Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions. *Educational Technology & Society*, *25* (1), 28-47.

Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions

Xieling Chen[1], Di Zou[2*], Haoran Xie[3], Gary Cheng[1], and Caixia Liu[1,4]

[1]Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR // [2]Department of English Language Education, The Education University of Hong Kong, Hong Kong SAR // [3]Department of Computing and Decision Sciences, Lingnan University, Hong Kong SAR// [4]Institute of EduInfo Science & Engineering, Nanjing Normal University, China // xielingchen0708@gmail.com // dizoudaisy@gmail.com // hrxie2@gmail.com // chengks@eduhk.hk // cxsqz@126.com

[*]Corresponding author

(Submitted January 12, 2021; Revised March 4, 2021; Accepted March 15, 2021)

ABSTRACT: With the increasing use of Artificial Intelligence (AI) technologies in education, the number of published studies in the field has increased. However, no large-scale reviews have been conducted to comprehensively investigate the various aspects of this field. Based on 4,519 publications from 2000 to 2019, we attempt to fill this gap and identify trends and topics related to AI applications in education (AIEd) using topic-based bibliometrics. Results of the review reveal an increasing interest in using AI for educational purposes from the academic community. The main research topics include intelligent tutoring systems for special education; natural language processing for language education; educational robots for AI education; educational data mining for performance prediction; discourse analysis in computer-supported collaborative learning; neural networks for teaching evaluation; affective computing for learner emotion detection; and recommender systems for personalized learning. We also discuss the challenges and future directions of AIEd.

Keywords: Artificial intelligence in education, Structural topic modeling, Bibliometric analysis, Research topics, Research evolution

1. Introduction

Artificial intelligence (AI), as a machine-based technique with algorithmic power for making predictions, diagnoses, recommendations, and decisions, has grown in importance within the educational community for its potential to support learning in diverse contexts in recent years (Hwang et al., 2020a). The field of AI in education (AIEd) has demonstrated technological advances, theoretical innovations, and successful pedagogical impact (Roll & Wylie, 2016), with diverse applications such as intelligent tutors for content delivery, feedback provision, and progress supervision (Bayne, 2015). The affordances of AIEd are widely recognized. AI can be used to provide specialized support and raise knowledge-gap awareness, which enables instructors to teach effectively and efficiently through personalized and adaptive instruction (Guan et al., 2020). AI also provides algorithm-based decisions which enable effective real-time assessment of complex skills and knowledge (Chen et al., 2021). Additionally, AI-empowered educational systems can be used to analyze classroom dynamics and student engagement, which in turn helps to identify at-risk students in real-time mode, thus enabling timely intervention (Tsai et al., 2020).

Researchers and practitioners have been promoting AI and exploiting its pedagogical potential; consequently, scientific output on AIEd has increased significantly (Hinojo-Lucena et al., 2019). Scientific literature is valuable for thoroughly understanding the history and status of a field and can be analyzed through research motivation identification, scientific collaboration evaluation, and research theme detection (Chen et al., 2020a). Given the rapid growth of AIEd research, a synthesis of the extant literature for a summarized overview appears timely.

Several reviews that applied narrative synthesis or the systematic review of small samples have been conducted. Chassignol et al. (2018) reviewed AIEd literature from four perspectives, i.e., personalized instructional materials, innovative instructional strategies, technology-assisted assessment and communications between learners and instructors, based on 47 publications in the *International Journal of Artificial Intelligence in Education* (IJAIED) in 1994, 2004, and 2014. Roll and Wylie (2016), who explored AIEd's strengths and opportunities, found there was an evolutionary process regarding in-class learning practices and interactions with instructors supported by diversified AI technologies and a revolutionary process regarding AI technologies' adoption in students' daily life and community activities. Zawacki-Richter et al. (2019) systematically reviewed 146 publications about AI in higher education, identifying AI's applications for profiling and prediction, assessment and evaluation, adaptivity and personalization, and intelligent tutoring systems (ITSs) to support academic, institutional, and administrative services.

There are reviews on AIEd based on quantitative methodologies. Goksel and Bozkurt (2019) adopted social network analysis in reviewing AIEd publications from 1970 to 2018. They identified three themes, i.e., adaptivity/personalization and learning styles; expert systems and ITSs; and AI as an integrated component during instruction. Hinojo-Lucena et al. (2019) bibliometrically analyzed 132 AIEd publications from 2007 to 2017; their review showed there was a global interest in AIEd, and the period represented an incipient stage for publications in the area. Chen et al. (2020b) reviewed 45 AIEd-related publications in terms of annual distribution, major journals, institutions, countries/regions, research issues, and the theories and technologies involved in order to highlight gaps in AIEd applications and theory. Guan et al. (2020) analyzed 400 articles on AI and deep learning (DL) in education through manual coding and keyword analysis. Their review indicated increasing interest in implementing and designing online education from 2000 to 2009 and the prevalence of personalized learning supported by learner profiling and learning analytics (LA) from 2010 to 2019. Tang et al. (2021) systematically reviewed publications about the application of AI in e-learning, focusing on leading journals, countries, disciplines, and applications, with a co-citation network analysis examining relations among core-cited references to predict future research directions. Their review revealed that AI-based personalized learning scenarios and student characteristic prediction using Bayesian networks were prevalent. These reviews, however, have mostly adopted qualitative methods, with limited studies analyzed and specific results discussed, failing to present a thorough understanding of the general field, particularly about research topics and topic evolution. Such traditional analysis of the full contents of a publication through manual coding and synthesis, however, is time-consuming and laborious, and as the published literature rapidly increases, is becoming outmoded.

Given the prevalence of AIEd and the lack of a quantitative analysis of its copious literature, a review providing a comprehensive understanding of AIEd using rigorous machine learning (ML) appears timely. Owing to ML's rapid development, diverse approaches capable of analyzing large volumes of data are now available, among which topic models are effective and efficient for inferring latent topics from large amounts of literature (Chen et al., 2020a). The inferred information reveals a better understanding of historical and extant research progress, development of technologies applied, and drivers of fresh ideas, all of which can help researchers and educators decide upon research topics and project planning.

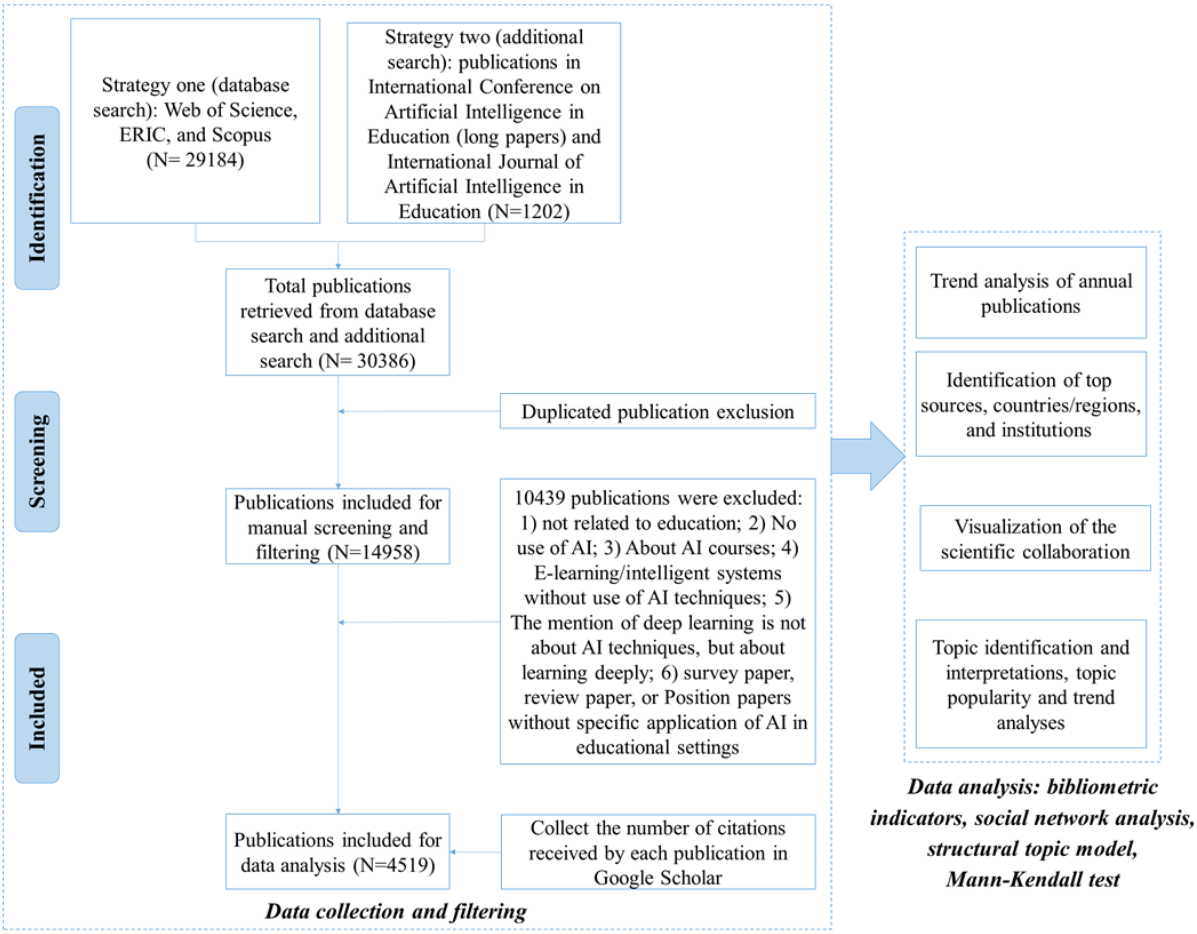
Accordingly, we applied topic-based bibliometrics to quantitatively examine 4,519 AIEd literature from 2000 to 2019 to uncover topic trends and predict the future of AIEd, focusing on the following: changes in topic popularity; major publication sources, countries/regions and institutions; and scientific collaborations. Our review was guided by five research questions:

RQ1: What were the number of AIEd articles published from 2000 to 2019?
RQ2: What were the top publication sources, countries/regions, and institutions?
RQ3: What was the nature of collaboration among countries and institutions?
RQ4: What were the most investigated research topics?
RQ5: How did the intensity of research interest in these topics change?

