Vol.6 No.2 (2023)

Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : http://journals.sfu.ca/jalt/index.php/jalt/index

A standardised PRISMA-based protocol for systematic reviews of the scientific literature on Artificial Intelligence and education (AI&ED)

Christian M. Stracke[A] A University of Bonn, Germany

Irene-Angelica Chounta[B] B University of Duisburg-Essen, Germany

Wayne Holmes[C] C University College of London, UK

Ahmed Tlili[D] D Beijing Normal University, China

Aras Bozkurt[E] E Anadolu University, Turkey

Keywords

AIED;
AI literacy;
AI&ED;
Artificial Intelligence;
education;
ethical;
learning;
PRISMA;
systematic review;
trustworthy.

Correspondence

stracke@uni-bonn.de [A]

Article Info

Received 31 October 2023
Received in revised form 25 November 2023
Accepted 26 November 2023
Available online 27 November 2023

DOI: https://doi.org/10.37074/jalt.2023.6.2.38

Abstract

By using standardised approaches, systematic reviews of the educational, scientific literature can inform educational research and influence educational policies and practices. However, the various systematic reviews of the scientific literature in the field of Artificial Intelligence (AI) and education all adopt individual approaches, making it challenging to systematically compare their conclusions. Accordingly, this paper presents a standardised protocol for conducting systematic reviews of the scientific literature on AI and education (AI&ED), including both literature on teaching and learning with AI (AIED) and literature on teaching and learning about AI (AI literacy). Our protocol applies the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and is presented here for the purpose of replication and validation. We exemplify our protocol by means of a systematic review of the scientific literature on *trustworthy and ethical AI&ED*, which was undertaken iteratively in symbiosis with the development of the protocol, informing each other throughout. In the future, we intend to apply our novel protocol for other search terms of relevance to AI&ED, as well as for the same search terms over a longer time period, in order to allow comparisons and the exploration of trends.

Introduction

Artificial Intelligence (AI) has been controversial since it was first introduced (Aiken & Epstein, 2000; Chaka, 2023; Huang et al., 2023; McCarthy et al., 1955). Nonetheless, it has now infiltrated almost all academic disciplines and most aspects of life outside academia (Borenstein & Howard, 2021). In particular, teaching and learning with AI (AIED) has been researched for around 50 years (Dillenbourg, 2016; Holmes et al., 2019; Holmes & Tuomi, 2022; Ifelebuegu et al., 2023; Mills et al., 2023; Pinkwart, 2016), both in K-12 education (Hrastinski et al., 2019) and in Higher Education (Crompton et al., 2020; Rasul et al., 2023). However, the implementation of AIED in classrooms, although growing rapidly, is still in its early stages. The same is true of teaching and learning about AI (AI literacy) (Holmes et al., 2022a).

Systematic reviews of the educational scientific literature using standardised approaches can inform educational research and can influence policies and practices in education. In fact, several systematic reviews of AIED have been published, including reviews about AIED in K-12 (Crompton et al., 2022; Sanusi et al., 2022), AIED in Higher Education (Zawacki-Richter et al., 2019), and AIED for specific educational purposes (Kurdi et al., 2020; Sottilare et al., 2018).

Meanwhile, global sustainability is emerging as an ambitious objective of AI developments. This is particularly true in the context of education (Chen et al., 2020; Chounta et al., 2022; European Parliament, 2021; Miao et al., 2021). The argument is that, to achieve a sustainable society, one thing that is necessary is to improve education about technology's (especially AI's) impact on humans, society and the environment (AI literacy) (Holmes et al., 2022a; Holmes & Tuomi, 2022; Holmes, 2023). In other words, to help ensure a sustainable society, we need students and citizens to have competences in both, the human and technological dimensions of AI, alongside other digital competences (Holmes et al., 2022b; Stracke et al., 2022a, 2022b, Vuorikari & Holmes, 2022). However, to date there has been limited research on the teaching and learning about these human and technological dimensions of AI literacy.

