2

Artificial Intelligence and Education: Is It Necessary, Is It Convenient?

PhD Enrique Estellés-Arolas
Universidad Católica de Valencia, Spain
enrique.estelles@ucv.es
https://orcid.org/0000-0002-6067-5064

PhD Javier Pérez Bou
Universidad Católica de Valencia, Spain
javier.perez@ucv.es
https://orcid.org/0000-0001-7298-0499

Abstract
Blockchain, cryptocurrencies and metaverse are technologies that have been all the rage in recent years. One might be tempted to add Artificial Intelligence to this group of technologies as just another fad, but unlike these, AI has been able to integrate into many areas of people's lives and find practical use cases. It was already doing so implicitly through virtual assistants (Siri, Alexa, etc.), but now it is doing so openly, with users being aware that they are using AI tools.

What is happening with AI, as has happened with other technologies throughout history, is that its supporters and detractors quickly emerge. And even more so when dealing with a subject as sensitive as education. Some tend to idealize its use, minimizing possible problems or risks, while others tend to fatalize about it and about the havoc it will cause.

Given this situation, it is worthwhile to critically analyze the advantages and disadvantages of AI as an educational tool, always asking the same question: what is in the best interest of the students?

In this book chapter we analyze different use cases and technical reports that will allow us to identify advantages, disadvantages, and good practices.

Keywords: Artificial Intelligence, education, ethics, students, teachers.

2.1. Introduction

Learning from the past

"Books will soon be obsolete in public schools", "[this technology] will make the services of the best teachers available to people" or "children are learning twice as fast as they once did, and retaining what they learn". We might think that these statements were made by technologists, experts in the field of education talking about the use of technologies such as the Internet. The reality is that they were made, in order, by Thomas Edison talking about cinema in 1913, by Benjamin Darrow (founder and principal of a school) talking about the radio in 1932, and by U.S. President Lyndon Johnson talking about the television in 1968 (Cuban, 1986; Wang & Reeves, 2003). All of them were technologies that promised great changes but failed to deliver them.

The past shows us that the history of the use of technologies in education is cyclical and tends to repeat itself. Cuban (1986) identified the structure of this cycle and divided it into 4 phases: euphoria, scientific credibility, disillusionment, and blame. In the first phase, different groups and individuals such as governments, technology companies and the so-called "evangelists" of technology (Reich, 2020), advocate the adoption of technology in the educational environment to change it and improve it in a broad and profound way. In the second phase, numerous studies, often carried out by the very companies that manufacture such technologies (Wang & Reeves, 2003; Desmuget, 2015), are conducted to find credible evidence of the effectiveness of the pedagogical applications of such technologies. The third phase basically consists in the disillusionment and frustration produced by the realization that the technologies introduced in schools do not deliver what was promised at the time. The fourth and last phase is a reaction to the latter, which consists in looking for a culprit. Cuban (1986) mentions the blaming of teachers. Nowadays, digital devices and their ineffectiveness in certain contexts are also pointed out, as in the case of Sweden and the use of computers. In this case, the Minister of Education has paralyzed the digitization plan due to the loss of11 points in the Progress in International Reading Literacy Study 2021 (PIRLS)

report, deciding to limit digital devices and reintroducing textbooks (Crace, 2023).

Interestingly, the sequence identified by Cuban (1986) is similar to that known as the "Gartner Hype Cycle" (Gartner, n.d.), which analyzes the development of fashionable technologies in different fields.

The reasons for failure in the adoption of technologies in education can be multiple and diverse, such as exaggerated expectations that are impossible to meet, lack of understanding of the educational reality or lack of necessary resources. After all, the educational environment is a complex one, where teachers, students, resources, a given socio-cultural context and a series of other elements interact, sometimes in unexpected ways (Reich, 2020). Therefore, interventions that consistently and responsibly analyze the use of technologies in the educational setting, generating evidence to support or discourage it, become necessary (Wang & Reeves, 2003). As Cuban (2018) maintains, "trying to accelerate learning by ramping up technology is like putting rockets on butterfly wings. More force does not lead linearly to more progress."

