# Best Practices for Implementing AI in STEM Education: A Systematic Literature Review

**Taha Mansor Khawaji**
tmkhawaji@uqu.edu.sa
Umm Al-Qura University (UQU), Makkah, Saudi Arabia

**Abstract**
Artificial intelligence (AI) describes a variety of approaches in computer applications to mimic human learning. As this technology becomes increasingly prevalent, it is inevitable that it will enter the educational environment, as both an educational tool and topic of learning. STEM education, which deals with science, technology, engineering, and math, is perhaps the most appropriate educational field in which to introduce students to this new and rapidly growing technology. In recent years, educators, AI engineers, and educational researchers have published trial results of experimental curricula implementing AI technology in student and teacher education. This systematic literature review analyzed a sample of seven such publications to identify key trends in suggested best practices for the usage of AI in STEM classrooms. The sample was analyzed for keywords using MaxQDA. The results indicated three key trends among suggested best practices. The first was that AI is best taught to students when the technology itself is the topic of education. Another trend was that simulating real world applications of AI technology was most impactful in showing students the potential, limits, and ethical implications of AI. Finally, it was found that educator's familiarity with AI is an important factor in their ability to employ it in the classroom.

***Keywords:***
*Artificial intelligence (AI), Technology, STEM, Education, Systematic review.*

## 1. Introduction

Artificial intelligence (AI) represents a major advancement in computing technology that stands to impact a broad range of activities and institutions. The field of education is rapidly becoming an example of this. As students increasingly turn to AI-powered search engines and applications like ChatGPT for assistance with schoolwork, teachers are beginning to recognize that they must begin accounting for AI in their curriculums (Holmes & Tuomi, 2022). The integration of AI into educational curriculums is arguably most relevant to the disciplines of science, technology, engineering, and mathematics, a subset of educational disciplines often referred to by the acronym, STEM (Martin-Paez et al., 2019). Students aspiring for careers in these disciplines must be prepared for a future in which AI-powered technology is an integral part of their daily lives. Guiding students through their first hands-on experiences with this technology may be the best way that educators can help their students prepare for the promises, limitations, and ethical issues of AI technology in STEM (Chng et al., 2023).

This systematic literature review examines recent research into the use of AI technology in educating students in STEM with the aim of identifying themes and trends regarding suggested best practices for the use of AI in teaching STEM. A sample of publications describing experimental pilots of curriculums employing AI technology was assembled and reviewed. Many of these curriculums were focused on AI literacy; however, some of them employ AI technology in teaching creative problem solving and other aspects of scientific inquiry. Keyword analysis was employed to thematically analyze the sample of publications. This yielded a sample of 11 keywords that, when contextually examined, offered valuable insights into what educators are finding to be most effective in bringing this new technology into their classrooms. From the contexts surrounding these 11 keywords, three major themes surrounding the application of AI in STEM classrooms were derived. These themes informed an information synthesis concerning how quickly best practices were emerging for the use of AI technology in STEM education.

### *AI Technology*

AI technology can refer to a wide variety of related technologies that promote the appearance of an intelligent system. All current AI technology falls under the category of narrow AI, which describes intelligent systems designed for a single, specific task (Page et al., 2018). The common technologies underlying narrow AI applications consist of machine learning, neural networks, deep learning, natural language processing, and computer vision. Machine learning is a broad term that covers any technology that uses algorithms and specially formatted data sets to mimic the pattern of human learning (Jordan & Mitchell, 2015). Deep

Manuscript received August 5, 2024
Manuscript revised August 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.8.1

learning is similar to machine learning but represents a more advanced approach that does not require labeled and formatted data sets to inform its learning algorithm. Image generation programs, which are informed by raw images alone, are an example of deep learning-powered AI (LeCun et al., 2015). Neural networks are a type of learning algorithm that can be applied to both machine learning and deep learning. It is a node-based structure for data analysis that mimics the structure and function of a human brain (Jordan & Mitchell, 2015). Computer vision is an application of deep learning that enables technology to actively process real-time visual information through a camera. This camera technology is an important part of developing innovations like autonomous vehicles (LeCun et al., 2015). Natural language processing describes a machine learning application in which AI can comprehend and mimic human language, often in real time. This technology is accessible through programs like chat bots and ChatGPT (Qin et al., 2023).

As these descriptions show, the growing technologies facilitating the AI revolution are interconnected with one another and present a complex web of structures. Teaching these complicated technological concepts can present a challenge for educators, especially those who are not already familiar with AI (Lin et al., 2021). The publications reviewed in this study utilize these technologies in a wide variety of combinations and applications. Selection of which AI technologies to use often reflects the AI experience and overall intentions of the curriculum designer. Suggestions for best practices in each publication are further reflective of which AI technologies are believed to have the greatest educational value.

***STEM Education***

Science, technology, engineering, and mathematics are the four disciplines that make up the STEM acronym. Education in these disciplines has been deemed valuable in modern society for its ability to perform tasks that are often time-consuming and difficult or even impossible for humans. These include analyzing large data sets, making rapid informed decisions in complex scenarios, and automating machinery operations. To prepare their populations to effectively utilize these capabilities, many nations have instituted country-wide campaigns to build their education system’s capacity for successfully teaching STEM topics (Gonzalez & Kenzi, 2012). STEM education typically relies on hands-on activities through which students learn how to independently conduct an inquiry, creatively solve problems, collaborate in a team towards a shared aim, and understand the value of evidence (Wahono et al., 2020). Even for students not aspiring towards STEM careers, the skills learned in STEM are considerably highly valuable and widely transferable (Gonzalez & Kenzi, 2012). Skills related to STEM are also becoming increasingly necessary for a wide variety of occupations, such as nursing, manufacturing, and industrial operations, that did not previously require them. Estimates suggest that over 80% of jobs becoming available in the near future will require at least a moderate level of STEM skills (Van Tuijl & Van der Molen, 2016).