While, as noted, several systematic reviews have analysed the state-of-the-art of AIED (e.g., Chen et al., 2020; Crompton et al., 2022; Sanusi et al., 2022; Zawacki-Richter et al., 2019), they have mostly adopted individual approaches, making it challenging to systematically compare their outcomes and conclusions. Meanwhile, Tlili et al. (2023) have shown that the transparency level of literature reviews of AIED has been low. This low transparency level and the lack of comparability, due to there being no agreed or common approach, together highlight the need for a standardised protocol that might be used by researchers to enhance the transparency, comparability and quality of AIED (and, by extension, AI literacy) literature reviews. Such a protocol might better advance AI&ED (following Holmes et al., 2022a, we use AI&ED as shorthand for the combination of AIED and AI Literacy). However, prior to the reported study, no such common or standardised protocol or approach appeared to exist.

We contend that a key issue meriting systematic review within AI&ED is the intertwined conceptual pair of trustworthiness and ethics. Indeed, trustworthiness and ethics in AIED have been discussed in various recent publications (e.g., Bozkurt et al., 2023; European Commission, 2022; HLEG on AI, 2019; Holmes, 2023; Holmes et al., 2022a; Kazim & Koshiyama, 2021; Miao & Holmes, 2023; UNESCO, 2021). However, to the best of our knowledge, no study has systematically reviewed the relationship between education and trustworthy and/or ethical AIED, or indeed between education and trustworthy and/or ethical AI literacy.

To address these research gaps, the lack of an agreed protocol for systematic reviews of AI&ED and the lack of research into trustworthy and ethical AI&ED, this present study developed a standardised Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol for systematic reviews of AI&ED and, in symbiosis (both to illustrate and to inform the protocol), we undertook an example systematic review of trustworthy and ethical AI&ED. The protocol was designed to both, inform future research and to be used as a framework to help differentiate and classify theoretical concepts and practical approaches.

Our proposed standardised protocol for systematic reviews into AI&ED builds upon the PRISMA model in two ways. First, we identify and recommend particular starting search terms for the PRISMA phase 1, to which further terms can be added to narrow the search to the particular AI&ED issue of interest. Second, we identify and recommend particular inclusion and exclusion criteria for the PRISMA phases 2 and 3. In this way, our novel standardised protocol might help the comparability of future AI&ED systematic reviews, enhance the quality of AI&ED research findings, and increase the replicability of the methods adopted by AIED researchers and applications (in this sense, our protocol complements Ismail et al., 2023, which proposes a future systematic literature review on AIED in higher education also based on the PRISMA guidelines).

In summary, the present paper reports a standardised PRISMA-based protocol that researchers might adopt to conduct systematic reviews of scientific literature of AI&ED, to enhance the robustness of results and to enable systematic comparisons of results, that was developed in symbiosis with (i.e., each informing the other) an example project on trustworthy and ethical AI&ED. In the following sections, the full procedure of our example systematic review is presented in italic text. The outcomes of that example systematic review, which is not the core focus of this paper, will be presented in a separate paper; in this paper, we focus on the protocol. The pre-stage of the protocol, the selection of appropriate search terms, is described in the following section.

Pre-selection of search terms

The pre-selection of the search terms used for a systematic review is critical to ensure its feasibility. We propose that a fundamental requirement is that the records identified by the selected search terms can be handled by the researchers (*in our case, the co-authors of this paper*). Therefore, we

adopt the principle that the number of identified records should be higher than 50, to allow a meaningful analysis but lower than 1,000 to avoid an impractical workload. Naturally, a higher number of reviewers would be able to handle a higher number of papers.

For our standardised protocol for future systematic reviews, we propose the use of the electronic database Web of Science (WoS, www.webofscience.com), because it offers rigorous indexing services available for scientific and peer-reviewed publications. In WoS, putting "TS" ("Topic") in the Advanced Search Query Builder search string causes the search to be undertaken in the following fields within each record: Title, Abstract, and Author Keywords. Comparing the results of various search strings for our example systematic review, for our standardised protocol, we decided upon the search term "TS = ((Artificial Intelligence) AND (education*) AND ([ISSUE OF INTEREST]))". We also tried including alternative terms for "artificial intelligence", but it became clear that presumably because other terms and synonyms are only used in combination with "artificial intelligence", this did not reveal a noticeably greater number of records.

