Exploring the Methodological Contexts and Constraints of Research in Artificial Intelligence in Education

Irene-Angelica Chounta[1] [0000-0001-9159-0664], Bibeg Limbu[1] [0000-0002-1269-6864], and Lisa van der Heyden[1] [0000-0003-0197-7473]

Department of Human-Centered Computing and Cognitive Science, University of Duisburg-Essen, Duisburg, Germany
{irene-angelica.chounta,bibeg.limbu,lisa.van-der-heyden}@uni-due.de

Abstract. In this paper, we present a Systematic Literature Review (SLR) on the state-of-the-art in Artificial Intelligence in Education (AIED) focusing on methodological contexts and constraints of the research landscape. To do so, we built on existing works and extended them to cover the latest research advancements in the field over the past five years. We aimed at covering all educational levels and retrieving important data regarding the planning and execution of research studies and the robustness of results. In total, we reviewed 181 papers and answered three research questions, relating to the educational context of AI use, the methodology and study design utilized in AIED research, and the type of AI algorithms and technologies used in education. Our findings suggest that research in AIED primarily focuses on formal, higher education and that there is a demand for robust and rigorous scientific evidence of the effectiveness and impact of AIED. Furthermore, the findings indicate that the most popular AI technologies currently studied are traditional AI algorithms, usually used for prediction, classification, or clustering. Based on our analysis, we discuss practical implications that can serve as inspiration and guidance for future research initiatives.

Keywords: artificial intelligence in education · research · methodology · systematic literature review

1 Introduction

Research on Artificial Intelligence in Education (AIED) has been a prominent topic for over 30 years. However, recently it gained attention due to technological advances in Artificial Intelligence (AI) and is projected to reach its peak by the year 2025[1]. In the past five (5) years alone, we identified nine (9) systematic literature reviews (SLRs) about AIED that aimed to document the research landscape (see Table 1).

[1] https://www.holoniq.com/notes/2019-artificial-intelligence-global-education-report

SLR Title	Years Covered	#Reviewed Publications
Systematic review of research on artificial intelligence applications in higher education–where are the educators? [20]	2007 - 2018	146
Application and theory gaps during the rise of artificial intelligence in education [3]	1999 -2019	45
AI in education: A systematic literature review [16]	2010 - 2019	23
A systematic review of AI role in the educational system based on a proposed conceptual framework [18]	2005 - 2021	51
Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020 [11]	2011 - 2020	32
Affordances and challenges of artificial intelligence in K-12 education: A systematic review [6]	2011 - 2021	169
Artificial intelligence applications in K-12 education: A systematic literature review [19]	2011 - 2021	210
Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education [4]	2012 - 2021	92
Artificial intelligence in higher education: the state of the field [5]	2016 - 2022	138

Table 1. Existing Systematic Literature Reviews on AIED

Most of these reviews explored the integration and use of AI in specific levels of formal education, such as K-12 [19, 6] or higher education [5, 20], while one review explored the whole range [4]. A common theme addressed in the SLRs was the benefits, opportunities, challenges, and critical implications that follow AIED [4, 16, 6]. For example, [20] explored – among others – the ethical implications and risks associated with the use of AI in education, while [6] also highlighted several challenges, such as ethical concerns regarding AI use, ensuring fairness and equal opportunities, providing sufficient training and support for teachers, addressing technical infrastructure requirements, and ensuring AI's integration and long-term sustainability in education. [16] pointed out the scarcity of publications presenting the relationship between opportunities, benefits, and challenges. Some reviews focused on the geographical distribution of research, pointing out that certain countries lead the AIED-related publications [5, 3]. Finally, other topics discussed in these SLRs were the most prominent AI technologies and algorithms presented in research publications [3, 11], the purpose of use, roles, and functions of AI [11, 18, 6, 20].