2. Dataset and methods

Figure 1 depicts the steps of data collection and analysis. Detailed descriptions follow:

Figure 1. Data collection and analyses



2.1. Data retrieval and preprocessing

AIEd-related publications from 2000 to 2019 were collected on May 30th, 2020 using two strategies. First, Web of Science (WoS), Scopus, and Education Resources Information Center (ERIC) databases were searched. Two lists of search terms were considered, including AI-related terms ("artificial intelligence," "machine intelligence," "intelligent support," "intelligent virtual reality," "chat bot*," "machine learning," "automated tutor*," "personal tutor*," "intelligent agent*," "expert system*," "neural network*," "natural language processing," "chatbot*," "intelligent system*," and "intelligent tutor*") and education-related terms ("education," "college*," "undergrad*," "graduate," "postgrad*," "K-12," "kindergarten*," "corporate training*," "professional training*," "primary school*," "middle school*," "high school*," "elementary school*," "teaching" and "learning"). Specifically, in WoS, "TS" was searched with AI-related terms to include research articles and conference papers written in English, and these were categorized in *Education Educational Research*. In Scopus, "TITLE-ABS-KEY" was searched with AI- and education-related terms to include articles published in journals and conference proceedings, written in English, categorized in Social Sciences and further restricted to publication sources with "education*," "teaching," "learning," or "instruction*" in their names. In ERIC, titles and abstracts were searched using individual AI-related terms, with the results being aggregated and duplicated. The first strategy identified 29,184 publications.

Second, considering the close relevance of the International Conference on Artificial Intelligence in Education (ICAIED) and IJAIED to our research target, we conducted an additional search in these two and obtained 1,202 publications.

The 30,386 publications were checked for duplication via title comparison by calculating string similarity using the Python package called *strsim*. After calculation, titles of publications with a similarity degree equal to "1" were duplicated, with the rest being sorted in descending order of similarity values for manual checking. Specifically, for the 29,184 publications retrieved using strategy one, title comparisons of Web of Science and ERIC, Web of Science and Scopus, and ERIC and Scopus were conducted to eliminate duplications. Thereafter, titles of the remaining publications were compared against the 1,202 publications retrieved using strategy two to delete duplications, resulting in 14,958 publications for further data screening.

Two domain experts adopted the criteria in Figure 1 to determine publication relevance. They each assessed the same 300 randomly selected articles independently, leading to inter-rater reliability of 91%, with inconsistencies being discussed to resolve differences. Thereafter, they screened the remaining data separately, resulting in 4,519 eligible publications whose numbers of citations were collected in Google Scholar (see https://scholar.google.com/).

Preprocessing included manual supplementation of publication features, including the author's address information by referring to original full-texts and the identification of authors' institutions and their corresponding countries/regions. To analyze research topics, terms were extracted from titles, abstracts, and keywords with a weighting strategy (Chen et al., 2018). Additionally, term frequency-inverse document frequencies with a threshold of 0.05 was conducted for term selection.

2.2. Data analysis

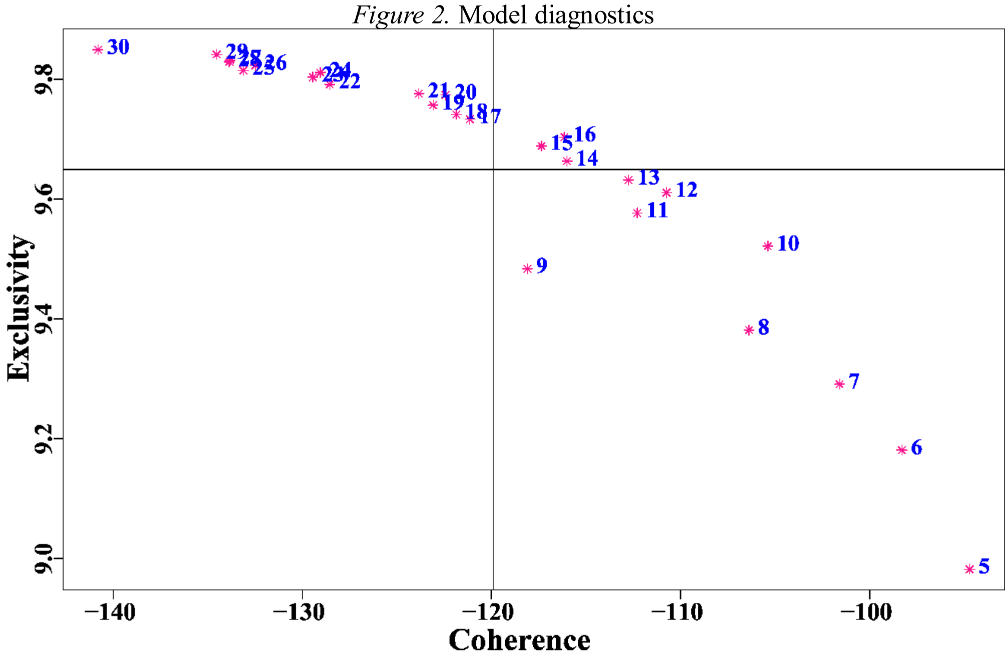
Four methodologies (i.e., bibliometric indicators, social network analysis, structural topic modeling (STM), and Mann-Kendall (MK) trend test) were used.

First, the publication count measured annual productivity. A polynomial regression analysis was further conducted to determine the developmental trend of AIEd research. Publication sources, countries/regions, and institutions were analyzed using publication count and the Hirsch index (H-index) to measure productivity and impact.

Second, social network analysis via Gephi (see https://gephi.org/) visualized relationships between institutions or countries/regions by treating institutions or countries/regions as nodes with the node size indicating their productivity and the link width indicating collaboration intensity.

Additionally, research topics in the 4,519 publications were identified using STM (Roberts et al., 2014; Roberts et al., 2019). We ran 26 models with the number of topics ranging from five to 30. Three models with 14, 15, and 16 topics each achieved higher values of semantic coherence and exclusivity measures (see Figure 2). For them, two domain experts conducted comparisons by examining representative terms and studies. The model with 16 topics (i.e., 16-topic model) was identified as it produced "the greatest semantic consistency within topics and exclusivity between topics (Chen et al., 2020a, p. 4)." To examine how the intensity of research interest in each topic changed over time, we employed the MK test (Mann, 1945).

Figure 2. Model diagnostics



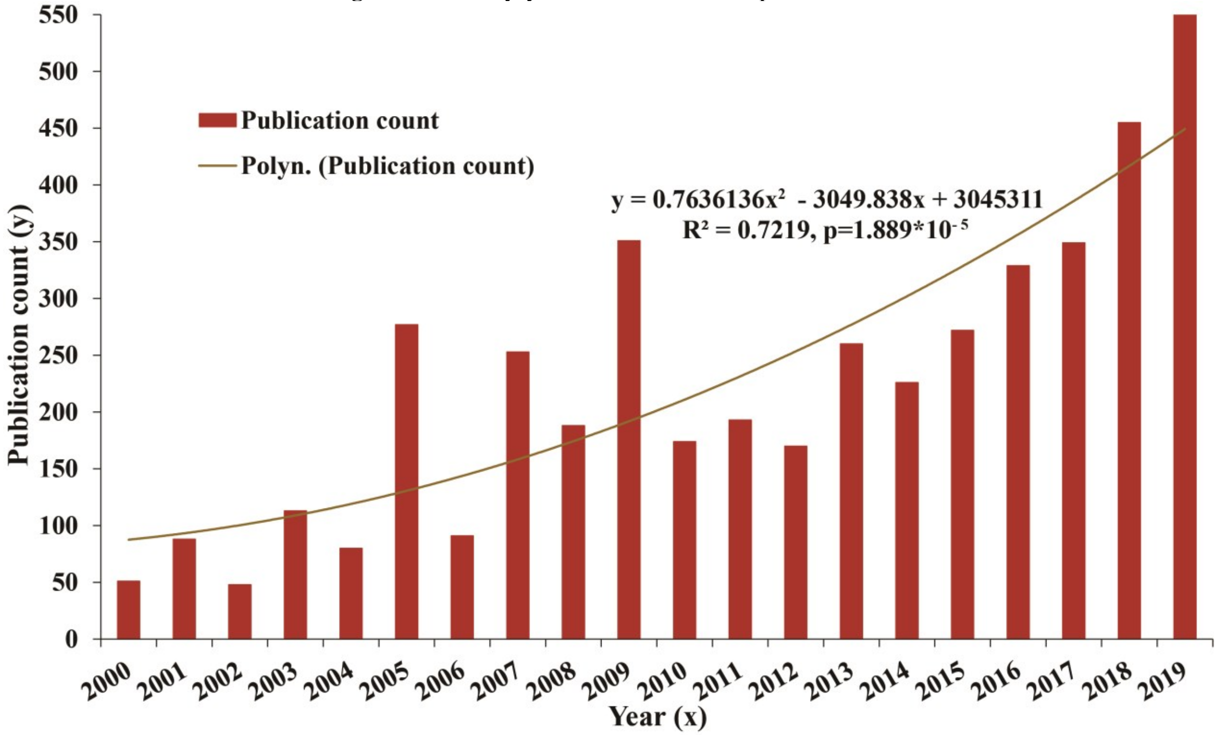
Note. Each node represents a topic model with blue labels indicating the number of topics.

3. Results

3.1. Annual numbers of AIEd publications

Figure 3 shows the number of AIEd articles published from 2000 to 2019, indicating an overall increasing tendency, particularly since 2012. The increasing interest in AIEd research is mainly due to the increased positive findings of AI's effects on learning performance and outcomes.