In our example systematic review, we submitted several "TS" search strings to the Advanced Search Query Builder on 21st of November 2022 (see Table 1). In order to identify papers that considered trustworthy and ethical AI&ED, we replaced "([ISSUE OF INTEREST])" with "((trust) OR (ethic*))".*

Table 1: Search strings used in the example systematic review's pre-selection, and the numbers of identified records.

#	Search string	Identified records
1	TS = (Artificial Intelligence)	79,429
2	TS = (education*)	1,203,303
3	TS = ((Artificial Intelligence) AND (education*))	3,694
4	*TS = (trust*)*	*164,738*
5	*TS = (ethic*)*	*265,794*
6	*TS = ((trust*) OR (ethic*))*	*423,626*
7	*TS = ((Artificial Intelligence) AND ((trust*) OR (ethic*)))*	*4,068*
8	*TS = ((Artificial Intelligence) AND (education*) AND ((trust*) OR (ethic*)))*	*324*
9	*TS = ((Artificial Intelligence) AND (education*) AND ((trust*) OR (ethic*)) AND (higher education*))*	*92*

For our example systematic review, we selected search term #8 (i.e., "TS = ((Artificial Intelligence) AND (education) AND ((trust*) OR (ethic*)))"), because it included our issue of interest ("trustworthy and ethical") and identified a number of records that could be properly analysed by a small team (as is typical of most research labs).*

The four selection phases of the proposed systematic review protocol, *illustrated by our example systematic review*, are detailed in the following section.

Standardised protocol methodology

Our standardised protocol defines the methodology for the systematic review of AI&ED. It strictly follows the PRISMA statement and its procedures while adapting and specifying the original four PRISMA phases for the specific selection of AI&ED articles (Moher et al., 2009; Page et al., 2021). The protocol was developed in symbiosis with our example *systematic review of trustworthy and ethical AI&ED* (both informing each other). The four PRISMA-based phases for the selection of articles are:

1. Identification,
2. Screening,
3. Eligibility, and
4. Included.

To ensure the reliability of the process, the four phases of the PRISMA process should be undertaken by at least two reviewers, each of whom having research experience in AI&ED. The rationale is that two knowledgeable researchers working independently and then sharing and discussing their results until they reach a consensus will reduce personal biases sufficiently.

In our example systematic review, the pre-selection of the search strings and the identification of the articles using the search terms were undertaken by two reviewers, supported by three further reviewers as required, all of whom have extensive research experience in AI&ED.

Phase 1

In the first PRISMA-based phase, *identification*, the selected database(s) is searched using the pre-defined search string. This phase is concluded by the elimination of duplicates.

The full procedure for the first phase of our example systematic review was as follows. The Advanced Search Query Builder of WoS was used with the search string ("TS = (("Artificial Intelligence") AND (education) AND ((trust) OR (ethic*)))") (see Section "Pre-selection of search terms"). As noted, this phase identified 324 records. The elimination of duplicates was not necessary here as only one data source was used.*

Phase 2

In the second PRISMA-based phase, *screening*, the title and abstract of the records identified in the first phase are reviewed, to identify and remove all records that do not meet the search aims. Table 2 gives an overview of the inclusion and exclusion criteria that we defined for the second phase of our standardised protocol.

As noted above, at least two reviewers should review in parallel the titles and abstracts of a randomly selected subset of all records. This subset should contain a minimum 5% of all records identified during the first phase because the reviewers should compare a substantial number of records after their independent reviews. Adopting a lower margin of 5% is a common sense decision designed to ensure that the

Table 2: Inclusion and exclusion criteria for screening the identified records.