The above signifies a plethora of research on AIED, especially after the COVID-19 pandemic, strengthened by recent advances in AI technologies, specifically generative AI. It is critical to monitor and document these works and keep an up-to-date state-of-the-art map regarding the AIED research landscape. However, there are evident gaps in the information and level of detail documented. For example, SLRs that examine the use of AI on all levels of education [3, 4] focus mostly on the geographical distribution of research, opportunities, and challenges. On the other hand, SLRs that focus on specific education levels (for example, higher education or primary education [19, 20]) delve into topics related to the use and purpose of AI in education. Another gap concerns the documentation of data that AI algorithms and AI systems use to achieve their goals. Existing SLRs provide minimal to no information about the data types or data sources, while few point out potential limitations or risks due to data privacy and ownership. We argue that there is a need to systematically document information regarding the AI technologies used for all education levels and their purpose. Most importantly, it is crucial to identify the data sources and types typically employed by AI applications in education. Additionally, there is no systematic evaluation of the research studies under review regarding their scientific rigorousness and robustness of the evidence. We acknowledge the difficulties that the organization and realization of such studies entail. Nonetheless, we argue that it is imperative to document the magnitude of studies in terms of population size, duration, reproducibility, and applicability, along with their methodological designs, to assess the robustness of their findings. This requirement is also imposed by the need to ensure fair, accountable, transparent, and ethical AI [9].

The contribution of this work is to gather insights regarding the methodological contexts and constraints of the AIED research landscape by building on existing works and extending them to cover the latest research advancements in the field over the past five years. In particular, we aimed to retrieve important information regarding the planning and execution of research studies that can potentially provide information about the robustness of results and findings.

For our research purposes, we formulated the following Research Questions (RQs):

RQ$_1$. *In which educational contexts is AI used?*

RQ$_2$. *What are the methodological and study designs employed in AIED research?*

RQ$_3$. *What AI algorithms and technologies are used in education?*

To answer our RQs, we used the SLR of [20] as a blueprint, and we extended it further to: a) cover all research literature from 2019 to 2023 (5 years) appearing in journal publications; b) cover all education levels, including adult and lifelong learning; c) retrieve information regarding the planning and realization of research studies and the robustness of results.

In the following sections, we discuss the methodological setup of this work. Then, we present our findings (section 3), and we answer our research questions along with a contextualized discussion on the need for evidence-based approaches for AIED (section 4). We conclude with a brief summary of our contribution, limitations and future work (section 5).

2 Methodology

To carry out the SLR, we followed the PRISMA statement [12, 14] that require four phases for selecting articles: a) Identification, b) Screening, 3) Eligibility and 4) Inclusion. For the Identification, we searched the databases Web of Science (WoS)[2], EBSCO[3] and Scopus[4] using the pre-defined search term:

> Search Query = ("artificial intelligence" OR "machine intelligence" OR "intelligent support" OR "intelligent virtual reality" OR "chatbot*" OR "machine learning" OR "automated tutor" OR "personal tutor*" OR "intelligent agent*" OR "expert system" OR "neural network" OR "natural language processing") AND ("higher education" OR college* OR undergrad* OR graduate OR postgrad* OR "K-12" OR kindergarten* OR "corporate training*" OR "professional training*" OR "primary school*" OR "middle school*" OR "high school*" OR "elementary school*" OR "vocational education" OR "adult education") AND (learn* OR student*)

The search string derived from the SLR of [20], and it was used to retrieve relevant journal publications from the pre-selected databases. After retrieving the search results, we concluded the *Indentification* phase by eliminating duplicates. For *Screening*, the publications were reviewed based on their title and abstract by two researchers in parallel who eliminated publications not considered relevant. During the *Eligibility* phase, the remaining records were reviewed based on their full texts: Two reviewers worked in parallel using the exclusion and inclusion criteria (Table 2) adapted from [14]. To support the reviewing process, we compiled a coding scheme with examples for each code. To train the reviewers, we used three papers to demonstrate the coding process. During training, the reviewers could ask for clarifications and take notes. Then, we carried out two reviewing steps. For the first step, we asked the reviewers to code together a set of ten (10), randomly selected papers to establish a common understanding. During this step, we provided additional clarifications and examples for the coding scheme and refined it further when needed. For the second step, we asked the reviewers to code separately a set of 20, randomly selected, papers. Then, we asked the reviewers to compare and discuss their results until they reached a consensus. If consensus was not possible, a third reviewer, an expert on AIED and educational technologies, was involved. This process was repeated until all publications were reviewed.