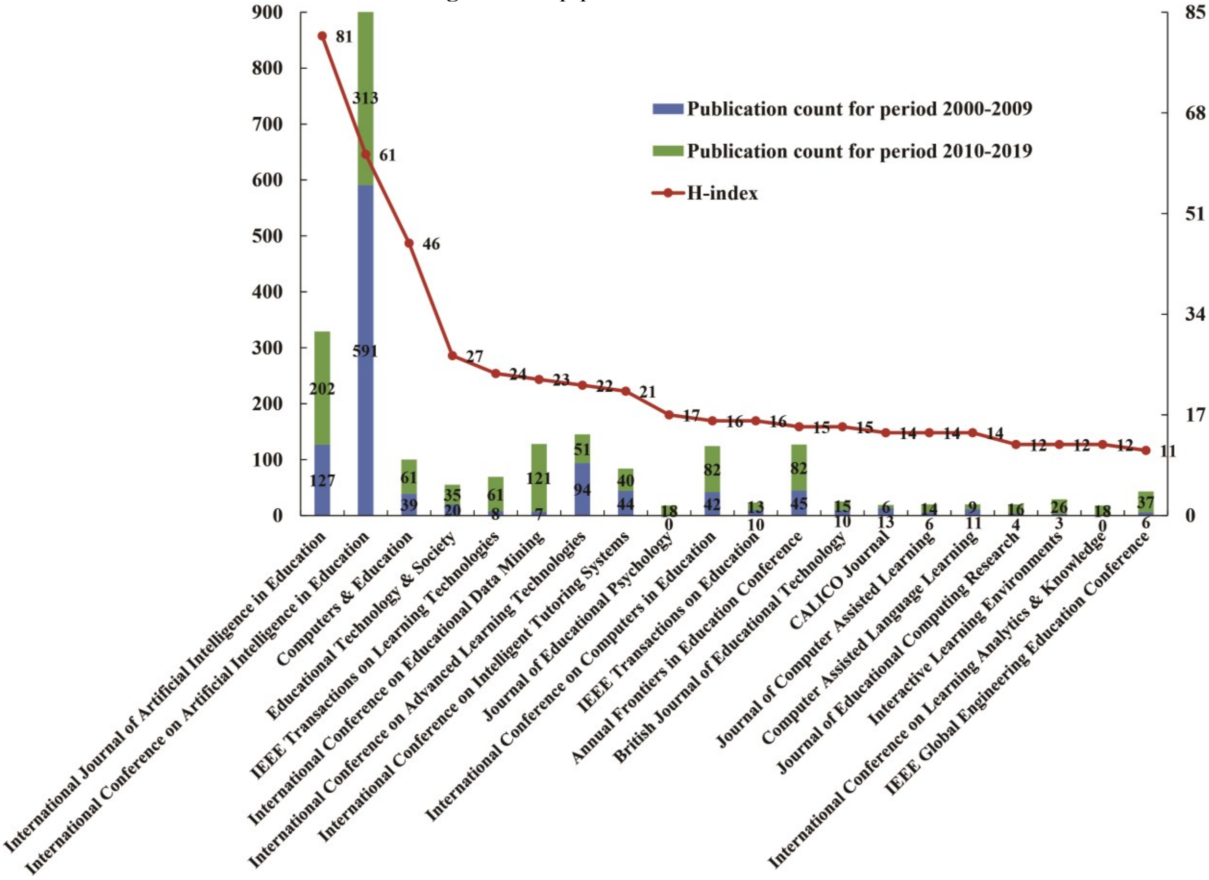
Figure 3. Year-by-year number of AIEd publications



3.2. Top publication sources

In total, 650 sources were identified, with the top 20 ranked by H-index (Figure 4) contributing to over 50% of the total. Eight were conferences, with IJAIED at the top with an H-index of 81 and 329 publications, followed by ICAIED, *Computers & Education*, and *Educational Technology & Society*. Comparing the publication counts of the first decade with the second, most sources became increasingly interested in AIEd in the latter.

Figure 4. Top publication sources



3.3. Top countries/regions and institutions

In total, 92 countries/regions were identified, with the top 20 ranked by H-index (Figure 5). The USA was at the top with 1,700 publications, 54,344 citations, and an H-index of 102. Based on the H-index, other important countries/regions included Canada, the UK, and Taiwan. We identified 2,296 institutions (top 20 in Figure 6), with Carnegie Mellon University, the University of Pittsburgh, and the University of Memphis holding the top three positions. Measured by publication count, the top three were Carnegie Mellon University, Arizona State University, and the University of Pittsburgh. Most countries/regions and institutions became increasingly interested in AIEd over the period.

Figure 5. Top countries/regions

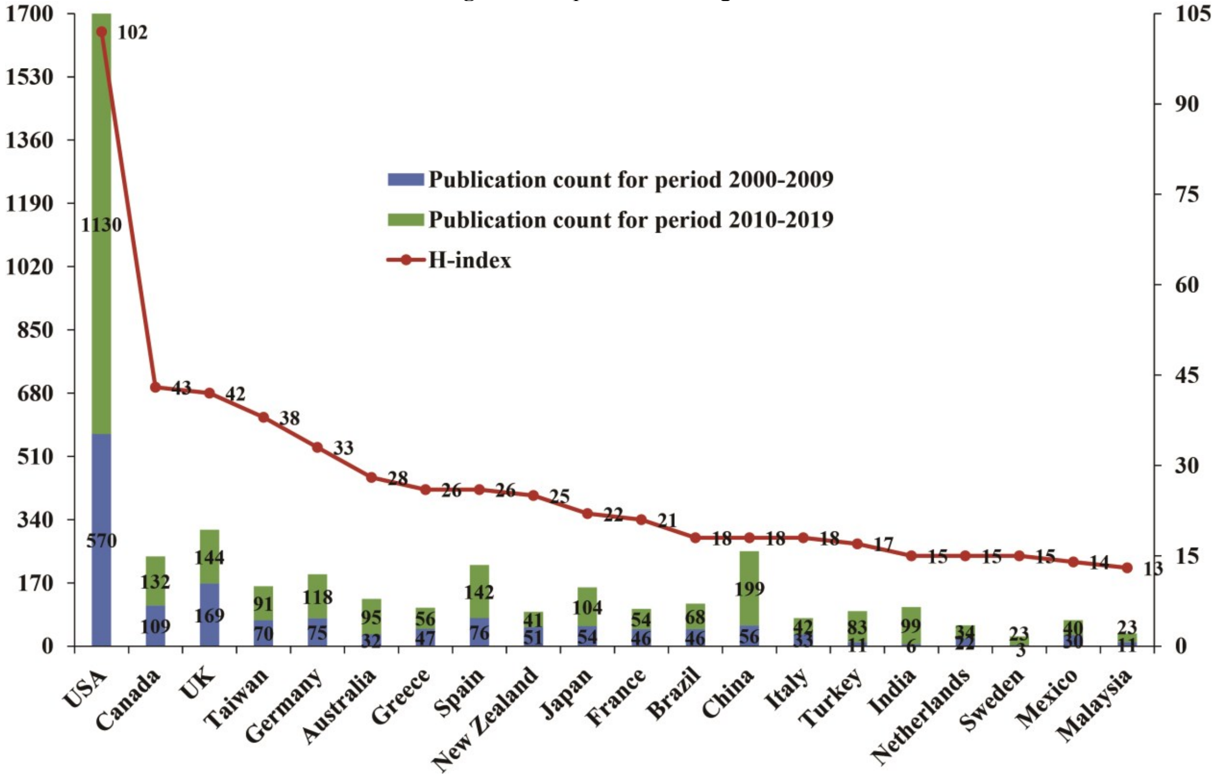
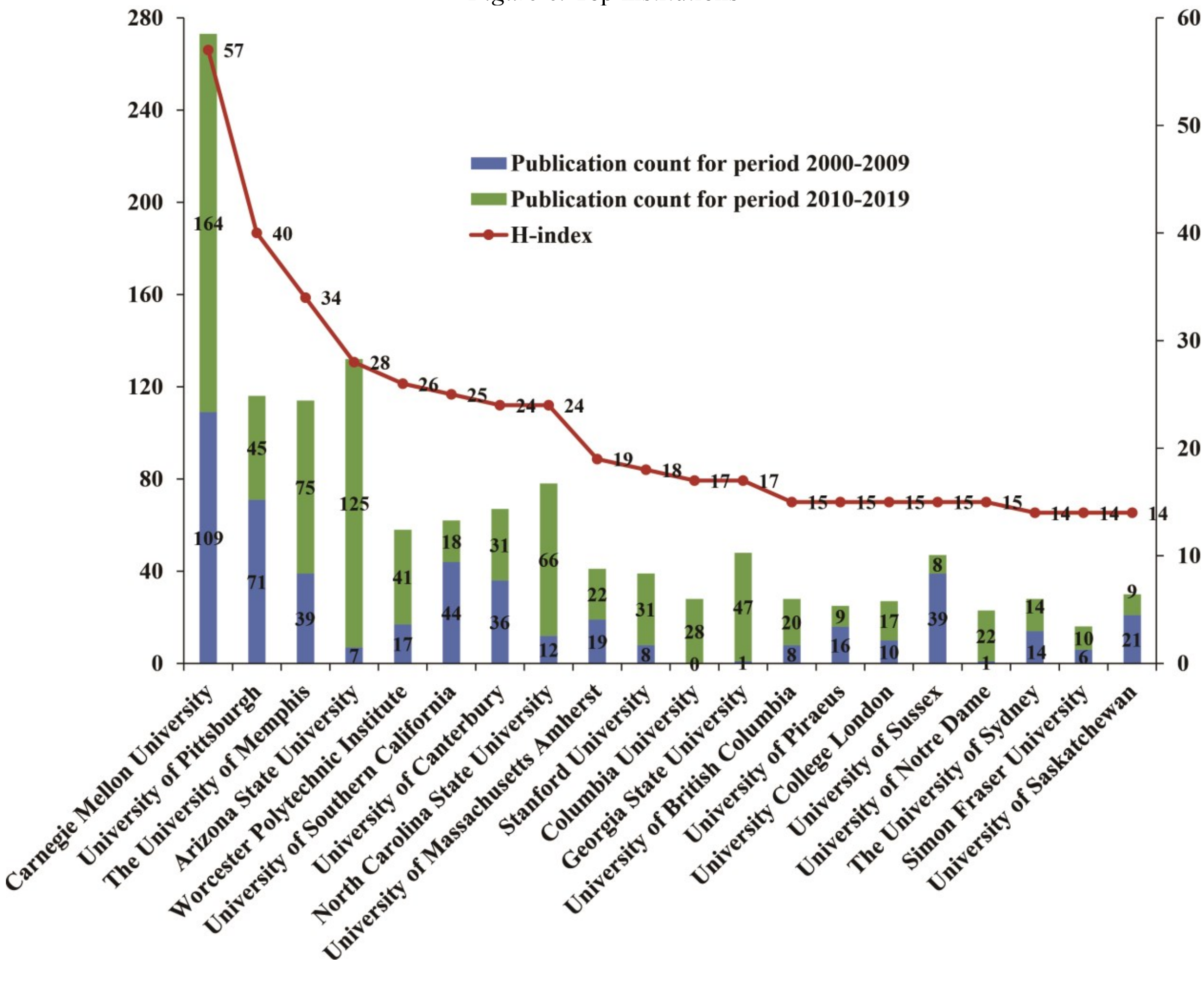


Figure 6. Top institutions



3.4. Scientific collaborations

Collaborations among the top countries/regions are visualized in Figure 7. The USA, the UK, Canada, Spain, and Australia were the most collaborative, with the USA and Germany being the closest partners. From an institutional perspective (Figure 8), Carnegie Mellon University, Arizona State University, and the University of Southern California were the most collaborative, with Georgia State University and Arizona State University being the closest partners.