Facing the present

Currently, the technology that promises to transform education is Artificial Intelligence. It is true that its application in education is not new (Chen, Chen & Lin, 2020; Zhai et al., 2021), but its use has been boosted by recent advances in the field of Generative Artificial Intelligence.

This type of AI makes it possible to generate content (text, images, etc.) in response to a request written in natural language called "prompt". Systems that produce textual content are called LLM (Large Language Models), and GPT (Generative Pre-trained Transformer) is a particular example of these models, which are trained with large amounts of data, allowing them to capture the particularities of language and generate coherent content (Miao & Holmes, 2023).

ChatGPT in particular has substantially changed the educational landscape for two reasons. The first reason is related to the types of tasks it can perform. With a variable level of correctness, ChatGPT and other language models can perform higher-order

cognitive tasks such as elaborating complex texts or summarizing texts, which are tasks that were previously reserved for humans. This has raised legitimate concerns among teachers at all levels: from whether AI will replace them as teachers, to what to do to detect when students use these technologies dishonestly (for carrying out assignments and essays) (OTS, 2023; Miao & Holmes, 2023).

The second reason is its level of popularization. As the first LLM accessible to the general public, it reached the number of 1 million active users in only 5 days and, for example, during the first months of 2023, it had more than 100 million active users (Miao & Holmes, 2023).

Moreover, as is always the case when a technology becomes popular, it is quickly proposed as a teaching tool, thinking that its use will motivate students more in the learning process (Baek, Yung & Kim, 2008). In this sense, numerous researchers have proposed different uses of ChatGPT in education, both for teaching and learning (Ilieva et al., 2023; Kadaruddin, 2023; Lo, 2023; Liu et al., 2024; Newton & Xiromeriti, 2024). One of the most frequently cited examples is the use of ChatGPT as a personal tutor, a type of tutoring with long-established benefits (Juel, 1996). In fact, work on its automation has been underway since the late 1960s, with varying levels of success (Miao et al., 2021; Ilieva et al., 2023). However, it should be noted that there is no universally accepted system for the design, development, and implementation of AI chatbots in educational settings, nor is there robust evidence of their effectiveness (Miao et al., 2021; Miao & Holmes, 2023).

Given the situation described above, in this chapter we will identify and analyze the characteristics of any LLM that must be considered to make a coherent analysis of its use in an educational environment, to attain the maximum benefit.

2.2. Framework of Analysis

To identify and analyze the characteristics mentioned above, it is important to determine the framework of analysis that is going to be used. It is assumed that all technology has a teleological nature, i.e., it is oriented to a specific end or goal (Rescher,

1999). Logically, to achieve this end or goal as efficiently as possible, these technologies are designed in a certain way. However, this does not mean that a technology can only be used for the purpose for which it was designed. That is why it is also said that technology is ambiguous (Ortega y Gasset, 1982) in that it can be used to achieve different ends. This ambiguity, which adds versatility to a technology, implies the possible variation in its efficiency in new uses. The technology will have a maximum degree of efficiency in the task for which it was created (provided it is well designed), but when it is used to achieve other objectives, its level may vary. It will depend on the alignment of the characteristics of the technology and the requirements of the task in question.

For example, video games were designed as a means of entertainment and, although many efforts have been made to use them in education, the result has not been as good as expected or desired (Desmuguet, 2015). On the contrary, the Internet was designed for the exchange of information, not for shopping. However, given its nature and through what Ciborra (2002) called "DIY" processes, today it can be used for many other purposes such as purchasing products or contracting services. In the case of video games, there is no alignment between the technology and the new task to be performed. In the case of the Internet, there is.

In addition, the use of technologies often involves unexpected effects, which are not contemplated in their design, as it is impossible to do so, and this may make their use inadvisable in certain areas. The use of social networks, for example, implies a high degree of disinhibition. This characteristic, which was not contemplated when computer-mediated communication systems were designed, makes their use inadvisable depending on the situation (Shalom et al., 2015).

So how does ChatGPT fit into this objective analysis - ambiguity - unexpected effects scheme?