### 1.1 Aims

The overarching aim of this research was to identify what best practice suggestions were emerging in the research literature regarding the usage of AI technology in STEM education. Facilitating this aim were the methodological aims of the research, which began with identifying a representative sample of rigorous studies testing the success of trial curricula implementing AI. Analysis of this sample was guided by the aim of identifying trending keywords shared across publications in sections regarding suggestions and best practices. The final aim was to synthesize the results of this analysis to present trends in best practices for AI implementation in STEM education and discuss what their potential implications were for the future of this new technology.

### 1.2 Research Questions

The primary research question was what best practices were being suggested based on the results of experimental STEM curricula implementing AI technology. Nesting questions related to this overarching question attempted to pre-empt different aspects of implementing AI for this purpose. The first was, what topics in STEM is AI technology being used to explore with students? Following this was what uses of AI technology in STEM education have been found to be most valuable by educators? Finally, are specialized educators necessary to effectively use AI with students in STEM education? It was expected that suggestions for best practices implementing AI in STEM education would answer these questions either directly or indirectly.

## 2. Methods

This systematic literature review was performed according to the approach developed by Torres-Carrion et al. (2018), which is further based off the work of Kitchenam et al. (2009) and Bacca et al. (2014). Their method for literature reviews considers the process of problem conceptualization to be the first stage in the design of literature collection. The underlying philosophy for this choice is that the justification for conducting a novel literature review may come down to nuanced details of an inquiry that are missed by previous relevant reviews of research. Torres-Carrion et al. (2018) recommend an extensive process of refining the central research question that concludes with a thorough examination of existing literature reviews. If a novel literature review is warranted, the following stage of development is developing the review protocol.

Several literature reviews focused on the application of AI technology in STEM education exist; however, these review approach AI from a novel technology perspective. Xu and Ouyang (2022) reviewed literature on AI in STEM education through the lens of what challenges this new technology may present to educators. Casal-Otero et al. (2023) focused their literature review on shortcomings in existing studies regarding AI applications in STEM education and drew conclusions regarding how future studies may improve on this. Chng et al. (2023) took a similar approach, reviewing existing studies with the intention of suggesting avenues of future research for AI in STEM education. Luzano (2024) studied applications of AI technology through the lens of equity and inclusion with the intent of discerning whether this new technology is helping or exacerbating the problem. The present systematic literature review shares the interest of these previous studies in how AI technology may enhance or hinder STEM education. A unique aspect of this study that necessitates a renewed literature review effort is its focus on best practices. The primary focus in the analysis and conclusions of this review is on what approaches to AI technology STEM educators have found most useful in accomplishing their classroom goals.

## 2.1 Search and Selection Method

Several well-known academic search engines were engaged to gather literature for this systematic review. These consisted of Google Scholar, JSTOR, ResearchGate, and ArXiv. A consistent series of research terms, presented in varying combinations via Boolean operators, were entered into each search engine. Because AI is often referred to by both its initialism and the term "artificial intelligence", it was consistently entered into search engines as ("AI" OR "artificial intelligence"). Searches were also alternatively performed in combination with terms related to specific types of AI, notably machine learning and expert systems, entered as ("machine learning" OR "expert system"). To capture a broad variety of educational environments, it was necessary to try a substantial variety of terms related to STEM education. Search terms were grouped by those related to STEM and those related specifically to education, connected by the Boolean operator AND in the middle. The search entry resulting from this strategy was "("education" OR "pedagogy" OR "teaching") AND ("science" OR "technology" OR "engineering" OR "mathematics" OR "math"). This search entry was alternatively combined with the two different AI search entries above.

## 2.2 Inclusion and Exclusion Criteria

The primary criterion for inclusion in this systematic literature review was that the published study be on the application of one or more AI-based technologies in an educational classroom setting focused on science, technology, engineering, or mathematics. There was no target age group or educational level for this literature review and the educational environment could be either formal or informal. Publications in English were exclusively considered for this review as this is a common language for academic publications and most international studies are translated, at least in part, to English. Only studies published in peer-reviewed journals were considered for this review as this ensured the study was subjected to the scrutiny of other experts in education research prior to publication.

Setting a temporal range for this review was difficult as AI has a long history of application in the educational field, with the earliest applications stretching back to the 1960's with the PLATO computer system (Cope & Kalantzis, 2022). These original applications are substantially different from modern applications of AI in which this technology has become widely available on a diverse array of platforms, including smart devices. To preserve a focus on the contemporary concept of AI, this review limited itself to studies published after the "Internet of Things" (IoT) paradigm was established. The IoT paradigm emerged around 2015 and describes the widespread presence of internet-capable technology beyond traditional computers. This paradigm has been instrumental in the development of modern AI applications and is considered a significant moment in the history of the current technological culture (Ghosh et al., 2018). Based on these considerations, the temporal range for the present literature review extended from 2016 to the present year of 2024.

After the initial literature search, all articles were vetted for exclusion. The first round of vetting examined publications at the title level. If the title did not indicate that the publication was primarily focused on a real-world application of AI technology to STEM education, it was excluded from the final selection. The second round of vetting considered the abstracts of the remaining publications. Abstracts indicating that the publication was only descriptive, and no quantitative or qualitative method was applied in the use of AI for STEM education were excluded. This decision was made to ensure this review limited itself to evidence-based research. Finally, a third round of vetting considered each remaining publication in its entirety. Publications were only excluded at this stage if there was no suggestion of best practices for AI in STEM education contained at some point in its discussion or conclusion.