Figure 7. Collaborations among top countries/regions

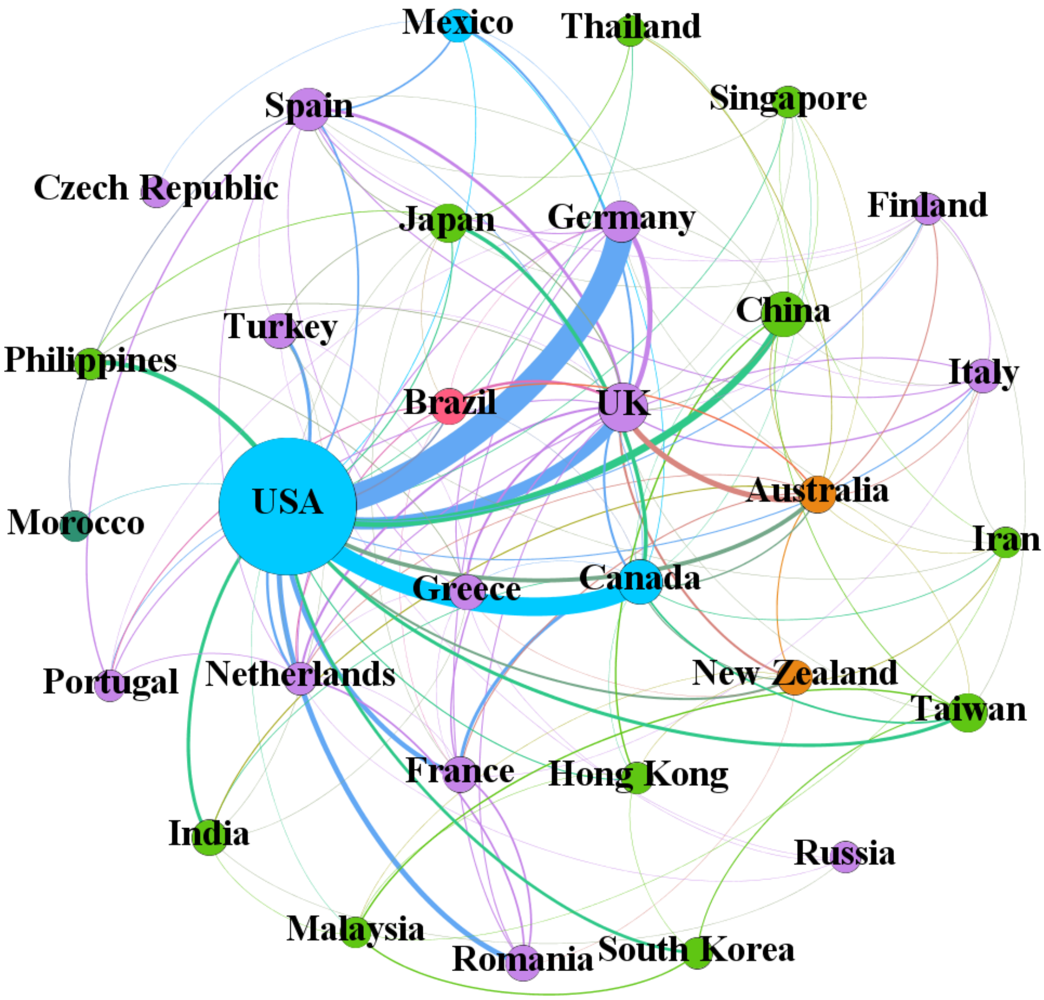
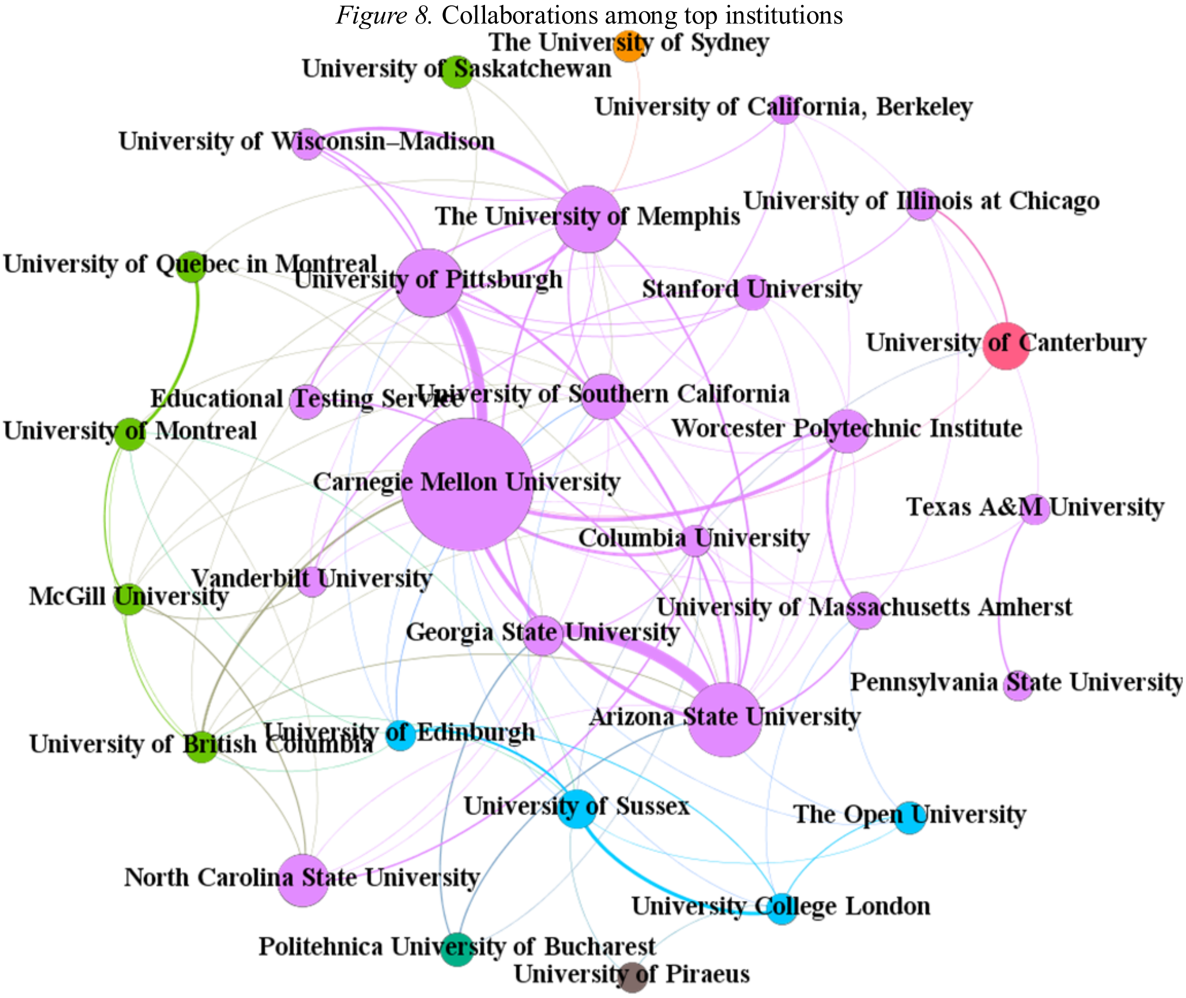


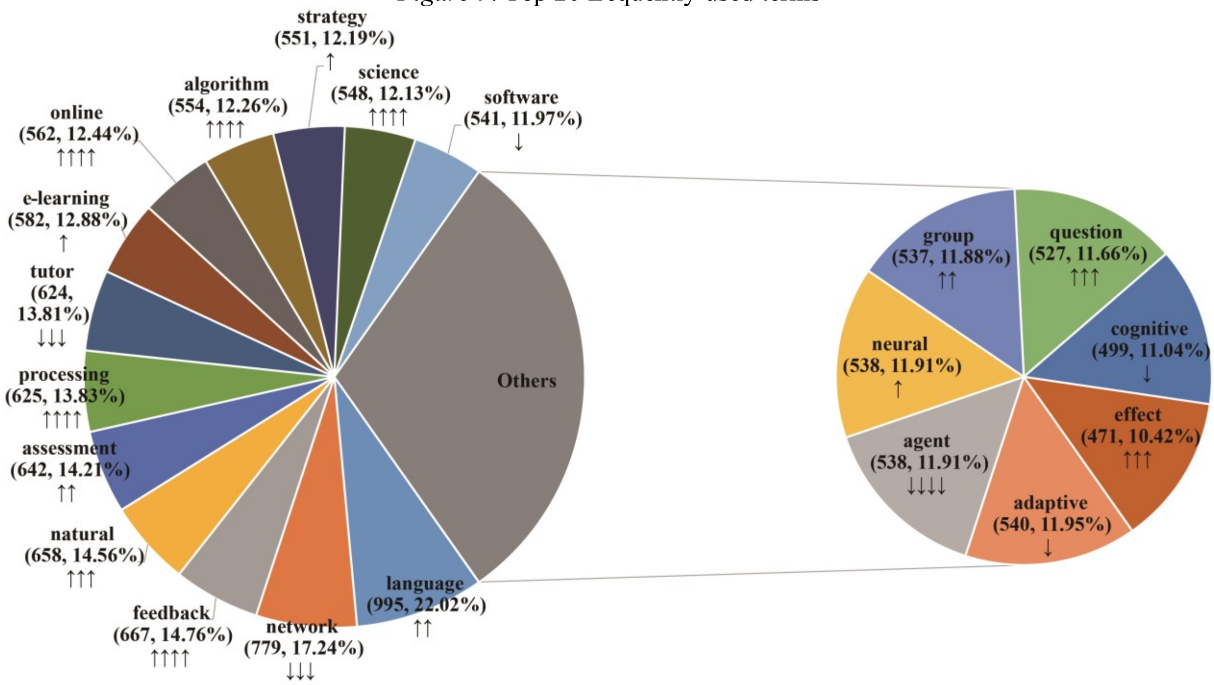
Figure 8. Collaborations among top institutions



3.5. Most frequently-used terms

Figure 9 shows the top 20 most frequently-used terms, with "language" at the top appearing in 995 publications. Other important terms included "network," "feedback," "natural" and "assessment." The trend test indicated that terms like "language," "feedback," "natural," "assessment," "processing," "online," "science," "group" and "question" experienced significant increases over the period.