Inclusion criteria	Exclusion criteria
The full text can be downloaded.	The full text of the record is not available.
The record is written in English.	The record is not written in English.
The record is published in a peer-reviewed journal.	The record is not published in a peer-reviewed journal (including conference proceedings).
The record is a scientific article.	The record is not a scientific article (instead it is an editorial, interview, research protocol, etc.).
The record is substantial (>2 pages).	The record is not substantial (<3 pages).

selection is representative while minimising the number of false positives. In their independent reviews, the reviewers should apply the inclusion and exclusion criteria outlined in Table 2. In cases of uncertainty, the related record should be kept. Afterwards, reviewers should compare their independent results for the random subset and discuss them in detail aiming to reach consensus on all decisions.

We propose that there are two possible outcomes of the independent reviews of the random subset during the second phase depending on the results of their independent reviews. The threshold of the criterion is set to zero because the reviewers should achieve common understanding and complete agreement about the inclusion or exclusion of records.

> Outcome 1: If the independent review results show one contradictory case or more, another random subset of records should be identified and independently reviewed in parallel. The subsequent independent review results should then again be compared and discussed as explained earlier.
>
> Outcome 2: If the independent review results are exactly the same, the remaining records should be shared among the reviewers to complete the second review phase. During that review, the reviewers should note all decisions about which they are not certain, for later discussion until a consensus is reached.

The full procedure for the second phase of our example systematic review was as follows. Two reviewers reviewed in parallel the titles and abstracts of a random subset of 5% of all the records identified in the first phase, using the inclusion and exclusion criteria outlined in Table 2. The researchers first worked independently, and then discussed their results until they reached a consensus (any records about which a researcher was uncertain or about which the researchers disagreed were discussed in depth in order to reach the consensus). In this way, personal biases were reduced. In our case, we took a first random subset of 24 records, leaving exactly 300 records that could be easily shared among the researchers to complete the second review phase.

Phase 3

In the third PRISMA-based phase, eligibility, the full texts of the remaining records are reviewed to finalise the selection of texts, ensuring that they all fulfil content-related requirements. Table 3 gives an overview of the inclusion and exclusion criteria that were defined for the third phase of our standardised protocol.

At least two reviewers should review in parallel the full text of a random subset of 50 records generated by the second phase, using the inclusion and exclusion criteria outlined in Table 3. The amount of 50 records is selected because the threshold can be set to two contradictory cases while still keeping the 95% margin of the normal distribution (2 cases out of 50 records are 4%).

Table 3: Inclusion and exclusion criteria for content-related screening of collected records.

Inclusion criteria	Exclusion criteria
The record addresses trustworthy or ethical AI.	The record does not address trustworthy or ethical AI.
The record addresses education.	The record does not address education.

As before, the researchers first should work independently, and then discuss their results until they reach a consensus (any records about which a researcher is uncertain or about which the researchers disagree should be discussed in depth in order to reach the consensus). Afterwards, reviewers should compare their independent results and discuss them in detail to reach consensus on all decisions. In cases of uncertainty, the records should be kept.

We propose that there are two possible outcomes of the third phase independent reviews of the random subset of 50 records depending on the results of their independent reviews. The threshold of the criterion is set to two because this limit keeps the 95% margin of the normal distribution (2 cases out of 50 records are 4%) and the reviewers need to discuss any contradictory case in details to achieve consensus in the end.

> Outcome 1: If the independent review results show more than two contradictory cases, then another random subset of records should be generated and independently reviewed in parallel. The independent review results should be compared and discussed in the same way as explained before.
>
> Outcome 2: If the independent review results are the same or differ in only one or two cases, the remaining records will then be shared among the reviewers to complete the third review phase. During that review, the reviewers should note all decisions about which they are not certain, for later discussion until a consensus is reached.

The full procedure for the third phase of our example systematic review closely followed the steps just outlined.

Phase four

In the fourth and final PRISMA-based phase, *included*, the remaining records are used for the actual systematic review, involving an in-depth analysis and discussion with respect of the research question(s). To begin with, the reviewers should propose an initial categorisation for the articles selected in the previous phases because such a categorisation is required for a systematic assignment and analysis of the articles. This categorisation should be discussed until a consensus about the terms and their categorical structure is reached. The discussion should include various dimensions of the topic in question, comments on general trends, and limitations of the systematic review and its analysis because all reviewers should be explicit about their analysis perspectives. The included papers should then be categorised according to this nominal taxonomy, using an iterative process. The systematic review will conclude with an outlook on future research needs and potential research questions.