### 2.3 Analysis

The framework provided by Torres-Carrion et al. (2018) gives only generic advice for data extraction and analysis, allowing their framework to be pared with a variety of research methods. This review adopted a qualitative approach known as “keywords-in-context” to synthesize the main arguments of the publications sampled (Onwuegbuzie et al., 2012). The qualitative paradigm was selected for this review because it offers a more in-depth approach to interpreting textual information than the more statistically based quantitative paradigm (Gelo et al., 2008). Keywords-in-context utilizes reoccurring key words and the context that surrounds them between publications to assess the similarities and differences of the sample (Onwuegbuzie et al., 2012).

The first step of this analysis method is to review each publication in depth and record key terms related to the description of the study. Once all publications have been reviewed, the resulting key terms can be compared and grouped by theme. The publications are then reviewed again for the contexts surrounding terms related to the reoccurring themes. An analytic process of comparing similarities and differences between these contexts then yields material for the researcher’s conclusion (Russell et al., 2016). For this research, keywords-in-context analysis was performed with the assistance of MaxQDA. MaxQDA is a specialized software suite uniquely designed for qualitative research that has been consistently found to perform well in keyword identification and analysis (Oliveira et al., 2013). In this research, MaxQDA presented an initial selection of potential keywords that the researcher then vetted based on their familiarity with the literature sample. A subset of these potential keywords was selected before MaxQDA was directed to analyze the literature sample for context surrounding keywords. Keywords and their context were presented via an annotated spreadsheet created by the software. The researcher was then responsible for the final analysis of the keywords and contexts presented by MaxQDA.

## 3. Literature Review

The initial search results yield 402,000 publications before vetting. Many of these initial publications were research and literature reviews that only covered prospective applications of AI in STEM classrooms. Filtering publication titles in the initial search results for the word “review” removed 201,248 publications from the prospective literature sample, leaving 200,752 results to vet. This was paired down to a more manageable number of 2,436 publications by filtering titles for the terms, “promise”, “trends”, “direction”, and “future”, all of which were found to be common among prospective reviews of potential future applications of AI in STEM education. A reading review of the remaining titles left a prospective sample of 947 publications to review at the abstract level. 812 of the publications reviewed at the abstract level described and synthesized other research, described tools for improving AI, or described theoretical practices that educators may use in the future. These were removed from the search results as they did not offer actionable suggestions for best practices based on actual case studies. This left 135 publications to review in depth for potential conclusion. Of these, only seven met the criteria of being a recent publication describing the actual application of AI in a STEM education setting that resulted in actionable results for best practices.

This section contains a brief synopsis for each study before discussing the results of the thematic analysis. The qualifications of the authors for each publication, their methods, study samples, outcomes and suggestions are all described in the synopses. This section is intended to give an idea of who is pursuing research into the actual application of AI in STEM education and what approaches they are using. Though these factors are not part of the analysis of these publications, they do provide useful context for the results of this study and potential suggestions for future research.

**Lin et al. (2021)**

This study was performed by several professors at Chung Yuan Christian University in Taiwan, all of whom work in teacher education at either the graduate or undergraduate level. Their intent was to design a training method for students without software engineering experience to build a skillset relevant to the application of AI in STEM education. To do so, they designed a three-week educational program with two main activities for the application and assessment of learned material. The authors’ desired outcome was that students preparing to become general educators would demonstrate a thorough understanding of how AI works and what role it may play in future academic settings. This outcome was deemed necessary based on two factors. First, the authors note that STEM education plays a crucial role in general education, especially in Asian countries. Second, because most practicing general educators did not themselves learn in an environment where AI technology was present, most have a very limited concept of what AI is and how it may impact their classroom. The guiding questions of the research were whether an “AI literacy” course would improve educator knowledge of how to use AI and what the related ethical issues are.

Though Lin et al. (2021) were interested in education students, they assembled their research sample of 328 participants from departments throughout their university, with most coming from accounting, business management, and information technology programs. The sample was also skewed female, with only 33% of the participants identifying as male. Each participant attended one week of lectures, followed by one week focused on the first AI project and another week focused on the second AI project. The students' first hands-on task consisted of training an AI art creation program by assembling a training data set of first five and then twenty images. This was followed by a second hands-on task in which students were tasked with training a remote-controlled vehicle that could function autonomously if properly trained to understand environmental variables.

The impact of the lectures and hands-on projects were measured using a previously vetted instrument, known as the "AI Literacy Scale", that was also developed by the authors. AI literacy tests were administered before and after the AI intervention and the two scores were compared via a paired samples T-test. Lin et al. (2021) found a significant increase in AI literacy scale scores after the AI lecture and training activities were completed, enough to conclude that their approach to AI training was successful. The authors conclude with several suggestions for best practices after their discussion, noting what worked best in their design and what could have been done better. Many of these suggestions focused on better assessing what level of understanding each participant had of AI prior to the training. Some of the participants already held a high understanding of AI and did not find the activities to be challenging or informative, while others had no previous understanding and found the activities more challenging than intended. Lin et al. (2021) recommend having a diversity of activities to potentially account for this variability in skill level. The authors also made the suggestion that a good practice would be to pair more material on AI ethics with the training, as many participants left the AI training with only a limited concept of its potential dangers if used unethically.

**Lee and Perrett (2022)**

Lee and Perrett (2022) describe a five-day educational program for secondary teachers in a variety of STEM disciplines focused on AI and AI ethics. This program and study were authored by two educational consultants working for the Education Development Center (EDC), an international nonprofit agency focused on educator and student development. Their AI training program built on existing foundational science programs, known as Science + C, that EDC currently offers in American high schools designed to implement computer science as part of the standardized STEM curriculum. Lee and Perrett (2022) worked with an advisory team of computer science undergraduate students at MIT in designing the AI program, which will eventually be oriented towards high school students (age range of 13 – 18 years old). The design team realized early in the process of creating the program that low levels of familiarity with AI among high school educators would present a significant barrier to any AI-focused program's success.