During the *Inclusion* phase, all remaining publications were included in the corpus for the systematic review. Two of the co-authors split the corpus and reviewed the papers again independently to confirm the original findings. If in doubt, the third co-author was involved. The number of publications over the different phases of the process is shown in Figure 1. The list of articles that

[2] https://www.webofscience.com/wos/woscc/
[3] https://www.ebsco.com/
[4] https://www.scopus.com/search/

Formal Screening	
Exclusion Criteria	**Inclusion Criteria**
The full text of the record is not available.	The full text can be downloaded.
The record is not written in English.	The record is written in English.
The record is not published in a peer-reviewed journal	The record is published in a peer-reviewed journal.
The record is not substantial (< 7 pages).	The record is substantial (>= 7 pages).
Content Screening	
Exclusion Criteria	**Inclusion Criteria**
The record does not address the use of AI.	The record addresses the use of AI.
The record does not address education.	The record addresses education.
The record does not involve a research study	The record involves a research study

Table 2. Exclusion and inclusion criteria for formal screening of collected records adapted from [14]. Formal criteria relate to the *formalization* of the publications, such as the working language, the number of pages or the record's availability. Content-related criteria address the topic of the publication and were established in accordance with our RQs.

were included in the review and the coding scheme are publicly available at `http://tinyurl.com/3rf5jzw6`.

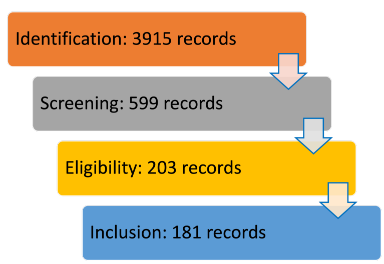

Fig. 1. The number of papers for each phase of the reviewing process according to PRISMA.

3 Results

In total, we reviewed one hundred and eighty-one (181) articles published in sixty-seven (67) journals from 2019 to 2023. The three (3) leading journals in terms of number of publications were the Education and Information Technologies journal (19 publications, 11%), the British Journal of Educational Technology (11 publications, 6%), and the International Journal of Educational Technology in Higher Education (9 publications, 5%). Figure 2 presents the distribution of publications over the last five (5) years. We note that for 2023, we only considered papers available by April 1st, 2023.

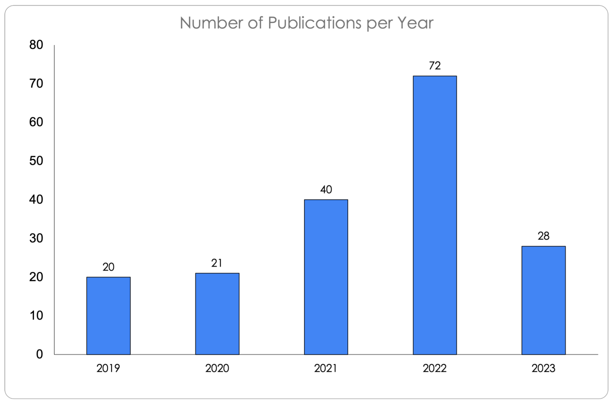

Fig. 2. The distribution of published papers over the past 5 years, from January 2019 to April 2023.