The full procedure for the fourth phase of our example systematic review will closely follow the steps just outlined. As this example systematic review is not the core focus of this paper, its outcomes will be presented in a separate paper.

Results

The results we report here are for the standardised protocol (*the results of the example systematic review will be presented in a separate paper*). Figure 1 presents a template (*illustrated with numbers from our example systematic review*) that may be used to report the results of the four PRISMA-based phases determining the final selection of articles for in-depth analysis. It should be used, adapted with the appropriate numbers, for the four protocol phases of any future systematic review and its results.

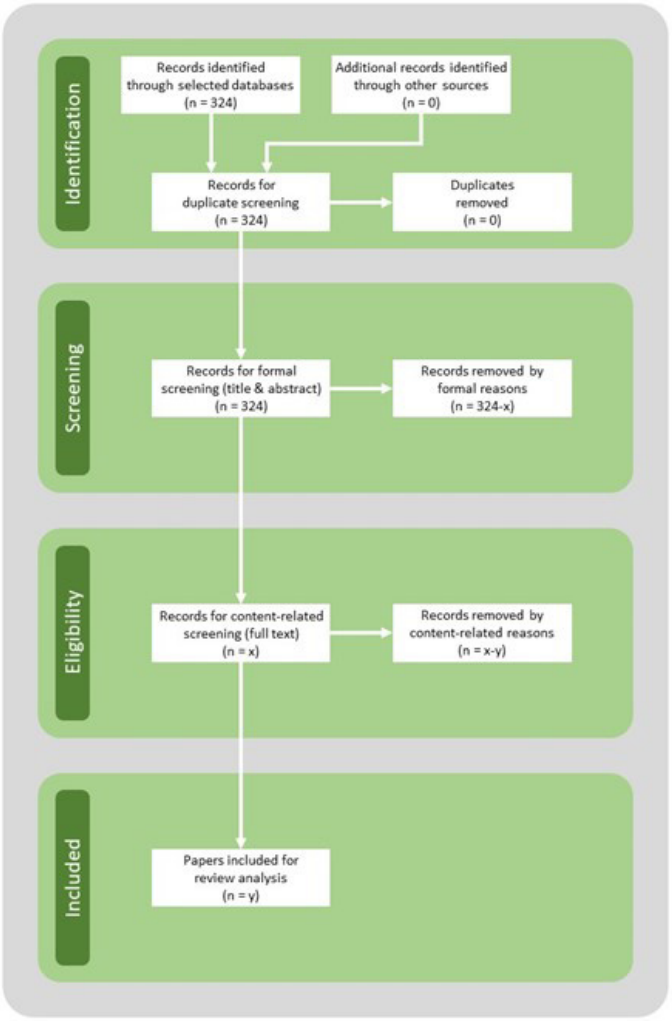

Figure 1: Template for the summary of the four phases of the standardised protocol.

Conclusions

This paper presented a standardised protocol which could serve as the basis of systematic literature reviews of AI&ED. It was developed in symbiosis with an example systematic review of trustworthy and ethical AI&ED research (the results will be published after finalising the analysis), informing each other throughout. The standardised protocol and the example systematic review were mutually informed by means of sense-testing and evaluation throughout.

To the best of our knowledge, no study in the literature has provided such a protocol on this increasingly important topic area before. The protocol identifies a suitable database (WoS), offers pre-defined search terms with the opportunity to fine-tune them to the issue of interest, and provides a structure that can be used in the systematic review of any aspect of AI&ED. By means of this standardised protocol, personal bias can be reduced, and the quality of the reported findings can be enhanced. It will be easier to systematically compare the results with those of other studies using the same protocol. In turn, this should advance AI&ED research, development and application.