Characteristics of ChatGPT

a) Objective. The original goal for which ChatGPT was created was to mimic human conversation. Thanks to the use of different AI techniques, ChatGPT is able to produce human-like

text and maintain a conversational style, allowing for more realistic natural and comprehensible dialogues (Tlili et al., 2023). In addition, and to facilitate this goal, other features have been added, such as the so-called "persona pattern", which allows the language model to mimic personalities, characters or emotions during its interactions to facilitate communication (Parra Pennefather, 2023).

b) Ambiguity. Like any technology, ChatGPT has this characteristic. Moreover, being able to simulate a fundamental human skill such as conversation (due the relational nature of the human being), its potential applications are numerous (Kocoń et al., 2023).
c) Unexpected effects. In this aspect, both positive and negative unexpected effects have been found. As positive effects, ChatGPT can perform relatively creative tasks (by composing the knowledge it already has), such as writing poetry or making up stories. It also allows finding alternative solutions (more or less valid) in problem solving (Tlili et al., 2023). Regarding negative unexpected effects, we find several in the literature, although we highlight three for the specific field of education: hallucinations, non-determinism and the existence of biases.

The positive effects extend their versatility even further; however, the negative effects have a very important weight for the case analyzed in the present work. In the following, we will analyze these three negative unexpected effects based on the literature consulted.

2.3. Unexpected effects on language models

Hallucinations

Hallucinations are defined as the production of "content that is nonsensical or untruthful in relation to certain sources" (OpenAI, 2023). This type of erroneous content can be classified in different ways (Van Deemter, 2022; Huang et al., 2023): omissions, wrong and/or invented data, answers that do not relate to the question posed (totally or partially), or logical inconsistencies among others.

This unexpected effect is known by OpenAI, which warns about it on the ChatGPT website and recommends that this technology should be used with special care in contexts where reliability is important. OpenAI (2023) analyzed the expert evaluations of ChatGPT-4 answers in different topics and, although it improved by 19% the correct answers of its previous versions, the correctness evaluation was between 70 and 80%. This problem is also identified in other studies that recommend human intervention for the evaluation of the accuracy and consistency of the answers (Ilieva et al., 2023).

Sometimes, with the aim of minimizing these hallucinations, as well as other problems arising from the use of LLMs such as the generation of inappropriate content, different technics called guardrails have been developed (Tonmoy et al., 2024). However, these guardrails do not work securely either. Liu et al. (2024) indicate that the level of success in using ChatGPT with guardrails in a programming course varied between the different calls, going from 88% success to 39%.

Regarding the area of knowledge, different studies indicate that it does not perform equally well in all areas: ChatGPT excelled in critical and higher order thinking and economics, but its performance was low in law, medical education and mathematics. It also presents problems in identifying sentiment in messages (Kocoń et al., 2023; Lo, 2023; Newton & Xiromeriti, 2024).

Although work is being done and progress is being made on different techniques, apart from guardrails, to mitigate the appearance of these errors (Tonmoy et al., 2024), according to some authors, it is something inherent to the language models themselves and it is difficult for them to disappear (Xu, Jain & Kankanhalli, 2024).

In fact, these hallucinations also occur in EdGPTs, which are models trained on education-specific data (Miao & Holmes, 2023).

Non-determinism

The non-determinism of LLM refers to their inconsistency in their responses given the same prompt, ChatGPT, for example, provides different answers (Tlili et al., 2023). Thus, for the same

question, two learners may randomly receive different, uncomplete or even contradictory information, which goes against fair access to education (Miao & Holmes, 2023).

This non-determinism not only affects the model's responses, but also manifests itself in the blocking or not of certain requests. For example, through the aforementioned guardrails, ChatGPT should not produce inappropriate content. However, the same question at different times may sometimes produce an answer justifying the non-generation of such content, and sometimes the requested content.

Therefore, this non-determinism affects not only the users in terms of the quality of the information they receive, but also the ChatGPT usage rules themselves.