The five-day program offered by Lee and Perrett (2022) in their study was designed to address the lack of professional development opportunities focused on AI for current educators. On each day of the program, educators were introduced to a new aspect of AI, with the week's curriculum proceeding through data analytics, decision trees, machine learning, neural networks, and transfer learning. Each day consisted of a lecture followed by a series of graduated tasks and projects that introduced educators to increasingly complex aspects of the day's topic. They assembled a sample of 19 STEM educators from the Northeast and Southwest United States to participate in an experimental pilot of the program. Lee and Perret (2022) did not use a previously vetted research instrument, but instead used an approach of assessment common to educational research. At the start and end of each day research participants were given a brief quiz to assess knowledge of the research area, which yielded evidence of how much they retained from the day's lesson when compared. Participants were also asked to participate in a daily survey of their perceptions of the lessons each day. Finally, an end of program survey was administered to measure how the sample of educators felt about AI and its potential in the classroom after learning about it.

Lee and Perret (2022) report that the training program was well received by educators and the most common theme in their answers was a new-found confidence in understanding AI topics. Many educators stated that the experience helped them better envision how AI and data science could be incorporated into their curriculums. All educators showed gains in knowledge after going through the lessons; however, there was substantial variability in how much knowledge was retained between participants. Suggestions for best practices came entirely from the suggestions of the educators in the end of program survey. The most prominent suggestion was that educators believed a more integrative approach to teaching AI might be more beneficial than focusing on a different aspect of AI each day. Many educators were concerned that their students would experience difficulties connecting what they learned at the beginning of such a program with that they learned at the end. A greater focus on real life applications of AI was also recommended as such applications were more likely to give strong examples of the ethical issues that accompany AI usage.

**Karampelas (2021)**

Karampelas (2021) is an effort by a single educator at an American Community School who designed an extensive assessment protocol and education unit to examine several specific research questions. These were: what prior knowledge and perceptions of AI did students have, what impact did AI-based learning experiences have on students' perceptions of AI and knowledge of the implications and limitations of AI, and if there were any correlations between age, gender, and AI knowledge. A sample of 62 international high school students were assembled for this experimental lesson unit, with the majority being female and in the 9th grade. Pre- and post-lesson assessments were conducted for every topic in the learning unit. In total, there were three topics; however, the amount of time spent on each topic and the AI unit overall is not specified. The three topics of the unit were: societal impact of artificial intelligence, machine perception, and machine learning.

Assessments before and after the unit on societal impacts revealed that most students, male and female, had generally positive impressions of the impacts AI would have on society before the unit and further education on AI had a minimal effect on this. There was evidence that some students changed their perception from "entirely positive" to "mostly positive" after this unit. Students generally showed a high degree of knowledge retention after the unit. Results were similar for the next unit, "machine perception", which was focused on self-driving cars. Students generally reported having pre-existing knowledge of the technology discussed but showed signs of knowledge gains after the lesson unit. It is important to note that Karampelas (2021) focused heavily on autonomous vehicles in this unit, which may have limited students' understanding of the wider field of AI-based sensory mechanisms. The final topic, machine learning, was the only AI lesson unit that the majority of students reported having little prior knowledge about. Many students struggled with this final unit and knowledge retention was far more varied, with at least 27% of the sample stating that they believed machine learning was too complex of a topic for high school students.

Karampelas (2021) chose not to offer hands-on activities as part of the AI educational unit, choosing instead to focus on discussion and repeatable exercises as a means of knowledge reinforcement. This was an intentional choice as the researcher was interested in more traditional, lecture-based modes of education as a means of introducing students to AI. AI was primarily present in the classroom through the brief use of online AI applications, such as image and text generating programs. Karampelas' (2021) suggestions for best practices were geared towards this more limited engagement with the technology. To this end, the researcher pointed out that students found AI chat bot programs (programs capable of mimicking human social interaction) to be exciting at first; however, this excitement quickly dissipated as students began to recognize the limits of the technology. The researcher found this to be a valuable outcome and suggested repeat experiences with specific AI apps to be a good practice to give students insight into capabilities and limitations.

**Sakulkueakulsuk et al. (2018)**

Sakulkueaskulsuk et al. (2018) is a publication by a diverse group of college professors from the United States and Thailand who collaborate via a lab known as the Future Research Cluster Thailand. This research team designed a unique AI learning project, known as "AI Challenge", that engaged students through an agriculture-based game in which they had to train an AI application to recognize the difference between ripe and unripe mangos, assess them for quality, and send them to market for a profit. The learning unit around AI Challenge was only designed to be three days long and was intended to serve as a complete introduction to AI for students who had little or no previous knowledge of the technology. A sample of 84 Thai middle school students was assembled for an experimental pilot of the educational unit. Pre- and post-unit assessments, along with a self-guided questionnaire, were used to measure the learning outcomes of the educational program.

AI Challenge proceeded in three phases. In the first phase, the student sample was divided into teams of six. Each team had to use an AI platform, RapidMiner, to train a machine learning program to differentiate between sweet and sour mangoes. To encourage more engagement with the lesson, the designers used a game-like approach in which each correctly predicted mango was worth 25 points and teams competed for the most points. In the second phase, the student teams were introduced to more advanced machine learning concepts that they were eventually tasked with using to improve their mango assessment models. The mango assessment program was now required to grade mangos based on their level of sweetness, with correct predictions again earning points for each team. In the final phase, students were introduced to real-world applications of AI by using their machine learning programs to market their mangos at auction. Students were encouraged to strategize in this phase and a number of other, non-AI options were available to them to enhance the effectiveness of their machine learning model.

