education», Int. J. Artif. Intell. Educ., vol. 26, n.o 2, pp. 582-599, jun. 2016, doi: 10.1007/s40593-016-0110-3.

[71]. R. S. Baker, «Stupid Tutoring Systems, Intelligent Humans», Int. J. Artif. Intell. Educ., vol. 26, n.o 2, pp. 600-614, jun. 2016, doi: 10.1007/s40593-016-0105-0.

[72]. K. Holstein, B. M. McLaren, y V. Aleven, «Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity», J. Learn. Anal., vol. 6, n.o 2, Art. n.o 2, jul. 2019, doi: 10.18608/jla.2019.62.3.

[73]. C. P. Rosé, E. A. McLaughlin, R. Liu, y K. R. Koedinger, «Explanatory learner models: Why machine learning (alone) is not the answer», Br. J. Educ. Technol., vol. 50, n.o 6, pp. 2943-2958, 2019, doi: 10.1111/bjet.12858.

[74]. F. Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press, 2015. Accedido: 13 de mayo de 2024. [En línea]. Disponible en: https://www.jstor.org/stable/j.ctt13x0hch

[75]. J. Whittlestone, R. Nyrup, A. Alexandrova, y S. Cave, «The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions», en Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, en AIES '19. New York, NY, USA: Association for Computing Machinery, ene. 2019, pp. 195-200. doi: 10.1145/3306618.3314289.

[76]. R. S. Baker y A. Hawn, «Algorithmic Bias in Education», Int. J. Artif. Intell. Educ., vol. 32, n.o 4, pp. 1052-1092, dic. 2022, doi: 10.1007/s40593-021-00285-9.

[77]. K. Kitto, S. Buckingham Shum, y A. Gibson, «Embracing imperfection in learning analytics», en Proceedings of the 8th International Conference on Learning Analytics and Knowledge, en LAK '18. New York, NY, USA: Association for

Computing Machinery, mar. 2018, pp. 451-460. doi: 10.1145/3170358.3170413.

[78]. A. Barredo Arrieta et al., «Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI», Inf. Fusion, vol. 58, pp. 82-115, jun. 2020, doi: 10.1016/j.inffus.2019.12.012.

[79]. P. Prinsloo y S. Slade, «An elephant in the learning analytics room: the obligation to act», en Proceedings of the Seventh International Learning Analytics & Knowledge Conference, en LAK '17. New York, NY, USA: Association for Computing Machinery, mar. 2017, pp. 46-55. doi: 10.1145/3027385.3027406.

[80]. S. Chaudhuri, U. Dayal, y V. Narasayya, «An overview of business intelligence technology», Commun. ACM, vol. 54, n.o 8, pp. 88-98, ago. 2011, doi: 10.1145/1978542.1978562.

[81]. B. P. Woolf, H. C. Lane, V. K. Chaudhri, y J. L. Kolodner, «AI Grand Challenges for Education», AI Mag., vol. 34, n.o 4, pp. 66-84, 2013, doi: 10.1609/aimag.v34i4.2490.

[82]. E. Mousavinasab, N. Zarifsanaiey, S. R. NiakanKalhori, M. Rakhshan, L. Keikha, y M. Ghazi Saeedi, «Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods», Interact. Learn. Environ., vol. 29, n.o 1, pp. 142-163, ene. 2021, doi: 10.1080/10494820.2018.1558257.

[83]. N. Nazari, M. S. Shabbir, y R. Setiawan, «Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial», Heliyon, vol. 7, n.o 5, may 2021, doi: 10.1016/j.heliyon.2021.e07014.

[84]. J. M. Gayed, M. K. J. Carlon, A. M. Oriola, y J. S. Cross, «Exploring an AI-based writing Assistant's impact on English language learners», Comput. Educ. Artif. Intell., vol. 3, p. 100055, ene. 2022, doi: 10.1016/j.caeai.2022.100055.

[85]. M. Vahid, «Design, Implementation and Evaluation of an AI-based Tutor for Persian Learners of English Language», 15 de marzo de 2024. doi: 10.21203/rs.3.rs-4055341/v1.

[86]. B. Berendt, A. Littlejohn, y M. Blakemore, «AI in education: learner choice and fundamental rights», Learn. Media Technol., vol. 45, n.o 3, pp. 312-324, jul. 2020, doi: 10.1080/17439884.2020.1786399.