Biases

In this case, the unexpected effect is the biases presented by the models' responses. By design, they tend to amplify the hidden features of their training data, thus reinforcing the positions they represent (Miao et al., 2021). This results in the emergence of political (Fujimoto & Takemoto, 2023), sexual (Miao et al., 2021), racial (Miao & Holmes, 2023), etc. biases. Being data-dependent, removed or fixed biases may re-emerge due to model updates, thus their periodic re-evaluation is inevitable (Fujimoto & Takemoto, 2023).

One way to mitigate these biases would be to use more representative and varied data. However, most of the training data are unknown: OpenAI, for example, partially reported ChatGPT3 data (Brown et al., 2020), but not version 4 data. This is a problem, as it is thus not possible to identify potential problems due to the use of inadequate or biased data sources and implies a significant lack of transparency that affects user confidence (Miao & Holmes, 2023).

Another problem associated with biases is the use that language models make of data from interactions with their users as part of their training (Tlili et al., 2023). This practice raises issues related to data security, but in terms of biases, it again prevents an adequate control.

2.4. Discussion

Language models have a series of unexpected effects that hinder their widespread use in the educational setting. In this sense, it is necessary to differentiate between their use by teachers and by students.

As for students, the key is to find the alignment between the characteristics of the technology and the requirements of the task to be carried out. It is clear, therefore, that if a task requires a language model to provide a 100% valid, reliable, and complete answer in its content, it is not advisable to use it. The same is not true if what matters about the answer is its grammatical structure and not its content, for example. Non-determinism is a problem if a concrete and unique answer is needed (which should also be correct), but it is not a problem if what is sought is the suggestion of topics, ideas, etc., where receiving different answers does not imply a comparative aggravation. A detailed study of what tasks could be carried out based on this technology-task alignment is therefore necessary, always bearing in mind that education is based on and requires truth (Barrio Maestre, 2008).

In addition, it is important to collect evidence of the effects of the use of these models on students to be able to make conscious decisions. For example:

- Its use can make learners lazy and those who are not motivated may use it as a shortcut (Tlili et al., 2023) or fail to adequately review the information provided by the model (Qureshi, 2023).
- Many learners tend to anthropomorphize the model, eventually establishing inadequate trust relationships that break down when hallucinations and inaccuracies are identified (Tlili et al., 2023; Liu et al., 2024).
- Although invited to always have a critical view on ChatGPT and other LLMs' answers (Miao & Holmes, 2023), due to the correct, convincing, and credible expression these models use, students tend to trust without questioning the answers, thereby reducing their critical thinking (OpenAI, 2023; Tlili et al., 2023). For example, 69% of the students who participated in the study of Liu et al. (2024) were very confident or generally

confident in the model's answers, which were valid between 39% and 88% of the time (in different calls).

It is also necessary to train students in the specific use of these models. In this way, problems such as those arising from the use of personal data can be avoided. Some models use interaction data as training data, even though these data are personal. In addition to the problem of a company storing and training an AI system with personal data, it has been shown that it is possible to obtain training data from the model, including such personal data, by means of given prompts (Nasr et al., 2023). These models can be configured not to use such data as training data; however, shouldn't it be configured that way from the start, assuming a data protection approach by design and by default?

Finally, UNESCO (2019) insists that the use of AI technologies in education should be aimed at enhancing human capabilities, not replacing them. On many occasions, what is important is not so much the result to be achieved as the learning involved in the process to be followed. It is therefore important to avoid model dependency, so as not to compromise the development of intellectual skills such as written expression.

As for teachers, AI can enable them to perform their tasks more effectively and efficiently in administrative and teaching tasks (Chen, Chen & Lin, 2020). In their case, the focus is different from that of students, as their job is not to learn content and/or skills, but to transmit them. Even so, teachers should be aware of all the limitations that language models include (biases, hallucinations, non-determinism, etc.), and thereafter use them ethically and professionally. It is important to always review the answers they provide and not to delegate the evaluation to these types of systems by adopting a "human in the loop" approach.

Furthermore, it is important to bear in mind the relevance of the teacher-student relationship. It is fundamental for the well-being of the student, as well as an important factor for ensuring a better academic performance. This relationship is generated based on a complex intersection of beliefs, attitudes, behaviors and interactions between both (Hamre & Pianta, 2006). The tools analyzed in this chapter should not hinder this relationship through, for example, the loss of credibility or trust by working with incorrect or incoherent content generated by the language model.