Sakulkueakulsuk et al. (2018) set a benchmark for accuracy of the machine learning programs to partially assess student knowledge retention. Though the accuracy of the machine learning programs varied substantially, only one failed to meet the benchmark. When compared, pre- and post-unit assessments gave further evidence that participants in the AI Challenge had managed to retain foundational knowledge of the technology. The most stirring evidence of the program's success came from the student self-assessment in which the majority of participants expressed that AI Challenge had been an enjoyable, informative, and motivating experience. The researchers make a point of noting that, though most students were successful in the program, it had not been easy for students to gain proficiency in AI. Many students took time to comprehend lesson materials and required at least some teacher assistance to successfully apply what they had learned in the game setting. Sakulkueakulsuk et al. (2018) conclude that scaffolding, a practice in which educators lightly guide students in projects, and gamification may both constitute best practices in getting students to successfully engage with AI technology.

**Yannier et al. (2020)**

Yannier et al. (2020) are three fellows of the Human-Computer Interaction Institute, which is hosted in the Psychology Department of Carnegie-Mellon University. This research team developed an AI-based educational tool known as the Intelligent Science Station (ISS), that assists students as they complete real-world applications of the scientific method. The aim of the ISS is to encourage successful active learning. Active learning is a philosophy in education that posits the most impactful way for students to learn a concept is by directly applying it through trial and error. The authors note that active learning is often difficult to facilitate as it requires providing students with rapid explanatory feedback on their efforts. In designing the ISS, the authors created a physical system for young students that could actively monitor their efforts and provide guidance as well as immediate feedback as they go about their project.

To test the efficacy of the ISS, a sample of 75 American school children in first or second grade was assembled to participate in a single educational experience with the ISS. Children worked in pairs to complete the educational tasks. The educational experience could take one of two directions. In both directions, children built structures out of wooden blocks on a vibrating table that was designed to mimic the conditions of an earthquake. Their goal was to create a balanced structure that could remain upright while the table vibrated. The first direction, referred to as "Guided-Discovery", gave students instructions on how to build towers on the table and use a touch-based tablet to predict which tower would fall first. An AI-powered camera would then monitor the block structures as the table shook and assess how well the students' designs survived the vibrations and how closely their predictions matched reality. An animated character on a video monitor would then give students feedback on their designs and predictions. The second direction, referred to as "Explore-Construct", did not offer students instruction on how to build stable towers and instead allowed a more open "trial and error" experience followed by feedback to help students learn what worked and did not. Yannier et al. (2020) also experimented with a third approach which combined elements of the two, providing students with differing levels of instruction depending on how quickly they learned to construct vibration-proof towers.

Pre- and post-lesson assessments were used to measure how well students learned from the educational experience. These assessments were very brief, due to the young age of the sample size, and only asked students to assess why they thought different tower designs would remain standing or fall. Post-lesson assessments also asked students how much they enjoyed the experience. The results of the experimental pilot of the ISS showed that children made much greater learning gains in all modes of AI-facilitated education than they did in traditional observational learning settings. Children also unanimously stated that they enjoyed the educational experience. An important outcome for best practices was that the "Guided-Discovery" approach to the activity had significantly higher educational value than the other two directions the educational experience could take. Students who were given instruction before attempting the activity were able to give much more robust explanations of tower stability. Yannier et al. (2020) concluded that, in terms of best practices, AI-facilitated education works best with some scaffolding or guidance by an educator. The authors noted that this result was valuable as many AI-based educational experiences in museums and similar settings used the "Explore-Construct" approach that was less impactful.

**Jang et al. (2022)**

Jang et al. (2022) are researchers at the department of engineering at Kyungpook National University in South Korea. To assess the potential for AI to impact the STEM curriculum of elementary students (ages 5 – 12), they created a semester-long educational program focused on AI. The program covered five topics over the course of the semester, each of which was assessed with a pre- and post-unit survey. These five topics consisted of the social impact of AI, AI and communication, interactions with AI, emotional exchanges with AI, and characteristics of AI. In conjunction with this lesson program, students worked in teams to apply AI technology in one of three different real-world settings. One project, known as "Mask On!", tasked students with building an AI model that could visually assess whether someone was correctly wearing a suitable mask for a COVID-19 outbreak. Another project, titled "Socially disadvantaged", tasked students with applying AI to mobility barriers faced by people living with disabilities. The final project, known as "eco conservation", tasked students with developing an AI model for tracking environmental degradation of a local forest. At the conclusion of the lessons and associated task, students wrote a free-form essay that was then qualitatively assessed to investigate knowledge retention and the overall experience of the semester-long unit.

A sample of 120 Korean elementary students, equally divided between males and females, were assembled for an experimental pilot of the education program. Each student was given a laptop with AI tools pre-installed on them for the activity component of the program. Educators teaching the unit were also vetted for experience and only those educators with ten or more years of STEM teaching experience were included in the pilot test of the program.

Analysis of the pre- and post-unit assessments indicated a high degree of knowledge retention from each section of the unit. The authors also interpreted the results to indicate that all student participants displayed an increase in creative problem-solving skills. Analysis of the free form essays revealed that the educational unit gave students an optimistic view of the impact AI would have on society. Furthermore, students unanimously enjoyed their experiences with AI during the semester.

Jang et al. (2022) attribute the success of their program to the strong connection it built between theoretical learning about AI and real-life applications of the technology. For best practices, they recommended that future efforts towards AI-powered STEM education maintain this approach, especially if students can directly experience the results of their AI work. The authors also note that including aspects of self-directed problem solving in AI educational units is a highly effective practice in encouraging students to engage with the technology. A third suggestion for best practices from this study is for future AI programs to focus on image and textual analysis applications of the technology rather than generative AI applications. Jang et al. (2022) concluded that analytical applications of AI technology have more educational value than generative applications.