In particular, our standardised protocol offers a template for undertaking future systematic reviews of AI&ED. The precise steps outlined above build a standardised protocol that anyone can easily repeat, and its repeated usage will lead to its validation and its continuous improvement. We envision that this protocol can contribute towards the standardisation of systematic reviews in the field of AI&ED, support the comparison of findings, enable the mapping of research trends over time, inform policymakers and educators, and influence policies and practices in AI&ED. Our current research focuses on applying our standardised protocol to a larger and up-to-date dataset of trustworthy and ethical AI&ED.

Acknowledgements

This paper was supported by Anadolu University as part of a project supervised by Aras Bozkurt (Grant number: SBA-2023-1852).

References

Aiken, R. M., & Epstein, R. G. (2000). Ethical guidelines for AI in education: Starting a conversation. *International Journal of Artificial Intelligence in Education, 11*, 163–176.

Borenstein, J., & Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. *AI and Ethics, 1*(1), 61–65. https://doi.org/10.1007/s43681-020-00002-7

Bozkurt, A., Karadeniz, K., Baneres, D., Guerrero-Roldán, A. E., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability, 13*(2), 1-16. https://doi.org/10.3390/su13020800

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., ..., & Jandrić, P. (2023). Speculative futures on ChatGPT and generative Artificial Intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education, 18*(1), 53-130. https://doi.org/10.5281/zenodo.7636568

Chaka, C. (2023). Fourth industrial revolution—a review of applications, prospects, and challenges for Artificial Intelligence, robotics and blockchain in higher education. *Research and Practice in Technology Enhanced Learning, 18*(2), 002. https://doi.org/10.58459/rptel.2023.18002

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

Chounta, I.-A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of Artificial Intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education, 32*(3), 725-755. https://www.doi.org/10.1007/s40593-021-00243-5

Crompton, H., Bernacki, M. L., & Greene, J. (2020). Psychological foundations of emerging technologies for teaching and learning in higher education. *Current Opinion in Psychology, 36*, 101–105. https://doi.org/10.1016/j.copsyc.2020.04.011

Crompton, H., Jones, M. V., & Burke, D. (2022). Affordances and challenges of Artificial Intelligence in K-12 education: A systematic review. *Journal of Research on Technology in Education*. https://doi.org/10.1080/15391523.2022.2121344

Dillenbourg, P. (2016). The evolution of research on digital education. *International Journal of Artificial Intelligence in Education, 26*, 544-60. https://doi.org/10.1007/s40593-016-0106-z

European Commission. (2022). *Ethical guidelines on the use of Artificial Intelligence (AI) and data in teaching and learning for educators*. https://data.europa.eu/doi/10.2766/153756

European Parliament (2021). *Report on Artificial Intelligence in education, culture and the audiovisual sector (2020/2017(INI))*. Committee on Culture and Education. https://www.europarl.europa.eu/doceo/document/A-9-2021-0127_EN.html

High-Level Expert Group (HLEG) on AI. (2019). *Ethics guidelines for trustworthy AI*. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Holmes, W. (2023). *The unintended consequences of artificial intelligence and education*. Education International Research. https://www.ei-ie.org/en/item/28115:the-unintended-consequences-of-artificial-intelligence-and-education

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in education. Promises and implications for teaching and learning*. Center for Curriculum Redesign. Boston, MA.

Holmes, W., Persson, J., Chounta, I.-A., Wasson, B., & Dimitrova., V. (2022a). *Artificial Intelligence and Education. A critical view through the lens of human rights, democracy and the rule of law*. https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680a886bd

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., ..., & Koedinger, K. R. (2022b). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education 32*(3), 504-526. https://www.doi.org/10.1007/s40593-021-00239-1

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Hrastinski, S., Olofsson, A. D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., ..., & Utterberg, M. (2019). Critical imaginaries and reflections on Artificial Intelligence and robots in post digital K-12 education. *Postdigital Science and Education, 1*(2), 427–445. https://doi.org/10.1007/s42438-019-00046-x