2.5. Conclusion

Cazzaniga et al. (2024), analyzing the possible effects of generative AI on the labor market, conclude that 60% of jobs in advanced economies are exposed to the effects of the appearance of AI. Of these, one half may benefit from its use, while the other half will be negatively affected. Training in the use of AI is therefore essential, especially at the university level.

On the contrary, the use of AI as a training tool should be taken with caution, analyzing its potential usefulness, and fleeing from the excitement caused by fads that, as seen at the beginning of the chapter, end up entailing frustration for not delivering what others promise for them.

Honest and responsible research is essential in the application of AI to the field of education, analyzing use cases and testing whether improvements in learning occur. As for teachers, they too must (we must) make a responsible and conscious use of AI tools, always prioritizing students' learning.

All those involved must be realistic and aware of the capabilities and limitations of these models, which were designed to replicate human conversation and not to tell the truth, so that then, in the words of the philosopher Emmanuel Mounier, we "do not demand virtues from them that they do not have and do not reproach them for not giving what they do not have to give" (Mounier, 1990).

References

Baek, Y., Jung, J., & Kim, B. (2008). What makes teachers use technology in the classroom? Exploring the factors affecting facilitation of technology with a Korean sample. *Computers & Education*, 50 (1), 224-234. https://doi.org/10.1016/j.compedu.2006.05.002

Barrio Maestre, J. M. (2008). Educación y verdad. *Teoría de la Educación, 20*, 83-102.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Proceedings of the 34th International*

Conference on Neural Information Processing Systems, USA, Article 159, 1877-1901.

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E., & Mendes Tavares, M. (2024). *Gen-AI: Artificial Intelligence and the Future of Work*. IMF Staff Discussion, note SDN2024/001, International Monetary Fund.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in education: A review. *IEEE Access*, 8, 75264-75278. http://doi.org/10.1109/ACCESS.2020.2988510

Ciborra, C. (2002). *The Labyrinths of Information: Challenging the Wisdom of Systems*. OUP Oxford.

Crace, J. (2023, September 11). Switching off: Sweden says back-to-basics schooling works on paper. *The Guardian*. https://www.theguardian.com/world/2023/sep/11/sweden-says-back-to-basics-schooling-works-on-paper

Cuban, L. (1986). *Teachers and Machines: The Classroom Use of Technology Since 1920*. Teachers College.

Cuban, L. (2018). *The Flight of a Butterfly or the Path of a Bullet? Using technology to transform teaching and learning*. Harvard Education.

Desmuget, M. (2015). *La fábrica de cretinos digitales. Los peligros de las pantallas para nuestros hijos*. Península.

Fujimoto, S., & Takemoto, K. (2023) Revisiting the political biases of ChatGPT. *Frontiers in Artificial Intelligence, 6*. https://doi.org/10.3389/frai.2023.1232003

Gartner (s. f.). *Hype Cycle*. https://www.gartner.es/es/metodologias/hype-cycle

Hamre, B. K., & Pianta, R. C. (2006). Student-teacher relationships. In G. G. Bear, & K. M. Minke (Eds.). *Children's Needs III: Development, Prevention, and Intervention* (pp. 59-71). National Association of School Psychologists.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2023). *A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions*. arXiv. https://doi.org/10.48550/arXiv.2311.05232

Ilieva, G., Yankova, T., Klisarova-Belcheva, S., Dimitrov, A., Bratkov, M., & Angelov, D. (2023). Effects of generative chatbots in higher education. *Information, 14* (9), Article 9. https://doi.org/10.3390/info14090492

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31*(3), 268-289.