**Chiu et al. (2022)**

Chiu et al. (2022) present a curriculum, known as AI4Future, that was developed by a team of 14 professors from the Chinese University of Hong Kong in collaboration with an advisory board of 17 principles and STEM educators from the Hong Kong school system. The resulting curriculum was a semester long and drew a direct analogy between the capabilities of AI and neurosensory network of a human. At the beginning of the curriculum, students learned about "core" AI concepts, such as machine learning, datasets, and the ethical and societal implications of AI. In the second phase of the curriculum, educators guide students in exploring how AI technology "perceives" it surroundings, such as using computer vision to "see", natural language capabilities to "hear", and neural networks to "think". The final stage draws a connection between the theoretical concepts learned and how these concepts are used in real-world AI applications. At the conclusion, students learn of how AI is currently estimated to reshape various aspects of the world as it continues to develop.

AI4Future also uses a nested approach to giving students hands-on experience with AI technologies. After learning of the core concepts behind the technology, students first began experimenting with basic AI applications like chat bots and image generators while learning about the "perceptive" abilities of AI. When students reach the real-world application unit of the curriculum, they are broken into teams and tasked with assembling a CUHKiCar, a robotic car built for educational applications that offers up to six AI-capable functions. A sample of 385 students in 7th-9th grade at Hong Kong secondary schools were selected to pilot this AI-based education program. Chiu, et al. (2022) chose a mixed methods approach to investigating both student and educator success. Student success was measured via pre- and post-curriculum assessments that were compared using a paired sample T-test. Both students and teachers also completed open-ended questionnaires regarding their experience and level of confidence with AI after the curriculum. Questionnaires were assessed qualitatively for more in-depth insights into the experience.

Chiu et al. (2022) had the misfortune of entering their initial pilot implantation of the curriculum during a COVID-19 outbreak in Hong Kong, thus the curriculum had to be hastily divided into a mixture of in-person and virtual classroom sessions. Despite this, reviews from teachers and students regarding how much they learned from the curriculum and how much they enjoyed it were high. Most students and teachers expressed much higher confidence in understanding how AI works and how it may impact the world in the future. The paired-samples T-tests of pre- and post-assessments also quantitatively indicated a high degree of knowledge retention at the end of the unit. Despite this success, Chiu et al. (2022) make the interesting suggestion that AI curriculums are best taught in-person rather than online. Their reasoning behind this is that learning about computers in a computer-mediated environment presents too much of a temptation for students to begin engaging with AI technology before their educator is ready to guide them. They also suggest that such curriculums may best serve students when they are designed to offer a high degree of flexibility. As the researchers explain, some students will rapidly learn the information while others will struggle. It is important for educators to have the option to slow the curriculum down or pivot to other sections if a large portion of students are struggling.

## 4. Results

***Keywords and Justification***

In assessing the literature sample, MaxQDA initially returned a selection of 27 potential keywords to consider between publications. Of these, ten were unrelated to AI and associated topics, focusing instead on terms common to classroom research such as, "educators", "students", "assessment", "pedagogy", and "curriculum". Another six keywords reflected terms unique to STEM education, such as "science", "math", and "engineering". The eleven remaining keywords were all relevant to the use of AI in STEM education and were used in the next round of analysis. These keywords consisted of: "integrative", "scaffolded", "ethics", "ethical", "data", "dataset", "training", "expert", "real-world", "hands-on", and "activity". The researcher selected these terms for inclusion in the next stage of analysis based primarily on their close reading of the publications used in the sample (see Table 1).

**Table 1**
*List of keywords and the publications they were found in*

| Keyword | Publications containing keyword |
|---|---|
| Integrative | Yannier et al. (2020); Sakulkueakulsuk et al. (2018); Lin et al. (2021); Lee and Perrett (2022) |
| Scaffolded | Jang et al. (2022); Yannier et al. (2020); Karampelas (2021) |
| Ethics | Chiu et al. (2022); Sakulkueakulsuk et al. (2018); Lee and Perrett (2022) |
| Ethical | Chiu et al. (2022); Lin et al. (2021); Karampelas (2021) |
| Data | Chiu et al. (2022); Jang et al. (2022); Sakulkueakulsuk et al. (2018) |
| Dataset | Chiu et al. (2022); Sakulkueakulsuk et al. (2018); Lee and Perrett (2022) |
| Expert | Jang et al. (2022); Lin et al. (2021); Lee and Perrett (2022); Karampelas (2021) |
| Real-world | Chiu et al. (2022); Jang et al. (2022); Lee and Perrett (2022) |
| Training | Jang et al. (2022); Yannier et al. (2020); Sakulkueaskulsuk et al. (2018); Lin et al. (2021) |
| Activity | Jang et al. (2022); Yannier et al. (2020); Lin et al. (2021); Karampelas (2021) |
| Hands-On | Chiu et al. (2022); Yannier et al. (2020); Lin et al. (2021) |

Integrative and real-world were both terms common to suggestions regarding best practices for AI training. Many of the publications concluded with suggestions to integrate multiple aspects of AI into each lesson and rely on real-world applications for examples. Scaffolded is a term frequently used in education that describes activities that are lightly guided by the educator, but largely rely on the initiative and ingenuity of students. Many of the activities designed to familiarize teachers and students with AI applications used a scaffolded approach in that they were encouraged to work independently with the guidance of an expert. Discussions of ethics and ethical issues were present in every publication reviewed. Collectively, the literature suggests this is a substantial topic of concern among those studying AI integration into classrooms, especially when it comes to the potential for plagiarism. Data and dataset are both necessary terms when discussing AI applications. AI programs rely on training datasets to familiarize themselves with a topic or practice, a fact that likely contributed to the prominence of the keyword, "training". Furthermore, AI programs are also frequently used when there is too large a volume of data to task a human with analyzing. The term training likely also emerged as a keyword due to the large number of training programs for AI education described in the literature. Expert emerged as a keyword due to frequent reference to "expert systems", an antiquated term for AI systems, or AI experts. Finally, hands-on activities were the most common type of training instrument used in the research.