Figure 9. Top 20 frequently-used terms

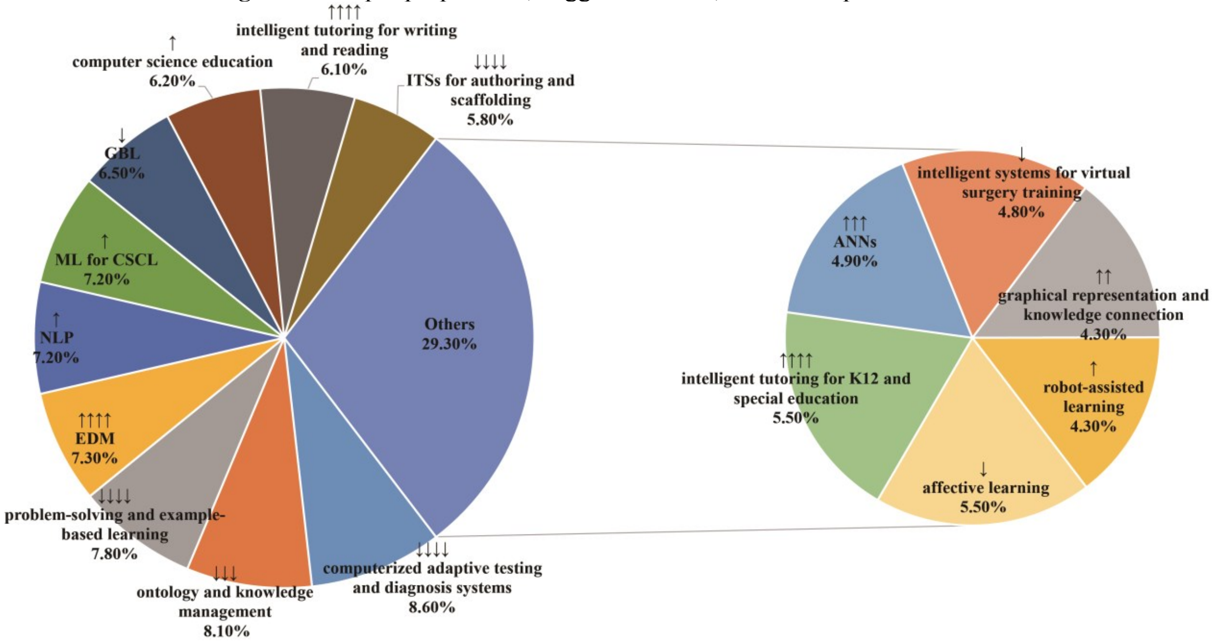


Note. inside the parentheses are term occurrence and proportion; ↑(↓): increasing (decreasing) trend but not significant ($p > .05$); ↑↑(↓↓), ↑↑↑(↓↓↓), ↑↑↑↑(↓↓↓↓): significantly increasing (decreasing) trend ($p < .05$, $p < .01$, and $p < .001$, respectively)

3.6. Research topics and topic trend

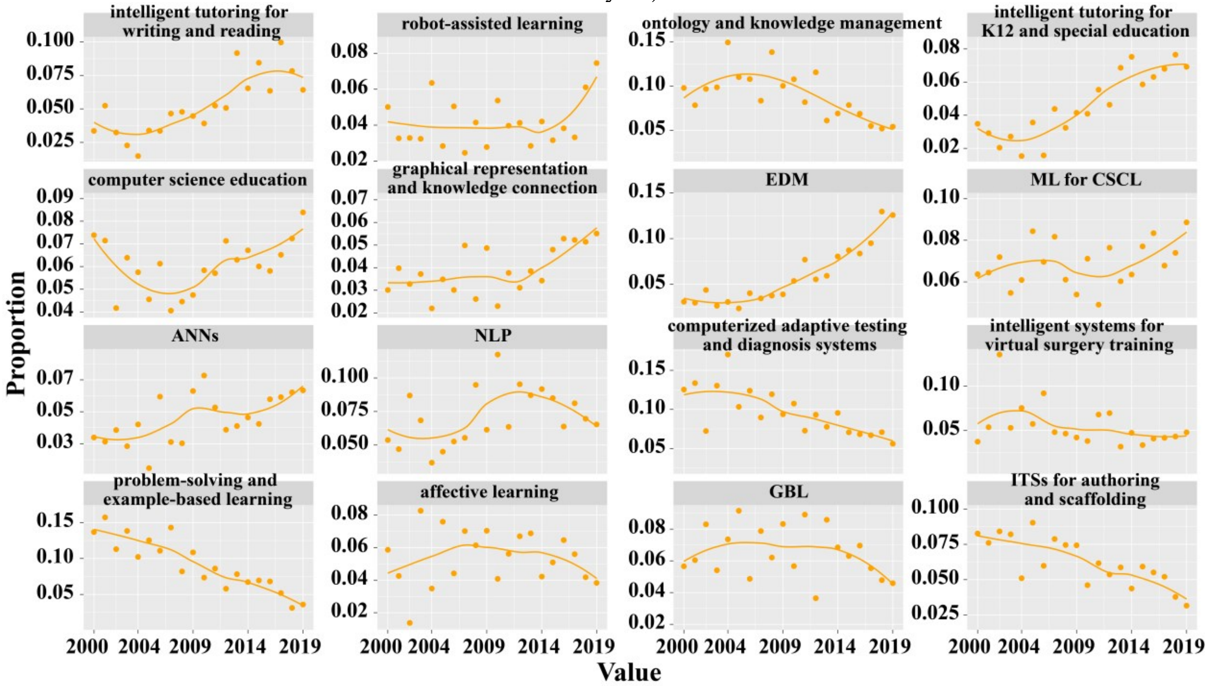
Figure 10 shows the results of the 16-topic model together with suggested labels, topic proportion, and trend test results. Five topics (i.e., *educational data mining* (*EDM*), *intelligent tutoring for writing and reading*, *intelligent tutoring for K12 and special education*, *artificial neural networks* (*ANNs*), and *graphical representation and knowledge connection*) enjoyed a significant increasing trend, whereas four topics (i.e., *computerized adaptive testing and diagnosis systems*, *ontology and knowledge management*, *problem-solving and example-based learning*, and *ITSs for authoring and scaffolding*) experienced a significant decreasing trend over the two decades. Figure 11 illustrates the annual topic proportion, indicating how popular each topic was in each year. Specifically, in the early years, AIEd scholars focused mainly on ontology use and knowledge management in ITSs to facilitate problem-solving and example-based learning for scaffolding purposes where computerized testing and diagnosis of learner knowledge and learning processes were frequently concerned. In the later years, articles on learner affect and emotion in diverse scenarios became more frequent, especially in GBL, where learners commonly experienced diverse emotions that directly impacted learning performance. Also, ITSs gradually extended their applications to facilitate the learning of diverse subjects, particularly NLP-assisted language education and in K12 and special education. Furthermore, increasingly diverse technologies were used for various educational goals, e.g., robot-assisted computer science education, ML-assisted CSCL, and ANN-assisted learning prediction and teaching evaluation. Additionally, EDM and LA were increasingly applied to visualize the learning process and knowledge acquisition for easy understanding. These foci in the later years point towards the future and challenges faced by AIEd scholars.

Figure 10. Topic proportions, suggested labels, and developmental trends



Note. % indicates topic proportion; ↑(↓), ↑↑(↓↓), ↑↑↑(↓↓↓), ↑↑↑↑(↓↓↓↓) are similar to *Figure 9*

Figure 11. Annual topic proportion (x-axis indicates publication year and the y-axis indicates topic popularity in each year)



4. Discussion

Focusing on the research questions, this section discusses the findings. For RQ1, consistent with previous reviews (Chen et al., 2020b; Hinojo-Lucena et al., 2019; Roll & Wylie, 2016; Tang et al., 2021; Zawacki-Richter et al., 2019), the overall growth of AIEd literature indicates a positive future with an expanding community and scientific output. Responding to RQ2, AIEd research is especially welcomed by interdisciplinary journals such as *Computers & Education* and *Educational Technology & Society* with their dual foci on education and technology; these journals are also highly ranked in publishing AI in e-learning studies (Tang et al., 2021). The results support Zawacki-Richter et al. (2019) and Tang et al. (2021), who highlighted AIEd's close relationship with computer science and software engineering. Consistent with Hinojo-Lucena et al. (2019) who identified the

USA as the dominant actor, our study further revealed that scholars in a variety of countries/regions (e.g., Canada, the UK, and Taiwan) and institutions were increasingly interested in AIEd. The higher AIEd research productivity in these countries/regions can be partially attributed to their governments' efforts to promote technology-enhanced learning through educational policy and funding (Chen et al., 2020c). In Tang et al., Taiwan was the top country, whereas, in our study, it was ranked 7th, which may reflect our wider focus on AI's application in education as a whole, rather than in e-learning alone. Carnegie Mellon University was the top in research productivity and impact. Responding to RQ3, the network visualization (Figures 7 and 8) revealed that the countries/regions and institutions that had intense scientific collaborations showed higher productivity and wider impact. We thus call for enhanced international collaborations to better embrace challenges as AIEd advances. Additionally, AIEd's interdisciplinarity was uncovered by the topic modeling, demonstrating effective and important AI technologies that originated from computer science.

The STM results respond to RQ4, revealing frequently occurring issues throughout the review period. These include computerized adaptive testing, diagnosis, and instruction systems integrated with varied AI technologies, especially NLP, ontology, ML, and ITSs. All of these facilitate diverse educational goals such as subject knowledge (e.g., language skills and programing) and ability (e.g., problem-solving) acquisition and innovative pedagogical strategy implementation (e.g., GBL and example-based learning). Consistent with several reviews (Chassignol et al., 2018; Guan et al., 2020; Tang et al., 2021; Zawacki-Richter et al., 2019) that identified the important roles of ITSs and AI in assessment, feedback, and learner performance prediction, we highlighted ITSs' popularity in various domain-specific types of education (e.g., K-12 education and language education), AI for computerized adaptive testing and diagnosis, and learner performance tracking and prediction using EDM. Similar to Roll and Wylie (2016) who highlighted AI's role in supporting collaborations in interactive learning, we identified CSCL assisted by ML, a technique also identified in Chen et al. (2020b). Consistent with Chassignol et al. (2018) who identified educational robots, Guan et al. who identified educational games and teaching evaluation, and Tang et al. (2021) who noted Bayesian networks and neural networks for learner learning characteristic prediction, we highlighted robot-assisted learning, GBL, and neural network-assisted teaching evaluation. Similar to Chen et al. (2020b) who identified NLP, we further highlighted its importance in language education. Just as several reviews (Chassignol et al., 2018; Guan et al., 2020; Tang et al., 2021; Zawacki-Richter et al., 2019) have identified the growing interest in AI-assisted personalization, we also noted the interest in personalization in adaptive testing and diagnosis. Roll and Wylie and Tang et al. (2021) highlighted an increasing interest in domain-level learning such as language and medical education and STEM (science, technology, engineering, and mathematics) education assisted by AI; we also revealed AI's use in various subjects and domains (e.g., computer science education, language education, K12 and special education, and surgery training). We additionally identified new topics such as problem-solving, example-based learning, authoring and scaffolding, and affective learning.