[87]. A. W. (Tony) Bates, Teaching in a Digital Age - Second Edition. Tony Bates Associates Ltd., 2019. Accedido: 15 de mayo de 2024. [En línea]. Disponible en: https://pressbooks.bccampus.ca/teachinginadigitalagev2/

[88]. J. A. Kulik y J. D. Fletcher, «Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review», Rev. Educ. Res., vol. 86, n.o 1, pp. 42-78, mar. 2016, doi: 10.3102/0034654315581420.

[89]. S. A. Friedler, C. Scheidegger, S. Venkatasubramanian, S. Choudhary, E. P. Hamilton, y D. Roth, «A comparative study of fairness-enhancing interventions in machine learning», en Proceedings of the Conference on Fairness, Accountability, and Transparency, en FAT* '19. New York, NY, USA: Association for Computing Machinery, ene. 2019, pp. 329-338. doi: 10.1145/3287560.3287589.

[90]. M. B. Zafar, I. Valera, M. Gomez Rodriguez, y K. P. Gummadi, «Fairness Beyond Disparate Treatment & Disparate Impact: Learning Classification without Disparate Mistreatment», en Proceedings of the 26th International Conference on World Wide Web, en WWW '17. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, abr. 2017, pp. 1171-1180. doi: 10.1145/3038912.3052660.

[91]. J. Fjeld, N. Achten, H. Hilligoss, A. Nagy, y M. Srikumar, «Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI», 15 de enero de 2020, Rochester, NY: 3518482. doi: 10.2139/ssrn.3518482.

[92]. G. D. Hager et al., «Artificial Intelligence for Social Good», 16 de enero de 2019, arXiv: arXiv:1901.05406. doi: 10.48550/arXiv.1901.05406.

[93]. A. Nguyen, H. N. Ngo, Y. Hong, B. Dang, y B.-P. T. Nguyen, «Ethical principles for artificial intelligence in education», Educ. Inf. Technol., vol. 28, n.o 4, pp. 4221-4241, abr. 2023, doi: 10.1007/s10639-022-11316-w.

[94]. D. Gunning y D. W. Aha, «DARPA's Explainable Artificial Intelligence

Program», AI Mag., vol. 40, n.o 2, pp. 44-58, 2019, doi: 10.1609/aimag.v40i2.2850.

[95]. P. Hacker, R. Krestel, S. Grundmann, y F. Naumann, «Explainable AI under contract and tort law: legal incentives and technical challenges», Artif. Intell. Law, vol. 28, n.o 4, pp. 415-439, dic. 2020, doi: 10.1007/s10506-020-09260-6.

[96]. R. Luckin y W. Holmes, «Intelligence Unleashed: An argument for AI in Education», UCL Knowledge Lab, London, UK, Report, feb. 2016. Accedido: 15 de mayo de 2024. [En línea]. Disponible en: https://www.pearson.com/content/dam/corporate/global/pearson-dot-com/files/innovation/Intelligence-Unleashed-Publication.pdf

[97]. R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, y D. Pedreschi, «A Survey of Methods for Explaining Black Box Models», ACM Comput. Surv., vol. 51, n.o 5, p. 93:1-93:42, ago. 2018, doi: 10.1145/3236009.

[98]. C. Conati, K. Porayska-Pomsta, y M. Mavrikis, «AI in Education needs interpretable machine learning: Lessons from Open Learner Modelling», 30 de junio de 2018, arXiv: arXiv:1807.00154. doi: 10.48550/arXiv.1807.00154.

[99]. V. Putnam y C. Conati, «Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS)», Los Angel., 2019.

[100]. D. Tempelaar, B. Rienties, y Q. Nguyen, «The contribution of dispositional learning analytics to precision education», Educ. Technol. Soc., vol. 24, n.o 1, pp. 109-122, 2021.

[101]. R. F. Kizilcec y H. Lee, «Algorithmic Fairness in Education», 11 de abril de 2021, arXiv: arXiv:2007.05443. doi: 10.48550/arXiv.2007.05443.

[102]. M. T. Ribeiro, S. Singh, y C. Guestrin, «“Why Should I Trust You?”: Explaining the Predictions of Any Classifier», en Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, en KDD ’16. New York, NY, USA: Association for Computing Machinery, ago. 2016, pp. 1135-1144. doi: 10.1145/2939672.2939778.