### *Context*

A contextual analysis of the eleven keywords described above was run on the literature sample using MaxQDA. This analysis returned a total of 119 excerpts from the publications sampled to consider in determining what trends the eleven keywords indicated in the literature. These contexts yielded three themes when compared, which were: a focus on AI as a topic rather than just an educational tool, the importance of hands-on AI activities that simulate real world applications, and the importance of educator familiarity with AI technology. These themes, their relevant keywords, and representative quotes for each are presented below in Table 2. For some of the keywords, there was a general agreement among the contexts mined from the literature, suggesting keywords were being used with same intent; however, most keywords showed some variation in intent based on the surrounding context. In total, three major themes emerged from examining these contexts. The terms "data", "dataset", "ethics", and "ethical" revealed a common theme of AI most frequently being employed in STEM classrooms to teach students about AI technology. "Integrative", "real world", "hands-on", and "activity" yielded a theme of direct application of AI in simulated real-world situations. Applying AI in hands-on activities that integrate theory and real-world applications was a frequent suggestion in best practices. The terms, "scaffolded", "training", and "expert" yielded a final theme highlighting the importance of educator competence in AI. Many pilot AI curriculums provided teacher training first so that students would have knowledgeable guidance as they navigated the AI activities.

**Table 2**
*Themes, Keywords, and Exemplifying Quotes from the thematic coding scheme*

| Theme | Relevant Keywords | Exemplifying Context Quote |
|---|---|---|
| Focus on AI as a topic rather than AI as a device for learning | Data, dataset, ethics, ethical | I think we must collect enough Data to create a good AI model (Lin et al., 2021, p. 235) |

| | | |
|---|---|---|
| The direct application of AI in real-world scenarios | Real world, hands-on, integrative | The only thing I would change would be framing some of the concepts in a bigger picture or real world before we jumped into what something did. It usually became clearer as we progressed through a new concept but I'm a big picture learner who tends to be lost until I can see where we are going (Lee & Perrett, 2022, p. 12789) |
| The importance of educator competence in AI | Scaffolded, training, expert | We see a general consensus on the importance of active learning and of engaging science students in inquiry activities with appropriate scaffolding, but what elements of active learning and which types of scaffolding are most effective? (Yannier et al., 2020, p. 94) |

## 5. Discussion

Collectively, the three themes that emerged from contextual analysis reflect a tendency to use AI in STEM classrooms primarily to give students an opportunity to directly work with technology in a scenario that mimics real world applications. The importance of having educators sufficiently trained in AI usage is crucial to this task. This section will examine the keyword contexts that yielded these three themes before synthesizing them to draw conclusions regarding best practices in AI.

### *Focus on AI as a topic rather than AI as a device for learning*

The majority of applications of AI in STEM education reviewed in this publication were focused on the study of AI itself rather than the use of AI to study other STEM topics. This became evident in the contexts surrounding the keywords "data", "dataset", "ethics", and "ethical". "Data" and "dataset" were often accompanied by contexts that referred specifically to the training of machine learning programs. Consider this excerpt from Lin et al. (2021), which describes one of the desired learning outcomes of the proposed curriculum, "I think we must collect enough Data to create a good AI model" (p. 235). A similar context can be found in Lee and Perrett (2022) when describe the gradual progression of their curriculum: "The progression of units ensured that learners inspected and analyzed datasets in the first unit prior to using the dataset in subsequent units for the training and testing of AI models" (p. 12785). Contexts like these, in which the terms data and dataset are used to describe the training of students for using AI, make clear that AI is most often applied in teaching students about the technology itself through hands on experience.

The terms "ethics" and "ethical" were found almost unanimously in contexts discussing how students were taught about AI technology through experiences with AI. Consider the following quotae from Jang et al. (2022) connecting AI ethics with perceptions of AI: "Attitude towards AI is an ethical area of AI, which is about the positive and negative effects of using AI, and the tendency to use AI in life" (p. 15179). An excerpt from Chiu et al. (2022) explains that ethics was often a focus of AI education as it was a less technical approach to teaching the subject: "The framework mostly relied on ethics, critical thinking, and creativity, rather than requiring primary and secondary teachers to become knowledgeable in coding or robotics" (p. 37).

### *The Direct Application of AI in Simulated Real-World Scenarios*

The most prominent suggestion for best practices among the publications was that students should directly engage with AI through hands-on activities that simulate real-world applications of the technology. This theme was commonly expressed around the keywords, "integrate", "hands-on", and "real world". These contexts largely indicate that this suggestion came directly from the experience of educators as they worked through different AI curriculums with students. Consider the following statement from Lin, et al. (2021) that describes the suggestions of educators after a pilot of an AI curriculum: "Furthermore, it was pointed out that hands-on activities in STEM courses are an important element that can effectively enhance students' active learning and increase their learning effectiveness" (p. 235). Educators generally seemed to believe the value of these hands-on experiences was in fully establishing new AI concepts for students in a way that would help them retain the concept as they moved on to be introduced to others. The following quote from Lee and Perrett (2022), taken from a post-curriculum questionnaire administered to teachers, is representative of this: "The only thing I would change would be framing some of the concepts in a bigger picture or real world before we jumped into what something did. It usually became clearer as we progressed through a new concept but I'm a big picture learner who tends to be lost until I can see where we are going" (p. 12790).