The findings of the topic analyses, and especially the trend analysis, answers RQ5, revealing there was a decreased interest in ITSs for authoring and scaffolding, whereas ITSs were increasingly used for NLP-assisted language education and K12 and special education. As for AI technologies, ontology use declined, whereas advanced techniques such as ML, ANNs, EDM gained popularity in scenarios such as CSCL; however, these were less popular in problem-solving and example-based learning. Compared to computerized testing and diagnosis, how AI facilitates subject knowledge acquisition became prevalent over the review period (e.g., robot-assisted programming education). These findings bring insight into important issues and the potential future directions of AIEd. We established eight themes by examining and interpreting the topics receiving increasing interest. For example, when considering the most representative studies of the topic *EDM* centering on EDM-assisted learning prediction, we formed a theme called "EDM for performance prediction." "NLP for language education" was established by examining three topics (i.e., *NLP*, *intelligent tutoring for writing and reading*, and *graphical representation and knowledge connection*) whose representative studies focused mainly on NLP use in language education. Other themes were formed similarly, except for "Affective computing for learner emotion detection" and "Recommender systems for personalized learning." The former was selected based on topic *affective learning*, which, although not found to increase significantly in popularity, was widely reported to facilitate instruction, particularly regarding personalization. The latter theme was included because of its prevalence in the data corpus. Although it was not identified as a separate topic due to topic overlaps in topic models, personalized learning was increasingly prevalent (e.g., Chassignol et al., 2018, Zawacki-Richter et al., 2019; Guan et al., 2020), particularly in the form of personalized material recommendation. Hereafter, tightly aligning to the eight themes, we discuss AIEd's challenges and the future effort needed to advance the field.

4.1. ITSs for special education

As an adaptive instructional system incorporating AI into educational methods, ITSs have been widely applied in various domains (e.g., STEM education, computer science education, and language education) with benefits and positive effects well documented. Instead of repeating what has been found in previous reviews, we would like to highlight an emerging need for the application of ITSs in special education, particularly among autistic students. ITSs' effectiveness for teaching autistic students owes much to their ability to provide immediate and personalized instruction and feedback, which is as effective as one-to-one tutoring. This overcomes the difficulties in anticipating and recognizing autistic students' negative behaviors (Mondragon et al., 2015). An integrative specialized learning application (ISLA) (Mondragon et al., 2016) can help autistic students manage emotions using learning trace analysis and learning performance evaluation. In ISLA, a virtual agent named Jessie adjusts an autistic learner's emotional state in real-time and provides personalized encouragement and support to assist problem-solving during learning. This feedback relieves the autistic learners' anxiety and frustration while keeping them engaged.

ITSs also benefit autistic learners in performing real-time learning tasks by monitoring and intervening when necessary. An intelligent LEGO tutoring system (Sun & Winoto, 2019) assists both instructors and autistic learners in brick playing. In the instructors' module, instructors design a new model of LEGO bricks; thereafter, visual and auditory step-by-step instructions for model completion are automatically generated. In the learners' module, the designed model is loaded with displayed instructions. Such systems benefit autistic learners by prompting instructors with necessary interventions and instructions while tracking learners' brick-building process in real-time with feedback and suggested corrections automatedly provided when a mistake is made.

4.2. NLP for language education

NLP is instrumental for computer-assisted language learning (CALL). First, many new CALL applications integrate various automatic speech recognition technologies to create realistic and engaging learning experiences by enabling computers to understand learners' speech and react accordingly or provide feedback on speech quality (Zhang & Zou, 2020). One recent call to researchers is to develop speech-to-text algorithms enabling seamless integration of speech recognition systems to enhance learners' real-time understanding of their adopted reading strategies for oral self-explanations on a given text (Panaite et al., 2018).

Second, word sense disambiguation facilitates effective vocabulary learning by resolving lexical ambiguity via automatically ordering dictionary definitions or assigning an appropriate meaning to a given context (Rosa & Eskenazi, 2011). In Eom (2012), a captioning tool facilitates listening by providing cues for ambiguous or difficult words, where a word sense disambiguation tool finds suitable definitions for words with multiple meanings.

Third, part of speech (POS) tagging is increasingly needed for language learners for effective word processing. The popularity of POS is mainly because of its ability to provide helpful information (e.g., language morphology, syntax, and phonology) to improve language proficiency (Hamouda, 2013). In an Indonesian computer-assisted self-learning system (Muljono et al., 2017), a POS, tagging with a hidden Markov model, deals with ambiguity by reducing tagging errors in unknown words.

Additionally, NLP also facilitates automatic feedback, i.e., grammar correction and writing evaluation and translation. In Lee et al. (2015), Genie Tutor assists English learning by identifying grammar mistakes and providing correction suggestions. With Genie Tutor, language learners know their mistakes in real-time and learn native expressions. An automatic translation chatbot (Sato et al., 2018) offers different types of second language translation along with first language texts during online interaction. By providing second language input and reducing learners' doubts about their second language competence, the chatbot lowers learners' anxiety and facilitates their language performance and motivation during online collaborations.

4.3. Educational robots for AI education

Educational robots are useful for motivating learners and solidifying abstract and complex topics (e.g., AI education). In Martínez-Tenor et al. (2019), Lego® Mindstorms robots teach reinforcement learning algorithms in a cognitive robotics course. Learners engage in lab exercises by implementing reinforcement learning in coding programs to control real robot movements (e.g., simple wandering, backward/forward motion, and detecting and avoiding obstacles). By converting reinforcement learning theory into real-world problems, learners create their own learning experiences by engaging with both theoretical algorithms and physical implementations. SyRoTek (Kulich et al., 2012) allows remote access to fully autonomous mobile robots placed around reconfigurable obstacles. With SyRoTek, learners control the robots in real-time using self-developed algorithms and then observe how the real robots behave through live videos, thus improving their problem-solving ability by integrating theory into practice.

4.4. EDM for performance prediction

Predicting student performance is important in EDM for mining meaningful patterns and knowledge from large-scale educational data using ML and data mining. EDM's effectiveness in learning to predict has been widely reported. Typical prediction scenarios include academic performance, learner enrolment, dropouts, retention, and early detection of at-risk learners. As for data used for predicting attrition, transcript-based features outperform those based on learner histories prior to college (Aulck et al., 2019). Features derived from institutions' routine data are effective for graduation and re-enrolment prediction. Considering algorithmic performance, Beaulac and Rosenthal (2019) highlight random forests' effectiveness for different prediction tasks, including the prediction of the number of registered learners in future years, learner distribution prediction across programs and at-risk learner identification (e.g., academic failure or dropping out).

4.5. Discourse analysis in CSCL

Collaborative dialogue analysis is essential for facilitating CSCL (Lin & Chan, 2018) as it promotes an understanding of the collaborative process and enables tailored interventions and appropriate scaffolding (Dowell et al., 2019). Jointly using time series analysis and semantic similarity can filter online discourse to identify learners' key collaborative moments (Samoilescu et al., 2019). Informed by the degree of collaboration, which is automatically assessed among learners in their conversations, instructors can provide feedback to promote learner involvement and collaboration in CSCL. Focusing on facilitating large-scale collaborative dialogue data analysis, Shibata et al. (2017) train and test an automatic coding approach based on DL, showing DL's superiority over naive Bayes and support vector machines for supporting authentic learning through monitoring and scaffolding non-activated groups in real-time.

4.6. Neural networks for teaching evaluation

With the rapid growth of higher education, teaching quality has been put in the spotlight. ANNs are revolutionizing teaching quality evaluation by avoiding human subjectivity to enhance evaluation accuracy and effectiveness (Hongmei, 2013). Such neural network-driven models can be further enhanced by particle swarm optimization for weight optimization and modification in accuracy calculation during model training (Rashid & Ahmad, 2016).

4.7. Affective computing for learner emotion detection

Affect in learning is receiving more attention to better understand learner emotions and cognition and to provide affective intervention and support to increase learner self-concept and motivation (Hwang et al., 2020b). Two affective computing techniques (i.e., emotion recognition from physiological or facial expression data and emotion recognition from texts) are widely embodied in ITSs. In Mehmood and Lee (2017), special school instructors teach learners with emotional disorders using wearable sensors and intelligent emotion detection technologies to identify useful information from brain signals. Then, the learners' feelings (i.e., happiness, calm, sadness, and fear) are extracted from the information and processed using support vector machines and near k-neighbor classifiers. In Su et al. (2016), emotions are identified through joint use of facial expression detection and textual sentiment analysis. Such a combined strategy strengthens recognition effectiveness and allows the detection of diverse emotions to facilitate personalized instruction and curriculum content provision.

4.8. Recommender systems for personalized learning

Recommender systems are increasingly integrated into ITSs to generate personalized recommendations about learning resources and paths by considering learners' background knowledge, behavioral preferences, profiles, and interests (Ma & Ye, 2018). In Liu et al. (2018), learners' quiz scores and multi-modal sensing data (i.e., heartbeats, blinks, and facial expressions) are measured to track learning processes and generate personalized guidance based on their present learning states. Such personalized systems can be improved by modifying dynamic key-value memory to design memory structures based on the course's concept list, plus by mapping exercise-concept relations during learners' knowledge tracing (Ai et al., 2019). This helps build learner simulators for exercise recommendation policy training to maximize learners' knowledge level through deep reinforcement learning.

4.9. Challenges and the future of AIEd

This section discusses the challenges existing within the above-discussed themes and points towards future efforts needed to resolve such challenges.

4.9.1. Personalization versus data privacy

The global prevalence of personalized learning calls for more investigations into AI's most effective use to support personalized learning (e.g., adaptively recommending learning materials and scaffolding learners' problem-solving) (Chen et al., 2021). However, to provide personalized experiences, large-scale learner data, which is sometimes highly personal, are required for AI model training. Models sometimes inadvertently store training data with sensitive information that is revealed through model analysis. However, an ML model's potential can only be realized by analyzing learners' data (Chan & Zary, 2019). Since most established models cannot guarantee output models' generalization away from individual learner specifics, plus the uncertainty of data protection places learners' data at risk and lowers AI societal acceptance, there is a need to limit instructors' access to learners' data to meaningfully bound learners' exposure to instructors' knowledge. Educational institutions should be transparent about learner data privacy practices to alleviate data use misperceptions and concerns.

4.9.2. Challenges and ways to increase instructors' AI acceptance

AIEd aims to use AI to facilitate the instruction process (e.g., understand and facilitate CSCL through discourse analysis and achieve performance prediction through EDM), during which instructors are essential, and their acceptance of AI is important. However, as AI is a relatively new concept for instructors, less-experienced instructors usually struggle to execute effective, on-the-spot responses to analytics from AI-empowered applications (Holstein et al., 2017), leading to their reluctance and lower acceptance of AI (Lin et al., 2017). This hinders AI's pedagogical purpose; thus, the improvement of instructors' acceptance of AI systems appears essential.