To align with existing SLRs, and to gain some understanding regarding potential hubs of AIED research, we will first present some "demographics" of the literature under review. Overall, the first authors of the research publications came from forty-eight (48) countries in terms of affiliation. The most prominent countries in terms of number of publications were the USA (37 publications, 20%), and China (18 publications, 10%). Next, we analyzed the scientific disciplines of the first authors of the publications. To classify various affiliations, we used the taxonomy proposed by the German Research Foundation (DFG)[5]. Our findings showed that most research comes from engineering sciences (88 publications, 49%) followed by humanities and social sciences (62 publications, 34%). Six (6) affiliations could not be retrieved.

3.1 Educational contexts and AI (RQ1)

From our results, we identified four cases regarding the educational contexts that appear in AIED research: a) Higher education (146 publications, 81%); b) K-12 Education (32 publications, 18%); c) Vocational Education (2 publications, 1%); d) Mixed Levels (K-12 and Higher Education) (1 publication, 2%).

[5] `https://www.dfg.de/download/pdf/dfg\_im\_profil/gremien/fachkollegien/amtsperiode\_2020\_2024/fachsystematik\_2020\-2024\_en\_grafik.pdf`

Higher education dominates AIED research. This finding was not surprising due to the affordances and potential that higher education offers for conducting research. One can argue that researchers are inclined to conduct studies within their own institutions since they are already familiar with the context. Other reasons could be stakeholders' familiarity with educational technologies and the level of integration of educational technologies within the curriculum. Also higher education students – as adults – are potentially the most accessible group in terms of ethical requirements that studies must adhere to.

3.2 Methodological and study designs in AIED (RQ2)

To answer this question, we collected information regarding the methodologies employed (experimental or observational) in state-of-the-art AIED research, the study design regarding population size, and repeatability (that is, whether the study took place only once or repeated times), the analytical methods employed (qualitative, quantitative or mixed-methods) and the data types and sources that researchers collect and/or employ in their approaches. We classified a study as experimental if researchers introduced an intervention, a test under controlled conditions, and studied the effects following the definition[6] provided by the American Psychological Association (APA). Studies that were not classified as experimental were tagged as observational. This scheme did not always align with publications' descriptions that used the term "experiment" in their methodological description but did not meet our criteria for an experiment.
Many studies involved the training of ML models for predictive purposes, such as predicting student dropouts. Although these studies do not qualify as experimental based on our criteria, we acknowledge their importance. Most publications presented studies that followed observational methods (135 publications, 75%) while the rest of the publications (46 publications, 25%) presented experimental studies with two conditions (control, experimental) or more. Among the observational studies, 121 employed quantitative methods (66.85%), 4 employed qualitative methods (2.21%), and 10 conducted mixed methods studies (5.52%). On the other hand, 31 experimental studies used quantitative methods (17.13%)), 1 study used purely qualitative methods (0.55%), and 14 used mixed methods (7.73%). The most common method for data collection was "surveys", used by 76 studies (41.99%) to collect student responses typically via tools such as questionnaires. This was closely followed by the students' "academic data" from platforms such as learning management and student information systems (75 studies, 41.44%). Other data types and sources employed involved assessment data (such as grades and knowledge-tests), student trace data from applications' logfiles, and other learning artifacts (such as essays, and student reports).

Regarding repeatability, most publications (177 publications, 98%) presented one study. Only four (4) publications (2%) reported repeated measures. Finally, in terms of study population size, most publications (53 publications, 29.28%) presented studies with less than 100 participants, followed by 44 publications

[6] `https://dictionary.apa.org/experiment`

(24.31%) ranging from 100 to 500 participants. Only 24 studies (13.26%) discussed studies with a population size of more than 10K (Figure 3). We want to note that, in our case, population size does not always refer to the number of participants. It can instead refer to data entries or data points (such as the matriculation records in a time period). However, we tried to match, based on the information that was available in the reviewed publications, the number of participants to the data records used in the studies.