Kadaruddin, K. (2023). Empowering education through generative AI: Innovative instructional strategies for tomorrow's learners. *International Journal of Business, Law, and Education, 4*(2), 618-625. https://doi.org/10.56442/ijble.v4i2.215

Kocoń, J., Cichecki, I., Kaszyca, O., Kochanek, M., Szydło, D., Baran, J., Bielaniewicz, J., Gruza, M., Janz, A., Kanclerz, K., Kocoń, A., Koptyra, B., Mieleszczenko-Kowszewicz, W., Miłkowski, P., Oleksy, M., Piasecki, M., Radliński, L., Wojtasik, K., Woźniak, S., & Kazienko, P. (2023). ChatGPT: Jack of all trades, master of none. *Information Fusion, 99*, 101861. https://doi.org/10.1016/j.inffus.2023.101861

Liu, R., Zenke, C., Liu, C., Holmes, A., Thornton, P., & Malan, D. J. (2024). Teaching CS50 with AI. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2024)*, USA. https://doi.org/10.1145/3626252.3630938

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Miao, F., & Holmes, W. (2023). *Guidance for Generative AI in Education and Research*. UNESCO.

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *Inteligencia artificial y educación: guía para las personas a cargo de formular políticas*. UNESCO.

Mounier, E. (1990). *Obras completas* (vol. 3). Sígueme.

Nasr, M., Carlini, N., Hayase, J., Jagielski, M., Cooper, A.F., Ippolito, D., Choquette-Choo, C.A., Wallace, E., Tramèr, F., & Lee, K. (2023). *Scalable Extraction of Training Data from (Production) Language Models*. arXiv. https://doi.org/10.48550/arXiv.2311.17035

Newton, P., & Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2023.2299059

OET, Office of Educational Technology (2023). *Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations*. U.S. Department of Education.

OpenAI (2023). *Gpt-4 technical report*. arXiv. https://doi.org/10.48550/arXiv.2303.08774

Ortega y Gasset, J. (1982). *Meditación de la técnica y otros ensayos sobre ciencia y filosofía*. Revista Occidente, Alianza.

Parra Pennefather, P. (2023). Generative AI with personalities. In *Creative Prototyping with generative AI* (pp. 65-89). Apress. https://doi.org/10.1007/978-1-4842-9579-3_3

Qureshi, B. (2023). ChatGPT in computer science curriculum assessment: An analysis of its successes and shortcomings. *Proceedings of the 2023 9th International Conference on e-Society, e-Learning and e-Technologies* (pp. 7-13). https://doi.org/10.1145/3613944.3613946

Reich, J. (2020). *Failure to Disrupt: Why Technology Alone Can't Transform Education*. Harvard University.

Rescher, N. (1999). *Razón y valores en la era científico-tecnológica*. Paidós.

Shalom, J. G., Israeli, H., Marko Vitzky, O., LiPsitz, J. D. (2015). Social anxiety and physiological arousal during computer mediated vs. face to face communication. *Computers in Human Behavior*, *44*, 202-208. https://doi.org/10.1016/j.chb.2014.11.056

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education? *Smart Learning Environments*, *10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Tonmoy, S. M. T. I., Zaman, S. M. M., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). *A Comprehensive Survey of Hallucination Mitigation Techniques in Large Language Models*. arXiv. https://doi.org/10.48550/arXiv.2401.01313

UNESCO (2019). *Beijing Consensus on Artificial Intelligence and Education*. UNESCO Publishing https://unesdoc.unesco.org/ark:/48223/pf0000368303

Van Deemter, K. (2024). The pitfalls of defining hallucination. *Computational Linguistics*, 1-10. https://doi.org/10.1162/coli_a_00509

Wang, F., & Reeves, T. C. (2003). Why do teachers need to use technology in their classrooms? Issues, problems, and solutions. *Computers in the Schools*, *20*(4), 49-65. https://doi.org/10.1300/J025v20n04_05

Xu, Z., Jain, S., & Kankanhalli, M. (2024). *Hallucination is Inevitable: An Innate Limitation of Large Language Models*. arXiv. https://doi.org/10.48550/arXiv.2401.11817

Zhai, X., Chu, X., Chai, C. S., Jong, M. S., Istenič, A., Spector, M., Liu, J., Yuan, J., & Li, Y. (2021) A review of Artificial Intelligence (AI) in education from 2010 to 2020. *Complexity*, *2021*, Article 8812542. https://doi.org/10.1155/2021/8812542