One way to enhance instructors' confidence in AI is to show the effectiveness of AI systems via robust experiments, particularly under the guidance of time-honored educational theories and philosophies. However, most current AIEd studies fail to positively assess AI system effectiveness through experiments that compare AI's use and traditional instructions (Zawacki-Richter et al., 2019). Albacete et al. (2019) evaluate the effectiveness of Rimac, a natural-language tutoring system capable of dynamically updating learner models, by comparing it with its control version without the updating function. Such an experimental design is challenging due to strict requirements, especially for AI system evaluation, where large samples are required to generate probabilistic results. Additionally, pre- and post-tests are fundamental to objective analysis, and participants should have a similar knowledge level before interventions. Consequently, the effectiveness of AI-driven educational systems is seldom assessed. Nevertheless, such experimental comparisons are indispensable for enhancing instructors' confidence in AI. Researchers should also reach beyond examining how AI improves subject-related outcomes to examining the effectiveness of systems in improving specific abilities (e.g., self-efficacy and higher-order thinking). Thus, in line with Tang et al. (2021), we suggest further investigations into AI's impact on learners' higher-order thinking skills to help deepen instructors' understanding of effective techniques specified for educational goals.

Another approach is to involve instructors in AI system design. Currently, most AI applications/mechanisms remain as proposals, i.e., they are still hypothetical without evidence for their effectiveness in the real world. Hence, real-life decision-support tool development should be promoted to see whether AI-oriented applications can adapt to realistic educational scenarios and be used as pedagogical instruments (Ijaz et al., 2017). However, developing such intelligent systems is complex when learning objectives are considered. Therefore, different types of design and prototyping approaches are desired to allow both data scientists and non-technical stakeholders such as educational experts to be meaningfully involved in system development (Holstein et al., 2018a). Engineers and data scientists are primarily concerned with AI system accuracy in predicting results and less about pedagogical practice. The development of efficient systems specified for particular learning objectives requires connecting closely with pedagogical innovations and carefully considering students' learning styles. Thus, researchers should actively collaborate with subject matter experts or professional educators to build educationally sound AI systems (Chen et al., 2020a). Involving subject matter experts is essential in the AI-building process to steer data scientists in the right direction (Burgess, 2017) to ensure that new models work properly and are applied correctly to whatever dataset is of interest.

Additionally, sufficient technical support is needed to assist instructors in understanding and using AI systems. Instructors are usually challenged by personalized ITSs as they are tasked with monitoring divergent activities simultaneously (Holstein et al., 2018b). Thus, there is a need to examine different types of real-time support offered by AI applications across instructors with varied experience levels. Specifically, researchers should explore how human and automated instruction can most effectively be combined to best support instruction. Such systems have been built on teachers' prior instruction to shape pathways for current instruction and provide guidance on future instruction. These personalized and adaptive AI systems suited to a variety of pedagogical needs are more accessible to instructors (Holstein et al., 2018a), leading to a greater level of personalization across education as a whole by helping instructors design the most effective classroom experience and drive digital transformation.

4.9.3. Shifting from ML to DL

Currently, prevalent techniques in AIEd involve EDM, NLP, discourse analysis, educational robots, ITSs, affective computing, recommender systems, and neural networks, while advanced DL algorithms are less adopted. Considering DL's advantages over traditional ML algorithms in various tasks such as prediction and classification, future studies may show how DL algorithms can replace the ML algorithms already integrated into the existing systems. This would validate DL's effectiveness for multi-task prediction in EDM (e.g., student dropout and use of hints) and reduce implementation time since many required modules already exist (Krouska et al., 2019).

Attention should be paid to DL's generalization ability for adapting or applying it to various new and unexpected tasks. Gray and Perkins (2019) highlight a shortcoming of current ML models' effectiveness for learner outcome prediction because, in many cases, different patterns are often detected for different learner cohorts progressing through courses. Thus, although current models generalize well to test sets, they may not work well for new cases due to implicit memorization of certain examples, leading to ongoing AI model training by constantly including new data and eliminating aging ones. Such processes are repetitive, tedious, and inefficient due to the challenge regarding whether and what attributes and variables within a new dataset should be exploited to improve model performance (Livieris et al., 2019). The following are directions to consider while developing DL-based generalized applications.

There are always new attributes potential to impact AI models' effectiveness that are either currently unavailable but can be collected by instructors or are hidden within students' learning interactions with educational systems (Livieris et al., 2019). There are also features that need constant adjustments, an example of which is the number of days absent indicating potential school-leavers. Thus, future work on automatic feature selection and adjustment is required to facilitate DL model training.

In feature design, we suggest integrating features available in the literature and variables obtained from various channels (e.g., learners' eye-tracking data and electrodermal activities) into modeling to enhance a models' predictive performance (Olive et al., 2019) via feature selection to identify valuable features to predict interested variables. The feature selection can be optimized by considering pedagogical practices and task independence. An example is a pedagogically and theoretically sound feature design assisted by a better understanding of manual grading criteria when developing AI systems for an automatic non-native learner essay assessment. Additionally, it is essential to develop an in-depth understanding of an input feature's relationships and roles to enhance its visibility on learning processes through straightforward visualization and statistical analysis (e.g., structural equation modeling to mediate affective factors' effects).

To initialize the model, most studies train separate classifiers for individuals, which is computationally expensive depending on the dataset, and it also burdens the system. General classifiers trained beforehand and capable of classifying an individual's learning states are needed. Alternative methods include: (1) initializing models with random weights for architecture evaluation with accurate non-linearities and pooling, and (2) exploiting hypernetworks for initialization by inputting learner model architecture and generating model weights. The latter strategy also reduces the learners' burden on model training and promotes the learners' perceived ease of use without requiring them to report learning states for classifier training.

Additionally, overfitting should be avoided and over-sampling impact reduced by testing a model's effectiveness in various scenarios, including: (1) experiments on large sample sizes, (2) applying it in different contexts (e.g., blended learning), courses, and institutions (e.g., middle school and college students), and (3) considering learners' demographic characteristics (e.g., gender, culture, and high/low performance) to validate a models' general effectiveness.

5. Conclusion

This first-in-depth topic-based bibliometric study tracks current advances in AIEd research in the first two decades of the 21st century, which is needed as AIEd is receiving increasing attention. Methodologically, bibliometric indicators such as the H-index and publication count measuring scientific impact and productivity help identify active sources, countries/regions, and institutions in AIEd research. This enables scholars to be more aware of channels to make contributions and important actors to learn from (Chen et al., 2020b). Social network analysis, through scientific collaboration visualization, also reveals an invisible collaborative network of participating countries/regions and institutions in AIEd research, intuitively helping to show collaborative relationships and potential scientific collaborators (Chen et al., 2018). Additionally, topic modeling, capable of mining themes from large-scale textual data, helps understand the past and present AIEd scientific structure (Roberts et al., 2014). The identified topics and themes were further analyzed using the MK test to reveal topic dynamics to indicate how research foci change and develop, providing insights into AIEd's future directions (Chen et al., 2020a). Increasingly diverse AI technologies are being incorporated into various applications (e.g., ITSs, robots, mobile devices, and digital games) to facilitate teaching and learning. Analytical techniques such as ML, EDM, NLP, ANNs, and affective computing are commonly adopted for analyzing large-scale data from various educational scenarios (e.g., computer science education, language education, STEM education, special education, virtual surgery training, CSCL, and flipped classrooms). Eight promising areas within AIEd include (1) ITSs for special education; (2) NLP for language education; (3) educational robots for AI education; (4) EDM for performance prediction; (5) discourse analysis in CSCL; (6) neural networks for teaching evaluation; (7) affective computing for learner emotion detection; and (8) recommender systems for personalized learning. Finally, we also highlight the need to: (1) be transparent about learner data usage to realize personalized learning, (2) enhance instructors' AI acceptance by involving them in system design and convincing them of AI's effectiveness through robust experimental design, and (3) move towards "DLEd" for educational system design with higher generalizability.

Acknowledgement

An abstract entitled "A Bibliometric Study on Artificial Intelligence in Education for Two Decades" from this paper was presented at the International Conference on Education and Artificial Intelligence 2020, The Education University of Hong Kong, 9-11 November 2020, Hong Kong. Haoran Xie's work in this research is supported by the Direct Grant (DR21A5) and the Faculty Research Fund (DB21A4) at Lingnan University, Hong Kong. Gary Cheng's work is supported by General Research Fund (No. 18601118) of Research Grants Council of Hong Kong SAR, China, and One-off Special Fund from Central and Faculty Fund in Support of Research from 2019/20 to 2021/22 (MIT02/19-20) of The Education University of Hong Kong.

References

Ai, F., Chen, Y., Guo, Y., Zhao, Y., Wang, Z., & Fu, G. (2019). Concept-aware deep knowledge tracing and exercise recommendation in an online learning system. *International Conference on Educational Data Mining (EDM)* (pp. 240–245). Montreal, Canada: International Educational Data Mining Society. Retrieved from https://files.eric.ed.gov/fulltext/ED599194.pdf

Albacete, P., Jordan, P., Katz, S., Chounta, I.-A., & McLaren, B. M. (2019). The Impact of student model updates on contingent scaffolding in a natural-language tutoring system. In *International Conference on Artificial Intelligence in Education* (pp. 37–47). https://doi.org/10.1007/978-3-030-23204-7_4

Aulck, L., Nambi, D., Velagapudi, N., Blumenstock, J., & West, J. (2019). Mining university registrar records to predict first-year undergraduate attrition. *International Conference on Educational Data Mining (EDM)* (pp. 9–18). International Educational Data Mining Society. https://files.eric.ed.gov/fulltext/ED599235.pdf

Bayne, S. (2015). Teacherbot: Interventions in automated teaching. *Teaching in Higher Education*, *20*(4), 455–467. https://doi.org/10.1080/13562517.2015.1020783

Beaulac, C., & Rosenthal, J. S. (2019). Predicting university Students' academic success and major using random forests. *Research in Higher Education*, *60*(7), 1048–1064. https://doi.org/10.1007/s11162-019-09546-y

Burgess, A. (2017). *The Executive guide to artificial intelligence: How to identify and implement applications for AI in your organization*. Springer. https://doi.org/10.1007/978-3-319-63820-1

Chan, K. S., & Zary, N. (2019). Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR Medical Education*, *5*(1), e13930. https://doi.org/10.2196/13930

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A Narrative overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, X, Xie, H., Wang, F. L., Liu, Z., Xu, J., & Hao, T. (2018). A Bibliometric analysis of natural language processing in medical research. *BMC Medical Informatics and Decision Making*, *18*, 1–14. https://doi.org/10.1186/s12911-018-0594-x

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020a). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A Retrospective of all volumes of Computer & Education. *Computers & Education*, *151*, 103855. https://doi.org/10.1016/j.compedu.2020.103855

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020b). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chen, X., Xie, H., & Hwang, G. J. (2020c). A Multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, *1*, 100005. https://doi.org/10.1016/j.caeai.2020.100005

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning: Topic modeling and knowledge mapping. *Educational Technology & Society*, *24*(1), 205–222. https://www.jstor.org/stable/10.2307/26977868

Dowell, N., Lin, Y., Godfrey, A., & Brooks, C. (2019). Promoting inclusivity through time-dynamic discourse analysis in digitally-mediated collaborative learning. In *International Conference on Artificial Intelligence in Education* (pp. 207–219). https://doi.org/10.1007/978-3-030-23204-7_18

Eom, S. (2012). *Automatic presentation of sense-specific lexical information in an intelligent learning system* (Order No. 3524085) [Doctoral dissertation, Georgetown University]. ProQuest Dissertations and Theses Global.