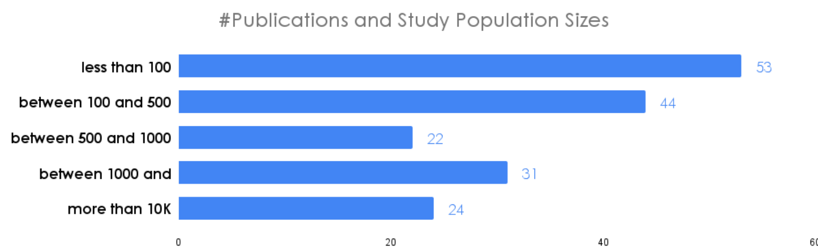

Fig. 3. The population sizes of research studies reported in the literature review.

3.3 AI algorithms and AI technologies used in education (RQ3)

We followed a combined bottom-up and top-down approach to retrieve and classify the AI algorithms and technologies used in education as referenced in research publications. First, two reviewers extracted this information from the publications under review, and then we categorized them based on existing literature. Overall, traditional machine learning algorithms were most frequently in the publications, either on their own or in combinations, such as Random Forest (42 references), Regression (36 references), Support Vector Machines (27 references), and Bayesian Networks (20 references). These algorithms were used mainly for prediction, clustering, and classification tasks. NLP methods were referenced 49 times, in relation to conversational agents and chatbots (17 references), and BERT models (11 references). Finally, publications referenced neural networks (12 references), Intelligent Tutoring Systems (5 times), and Recommender Systems (5 times).

Regarding the purpose of using AI, publications focused on predicting the academic performance of students (69 references), and predicting students at-risk of dropping out. Algorithms such as Random Forest, Support Vector Machines and Regression were commonly used for such purposes. Other usages of AI revolved around supporting students in their career choices (23 references), automated evaluation and assessment in academic contexts (12 references).

Our findings confirm prior research [3] suggesting that traditional machine learning algorithms are more popular and heavily used than advanced AI technologies such as deep learning and artificial neural networks. Additionally, we saw a growing interest in NLP and NLP-related technologies, such as chatbots and conversational agents, justified by recent advances in NLP.

4 Discussion

4.1 In which educational contexts is AI used? (RQ1)

AIED research focuses, primarily, on higher education (see section 3.1). This may suggest that AI applications are limited to specific settings and not generalized or adapted to serve multiple contexts. Also, the potential and impact of AI on other important sectors, such as workplace and life-long learning, is not sufficiently explored. We perceive this as a unique research opportunity to cross-validate existing AI approaches for learning, such as ITSs, over a wide range of learning contexts. To achieve this, it is necessary to involve stakeholders from diverse settings and explore different aspects of learning.

4.2 What are the methodological and study designs employed in AIED research? (RQ2)

There is a need for robust and rigorous scientific evidence for the effectiveness and impact of AIED (see section 3.2). Most research studies are observational in nature, with a limited population sample, limited applicability, and transfer in new contexts. For example, studies may focus on AI models' performance but fall short in experimentally testing models in authentic contexts.

It is evident that we need to focus on multilevel, mixed-methods approaches, as recommended by [15, 17], and involve triangulation to allow for cross-validation of findings.

Notably, a substantial number of studies have not made their data set open access for various reasons, such as privacy [10] etc. Some authors state willingness to make the data available on request [13, 8], however there is a lack of interoperability standards in the community for communication and data exchange. This affects the reproduction and cross-validation of findings necessary for open science. This lack of standardization was also observed in reporting, making it difficult to evaluate scientific rigor. For example, while reporting their choice of AI technology, some authors reported vaguely, simply stating *NLP or chatbots* while some authors were more specific, reporting specific family of NLP algorithms such as *BERT*.