Huang, R., Tlili, A., Xu, L., Chen, Y., Zheng, L., Saleh Metwally, A. H., ..., & Bonk, C. J. (2023). Educational futures of intelligent synergies between humans, digital twins, avatars, and robots - the iSTAR framework. *Journal of Applied Learning & Teaching, 6*(2), 1-16. https://doi.org/10.37074/jalt.2023.6.2.33

Ifelebuegu, A. O., Kulume, P., & Cherekut, P. (2023). Chatbots and AI in Education (AIEd) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching, 6*(2), 1-14. https://doi.org/10.37074/jalt.2023.6.2.29

Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning & Teaching 6*(2), 1-8. https://doi.org/10.37074/jalt.2023.6.2.34, last accessed 2023/10/26.

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns, 2*(9), 100314. https://doi.org/10.1016/j.patter.2021.100314

Kurdi, G., Leo, J., Parsia, B., Sattler, U., & Al-Emari, S. (2020). A systematic review of automatic question generation for

educational purposes. *International Journal of Artificial Intelligence in Education, 30*, 121-204. https://doi.org/10.1007/s40593-019-00186-y

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). *A proposal for Dartmouth summer research project on Artificial Intelligence*. https://www-formal.stanford.edu/jmc/history/dartmouth.pdf

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000376709

Mills, A., Bali, M., & Eaton, L. (2023). How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *Journal of Applied Learning and Teaching, 6*(1), 16-30. https://doi.org/10.37074/jalt.2023.6.1.34

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., …, & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews, 10*, 89. https://doi.org/10.1186/s13643-021-01626-4

Pinkwart, N. (2016). Another 25 years of AIED? Challenges and opportunities for intelligent educational technologies of the future. *International Journal of Artificial Intelligence in Education, 26*, 771-83. https://doi.org/10.1007/s40593-016-0099-7

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., ..., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching, 6*(1), 41-56. https://doi.org/10.37074/jalt.2023.6.1.29

Sanusi, I. T., Oyelere, S. S., Vartiainen, H., Suhonen, J., & Tukiainen, M. (2022). A systematic review of teaching and learning machine learning in K-12 education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-11416-7

Sottilare, R. A., Burke, S., Salas, E., Sinatra, A. M., Johnston, J. H., & Gilbert, S. B. (2018). Designing adaptive instruction for teams: A meta-analysis. *International Journal of Artificial Intelligence in Education, 28*, 225-64. https://doi.org/10.1007/s40593-017-0146-z

Stracke, C. M., Burgos, D., Santos-Hermosa, G., Bozkurt, A., Sharma, R. C., Swiatek, C., ..., & Truong, V. (2022a). Responding to the initial challenge of COVID-19 pandemic: Analysis of international responses and impact in school and higher education. *Sustainability, 14*(3), 1876. https://doi.org/10.3390/su14031876

Stracke, C. M., Sharma, R. C., Bozkurt, A., Burgos, D., Swiatek, C., Inamorato dos Santos, A., ..., & Truong, V. (2022b). Impact of COVID-19 on formal education: An international review on practices and potentials of Open Education at a distance. *The International Review of Research in Open and Distributed Learning, 23*(4), 1-18. https://doi.org/10.19173/irrodl.v23i4.6120

Tlili, A., Huang, R., Mustafa, M. Y., Zhao, J., Bozkurt, A., Xu, L., ..., & Burgos, D. (2023). Speaking of transparency: Are all Artificial Intelligence (AI) literature reviews in education transparent? *Journal of Applied Learning and Teaching, 6*(2), 1-12. https://doi.org/10.37074/jalt.2023.6.2.15

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2021). *Recommendation on the Ethics of Artificial Intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Vuorikari, R., & Holmes, W. (2022). DigComp 2.2. Annex 2. Citizens Interacting with AI Systems. In R. Vuorikari, S. Kluzer, & Y. Punie (Eds.), *DigComp 2.2, The digital competence framework for citizens: With new examples of knowledge, skills and attitudes* (pp. 72–82). Publications Office of the European Union. https://data.europa.eu/doi/10.2760/115376

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on Artificial Intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0