Goksel, N., & Bozkurt, A. (2019). Artificial Intelligence in Education: Current insights and future perspectives. In *Handbook of Research on Learning in the Age of Transhumanism* (pp. 224–236). IGI Global. https://doi.org/10.4018/978-1-5225-8431-5.ch014

Gray, C. C., & Perkins, D. (2019). Utilizing early engagement and machine learning to predict student outcomes. *Computers & Education*, *131*, 22–32. https://doi.org/10.1016/j.compedu.2018.12.006

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A Twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, *4*(4), 134–147. https://doi.org/10.1016/j.ijis.2020.09.001

Hamouda, A. (2013). A Study of dictionary use by Saudi EFL students at Qassim University. *Study in English Language Teaching*, *1*(1), 228–257. https://doi.org/10.22158/selt.v1n1p227

Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial intelligence in higher education: A Bibliometric study on its impact in the scientific literature. *Education Sciences*, *9*(1), 51. https://doi.org/10.3390/educsci9010051

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018b). The Classroom as a dashboard: Co-designing wearable cognitive augmentation for K-12 teachers. *The 8th international conference on learning Analytics and knowledge* (pp. 79–88). https://doi.org/10.1145/3170358.3170377

Holstein, K., McLaren, B. M., & Aleven, V. (2017). SPACLE: Investigating learning across virtual and physical spaces using spatial replays. *The Seventh International Learning Analytics & Knowledge Conference* (pp. 358–367). https://doi.org/10.1145/3027385.3027450

Holstein, K., McLaren, B. M., & Aleven, V. (2018a). Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. *International Conference on Artificial Intelligence in Education* (pp. 154–168). https://doi.org/10.1007/978-3-319-93843-1_12

Hongmei, L. (2013). Application research of BP neural network in English teaching evaluation. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, *11*(8), 4602–4608. http://dx.doi.org/10.11591/telkomnika.v11i8.3085

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020a). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Hwang, G. J., Sung, H. Y., Chang, S. C., & Huang, X. C. (2020b). A Fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Computers and Education: Artificial Intelligence*, *1*, 100003. https://doi.org/10.1016/j.caeai.2020.100003

Ijaz, K., Bogdanovych, A., & Trescak, T. (2017). Virtual worlds vs books and videos in history education. *Interactive Learning Environments*, *25*(7), 904–929. https://doi.org/10.1080/10494820.2016.1225099

Krouska, A., Troussas, C., & Virvou, M. (2019). SN-Learning: An Exploratory study beyond e-learning and evaluation of its applications using EV-SNL framework. *Journal of Computer Assisted Learning*, *35*(2), 168–177. https://doi.org/10.1111/jcal.12330

Kulich, M., Chudoba, J., Kosnar, K., Krajnik, T., Faigl, J., & Preucil, L. (2012). SyRoTek—Distance teaching of mobile robotics. *IEEE Transactions on Education*, *56*(1), 18–23. https://doi.org/10.1109/TE.2012.2224867

Lee, K., Kwon, O.-W., Kim, Y.-K., & Lee, Y. (2015). A Hybrid approach for correcting grammatical errors. In F. Helm, L. Bradley, M. Guarda, & S. Thouesny (Eds), *Critical CALL–Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 362–367). Dublin, Ireland: Research-publishing.net. http://dx.doi.org/10.14705/rpnet.2015.000359

Lin, F., & Chan, C. K. K. (2018). Examining the role of computer-supported knowledge-building discourse in epistemic and conceptual understanding. *Journal of Computer Assisted Learning*, *34*(5), 567–579. https://doi.org/10.1111/jcal.12261

Lin, C.-C., Liu, G.-Z., & Wang, T.-I. (2017). Development and usability test of an e-learning tool for engineering graduates to develop academic writing in English: A Case study. *Educational Technology & Society*, *20*(4), 148–161.

Liu, S., Chen, Y., Huang, H., Xiao, L., & Hei, X. (2018). Towards smart educational recommendations with reinforcement learning in classroom. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (pp. 1079–1084). https://doi.org/10.1109/TALE.2018.8615217

Livieris, I. E., Drakopoulou, K., Tampakas, V. T., Mikropoulos, T. A., & Pintelas, P. (2019). Predicting secondary school students' performance utilizing a semi-supervised learning approach. *Journal of Educational Computing Research*, *57*(2), 448–470. https://doi.org/10.1177/0735633117752614

Ma, X., & Ye, L. (2018). Career goal-based e-learning recommendation using enhanced collaborative filtering and prefixspan. *International Journal of Mobile and Blended Learning (IJMBL)*, *10*(3), 23–37. https://doi.org/10.4018/IJMBL.2018070103

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, *13*(3), 245–259. https://doi.org/10.2307/1907187

Martínez-Tenor, Á., Cruz-Martín, A., & Fernández-Madrigal, J.-A. (2019). Teaching machine learning in robotics interactively: the case of reinforcement learning with Lego® Mindstorms. *Interactive Learning Environments*, *27*(3), 293–306. https://doi.org/10.1080/10494820.2018.1525411

Mehmood, R. M., & Lee, H. J. (2017). Towards building a computer aided education system for special students using wearable sensor technologies. *Sensors*, *17*(2), 317. https://doi.org/10.3390/s17020317

Mondragon, A. L., Nkambou, R., & Poirier, P. (2015). Towards an Integrated Specialized Learning Application (ISLA) to support high functioning ASD children in mathematics learning. In *European Conference on Technology Enhanced Learning* (pp. 225–239). https://doi.org/10.1007/978-3-319-24258-3_17

Mondragon, A. L., Nkambou, R., & Poirier, P. (2016). Towards an effective affective tutoring agent in specialized education. In *International Conference on Intelligent Tutoring Systems* (pp. 402–408). https://doi.org/10.1007/978-3-319-39583-8_48

Muljono, M., Afini, U., Supriyanto, C., & Nugroho, R. A. (2017). The Development of Indonesian POS tagging system for computer-aided independent language learning. *International Journal of Emerging Technologies in Learning (IJET)*, *12*(11), 138–150. https://doi.org/10.3991/ijet.v12i11.7383

Olive, D. M., Huynh, D. Q., Reynolds, M., Dougiamas, M., & Wiese, D. (2019). A Quest for a one-size-fits-all neural network: Early prediction of students at risk in online courses. *IEEE Transactions on Learning Technologies*, *12*(2), 171–183. https://doi.org/10.1109/TLT.2019.2911068

Panaite, M., Dascalu, M., Dessus, P., Bianco, M., & Trausan-Matu, S. (2018). Identifying reading strategies employed by learners within their oral French self-explanations. *The 5th International Workshop on "Semantic and Collaborative Technologies for the Web", joint to the 14th International Scientific Confrerence on eLearning and Software for Education (eLSE 2018)* (pp. 362–368). Advanced Distributed Learning Association. https://doi.org/10.12753/2066-026X-18-120

Rashid, T. A., & Ahmad, H. A. (2016). Lecturer performance system using neural network with Particle Swarm Optimization. *Computer Applications in Engineering Education*, *24*(4), 629–638. https://doi.org/10.1002/cae.21737

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(2), 1-40. https://doi.org/10.18637/jss.v091.i02

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Rosa, K. Dela, & Eskenazi, M. (2011). Impact of word sense disambiguation on ordering dictionary definitions in vocabulary learning tutors. *The Twenty-Fourth International Florida Artificial Intelligence Research Society Conference* (pp. 507–512). Palm Beach, FL: AAAI Press. https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS11/paper/view/2545/3096

Samoilescu, R.-F., Dascalu, M., Sirbu, M.-D., Trausan-Matu, S., & Crossley, S. A. (2019). Modeling collaboration in online conversations using time series analysis and dialogism. *International Conference on Artificial Intelligence in Education* (pp. 458–468). https://doi.org/10.1007/978-3-030-23204-7_38

Sato, T., Ogura, M., Aota, S., & Burden, T. (2018). Examining the impact of an automated translation chatbot on online collaborative dialog for incidental L2 learning. In P. Taalas, J. Jalkanen, L. Bradley & S. Thouesny (Eds.), *Future-Proof CALL: Language Learning as Exploration and Encounters–Short Papers from EUROCALL 2018* (pp. 284–289). http://doi.org/10.14705/rpnet.2018.26.851

Shibata, C., Ando, K., & Inaba, T. (2017) Towards automatic coding of collaborative learning data with deep learning technology. *eLmL 2017: The Ninth International Conference on Mobile, Hybrid, and On-line Learning* (pp. 65–71). Nice, France: IARIA. https://laweb.cloud.teu.ac.jp/cms/wp-content/uploads/2017/08/contribution_50092_eLmL_201703032017ando2.pdf

Su, S.-H., Lin, H.-C. K., Wang, C.-H., & Huang, Z.-C. (2016). Multi-modal affective computing technology design the interaction between computers and human of intelligent tutoring systems. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, *6*(1), 13–28. http://doi.org/10.4018/IJOPCD.2016010102

Sun, Q., & Winoto, P. (2019). An Intelligent LEGO tutoring system for children with autism spectrum disorder. *The 2019 the International Conference on Pattern Recognition and Artificial Intelligence* (pp. 56–59). https://doi.org/10.1145/3357777.3357795

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2021). Trends in artificial intelligence-supported e-learning: A Systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2021.1875001

Tsai, S. C., Chen, C. H., Shiao, Y. T., Ciou, J. S., & Wu, T. N. (2020). Precision education with statistical learning and deep learning: a case study in Taiwan. *International Journal of Educational Technology in Higher Education*, *17*(1), 1–13. https://doi.org/10.1186/s41239-020-00186-2

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

Zhang, R., & Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 1–47. https://doi.org/10.1080/09588221.2020.1744666