4.3 What AI algorithms and technologies are used in education? (RQ3)

The most popular AI technologies currently explored in the research are traditional AI algorithms, such as random forest and SVMs (see section 3.3), often employed in the context of predicting academic performance and student dropouts. Additionally, the use of NLP is widespread either in combination with chatbots and conversational agents or with regard to text mining and automatic assessment. Although we acknowledge the growing body of literature regarding generative AI for teaching and learning, the research landscape is still dominated by traditional and basic AI approaches rather than performance-demanding AI technologies. This is perhaps due a limitation of our corpus containing studies only up to early 2023.

4.4 Theoretical and Practical Implications

The SLR findings entail critical theoretical and practical implications for AIED (in general) and ITS (in particular) research and practice. We envision that these implications can serve as input and guide future research initiatives, ensuring significant impact and innovation. We argue that the findings of this SLR make evident the need for: a) transparent communication of the use of AI in education, its potential benefits and capabilities, and its challenges and opportunities to stakeholders and the general public; b) the need for evidence-based approaches and large-scale, longitudinal studies that will provide sufficient and acceptable indications of the benefits of AI in teaching and learning; c) standardization and open access of data for reproduction and cross-examination of findings experimentally; and, d) the opportunity to revisit traditional AI approaches and cutting-edge technologies from the perspective of feasibility and applicability.

5 Conclusion

This paper presented an SLR about the research landscape in AIED that aimed to provide complementary insights to the findings of [20] over the past 5 years, regarding the robustness and rigorousness of methodological approaches applied in the field and in relation to AI technological advancements. To that end, we reviewed 181 research publications published from 2019 to 2023.

In this paper, we retrieved journal publications from three digital libraries (WoS, EBSCO and Scopus). Thus, we acknowledge that we may have failed to include in our review relevant and important publications. To identify AI technologies used in research, we used the descriptions that the authors offered in their publications. Thus, if some AI technology were inadvertently misrepresented, we would not be able to identify it. Similarly, AI technologies that are part of a bigger group or known by different names, might have been not accounted for. For example, BERT models have been categorized as either deep learning or NLP, depending on the authors' descriptions. In the case of reporting population or sample size, we attempted to record the number of participants per study. However, this was not always possible since some papers reported instead population samples in terms of data points or data entries, for example, the number of matriculations per academic year. Although, in some cases, making the connection between data points and participants was straightforward, this was not always the case. We noticed some discrepancies between how authors report their studies' setup and the definitions of these designs, especially regarding experimental studies. We acknowledge that our classification (experimental vs. observational) is restrictive, and one would like to gain further insights regarding diverse study designs employed in AIED research. Therefore, we see a need for future work that will aim to document the methodologies employed in AIED research in relation to the outcomes and robustness of the evidence they produce. At the same time, we acknowledge the complexity, and diversity of study designs in AIED research and the importance of Design-Based Research for the field.

Evidence-based practice indicates the need for building on solid scientific foundations when introducing effective practices into daily life, as a means to eliminate ineffective or plainly wrong approaches that rather rely on tradition, personal beliefs, and assumptions [7]. This is especially important nowadays because rapid technological advances require quick reactions regarding their integration and adoption, as several published calls for evidence regarding the effectiveness and efficacy of AI and education indicate [2, 1, 9]. Creating opportunities for in-situ research in coordination with practitioners, establishing good practices of rigorous reporting and data sharing, and promoting repeatability and reproducibility efforts can contribute towards bridging the gap between research findings and practical implementation.

We acknowledge the opportunities that AI can introduce for modern education. At the same time, we argue that it is critical to establish solid evidence regarding the impact of such technologies to leverage their potential.

Acknowledgements This research was developed within the augMENTOR ("Augmented Intelligence for Pedagogically Sustained Training and Education") project, co-funded by the European Commission under HORIZON-CL2-2021-TRANSFORMATIONS-01-05, project number: 101061509.

References

1. Aitken, M., Ashurst, C., Atabey, A., Briggs, M., Bright, J., Cross, M., Crowcroft, J., Mahomed, S., Margetts, H., Morgan, D., Straub, V., Wong, J.: Generative ai in education: Call for evidence. https://www.turing.ac.uk/sites/default/files/2023-08/generative_ai_in_education_-_the_alan_turing_institute_0.pdf, accessed: 2024-03-01
2. Cardona, M.A., Rodríguez, R.J., Ishmael, K., et al.: Artificial intelligence and the future of teaching and learning: Insights and recommendations (2023)
3. Chen, X., Xie, H., Zou, D., Hwang, G.J.: Application and theory gaps during the rise of artificial intelligence in education. Computers and Education: Artificial Intelligence **1**, 100002 (2020)
4. Chiu, T.K., Xia, Q., Zhou, X., Chai, C.S., Cheng, M.: Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence **4**, 100118 (2023)

5. Crompton, H., Burke, D.: Artificial intelligence in higher education: the state of the field. International Journal of Educational Technology in Higher Education **20**(1), 1–22 (2023)
6. Crompton, H., Jones, M.V., Burke, D.: Affordances and challenges of artificial intelligence in k-12 education: A systematic review. Journal of Research on Technology in Education pp. 1–21 (2022)
7. Dekker, I.: Academic thriving: Optimising student development with evidence-based higher education. Erasmus University Rotterdam (2022)
8. Eegdeman, I., Cornelisz, I., Meeter, M., Van Klaveren, C.: Identifying false positives when targeting students at risk of dropping out. Education Economics **31**(3), 313–325 (May 2023). https://doi.org/10.1080/09645292.2022.2067131, read
9. Holmes, W., Persson, J., Chounta, I.A., Wasson, B., Dimitrova, V.: Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law. Council of Europe (2022)
10. Nagy, M., Molontay, R.: Interpretable Dropout Prediction: Towards XAI-Based Personalized Intervention. International Journal of Artificial Intelligence in Education (Mar 2023). https://doi.org/10.1007/s40593-023-00331-8
11. Ouyang, F., Zheng, L., Jiao, P.: Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. Education and Information Technologies **27**(6), 7893–7925 (2022)
12. Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., et al.: The prisma 2020 statement: an updated guideline for reporting systematic reviews. International journal of surgery **88**, 105906 (2021)
13. Phillips, T.M., Saleh, A., Ozogul, G.: An AI toolkit to support teacher reflection. International Journal of Artificial Intelligence in Education **33**(3), 635–658 (Sep 2023). https://doi.org/10.1007/s40593-022-00295-1
14. Stracke, C.M., Chounta, I.A., Holmes, W., Tlili, A., Bozkurt, A.: A standardised prisma-based protocol for systematic reviews of the scientific literature on artificial intelligence and education (ai&ed). Journal of Applied Learning and Teaching **6**(2), 64–70 (2023)
15. Suzuki, H., Hong, M., Ober, T., Cheng, Y.: Prediction of differential performance between advanced placement exam scores and class grades using machine learning. Frontiers in Education **7**, 1007779 (Dec 2022). https://doi.org/10.3389/feduc.2022.1007779
16. Tahiru, F.: Ai in education: A systematic literature review. Journal of Cases on Information Technology (JCIT) **23**(1), 1–20 (2021)
17. Taskiran, A., Goksel, N.: Automated feedback and teacher feedback: Writing achievement in learning english as a foreign language at a distance. Turkish Online Journal of Distance Education **23**(2), 120–139 (2022). https://doi.org/10.17718/tojde.1096260
18. Xu, W., Ouyang, F.: A systematic review of ai role in the educational system based on a proposed conceptual framework. Education and Information Technologies pp. 1–29 (2022)
19. Zafari, M., Bazargani, J.S., Sadeghi-Niaraki, A., Choi, S.M.: Artificial intelligence applications in k-12 education: A systematic literature review. IEEE Access **10**, 61905–61921 (2022)
20. Zawacki-Richter, O., Marín, V.I., Bond, M., Gouverneur, F.: 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 (2019)