The context derived from a smaller portion of the publications acknowledged that the use of hands-on, real-world applications of AI had become standard practice in the application of AI in STEM classrooms. These contexts often refer to previous research on AI in classrooms. Consider the following representative quote from

Sakulkueakulsuk et al. (2018): "In the context of AI education, researchers have proposed the strategies to integrate the learning of artificial intelligence with the social and real-world problems by implementing the idea of "AI challenges" as a way to motivate students to learn about AI" (p. 1008).

***The Importance of Educator Competence in AI***

There was a shared concern between publications regarding the ability of teachers without previous knowledge of AI-based technologies to adequately integrate this technology into their classroom and/or teach it as a topic. Some of the programs reviewed were specifically targeted at educators while others used educators as a pilot sample for their educational program with the aim of preparing these educators to teach the curriculum. Some of the contexts relevant to this theme made note of efforts to connect AI to major theoretical concepts in pedagogy, especially the practice of scaffolding. This term describes an educational activity focused on student problem solving skills that is lightly guided by an educator (Anghileri, 2006). Consider the following quote from Yannier et al. (2020), "We see a general consensus on the importance of active learning and of engaging science students in inquiry activities with appropriate scaffolding, but what elements of active learning and which types of scaffolding are most effective?" (p. 95). This quote exemplifies a desire to move beyond connecting educational theory with AI education and begin establishing which educational approaches are most effective in the teaching of this new technology.

Relevant to this theme, there was also some disagreement among publications regarding the value of directed educator training about AI-based technology. Some publications made training and related educator focus groups a central part of their pilot curriculum. This quote from Jang et al. (2022) exemplifies this as the authors assembled a focus group of the most experienced educators to make training decisions, "Teachers with more than 10 years of experience held a pre-training session and a monthly meeting to discuss the progress of the class and effective learning process to improve their competency to operate the developed program" (p. 15182). In contrast, other publications put less emphasis on training or argued that direct training was not necessary. These publications often stated the belief that educators were capable of learning along with students as they moved through the curriculum or, in some cases, that helping design the curriculum alone was enough to prepare educators. Chiu et al. (2022) was among the publications that argued dedicated educator training was not necessary. These researchers instead chose to involve educators directly in the curriculum design process to help familiarize them with AI concepts. This is reflected in the following quote, "Analyses reveal that all the participating teachers did not receive formal AI training, and they were able to learn the necessary AI knowledge for curriculum design from the co-creation process. They felt more qualified and confident to teach AI." (Chiu et al., 2022, p. 36).

A subset of publications designed AI curriculums for STEM education that were specifically targeted at educators. The aims of these curriculums were mixed between preparing teachers to educate students about AI and preparing educators to incorporate AI into their classrooms for general education purposes. A common theme among these publications was bridging the gap between increasing expectations about AI literacy and a lack of familiarity with a technology that has only become widely accessible relatively recently. This is further complicated by a lack of certainty regarding whether teachers' expertise in the educational field is a factor in how well they will be able to teach and AI curriculum. The following quote from Lin et al. (2021) is representative of this concern: "Thinking about how to teach AI has become important because people's demand for AI applications has increased; however, it is not easy to design a proper AI course which matches students' expertise in the educational field" (p. 234). With this quote, the authors acknowledge that they are inevitably working with a demographic of mixed experience and skill when working with educators that is difficult to make assumptions about.

***What themes imply for best practices in using AI for STEM education***

The three themes listed here generally reflect best practices that have emerged in previous research on the application of AI in education; however, this is the first such systematic literature review to focus specifically on best practices in STEM education. Previous studies reflecting these themes have focused on other sectors, like e-commerce (Sharma & Mohan, 2022), public health (Sharma et al., 2023), and general education (Gibellini et al., 2023). The field of STEM education is unique from these other contexts in that it has a direct relationship with AI technology and is the most appropriate place for students to encounter educational material on how to best use AI. In and of itself, AI is a powerful example of the STEM disciplines coming together in a single technological field and offers a valuable educational opportunity for students being introduced to these disciplines (Beigman Klabonov et al., 2017). This systematic review has offered three considerations for how educators may seek to maximize this educational opportunity.

As exemplified by Jang et al. (2022) and Yannier et al. (2020), AI technology can be used to teach about a wide variety of STEM topics. The consensus among AI-based curriculums indicates that, despite this flexibility, educators and researchers believe it important to first use AI as a means of introducing students to the technology itself. Doing so provides a clear focus on AI that can aid students in drawing connections between how the technology works, what its applications are, and what limiting factors and ethical issues may arise in using it. This focus on the technology itself may also help educators who are previously unfamiliar with AI technology, either through direct training or participation in the design of an AI curriculum.

Direct engagement with AI technology in applications that simulate real world situations has emerged as the best approach to introducing students to this technology. Educators throughout the publications reviewed appreciated that this approach helped students build connections between how AI works, what its applications are, what its limitations may be, and what ethical issues can arise when using it. Among researchers reviewed in this study, there was considerable variation on what type of AI technology was most appropriate for educational purposes. Some studies used only simple AI programs, such as Chatbots, that were only capable of natural language processing. Educators like Karampelas (2021) argued that these simple applications were enough to intrigue students at first while also quickly revealing the limitations of AI technology. Other researchers, such as Chiu et al. (2021), used a variety of AI technologies, often beginning with simpler applications, such as text generation programs, and gradually moving on to hardware with multiple AI capabilities, such as autonomous model vehicles capable of computer vision.

Finally, significant consideration has been given to assessing what kind of training educators need to be able to teach with and/or about AI technology. In this theme, there was some disagreement, with some publications arguing that teachers would learn as they engaged with the course material. Most publications recommended some level of training, with a portion of the curriculum reviewed here, such as Lee and Perret (2022), were focused specifically on this. Others trained teachers on AI by inviting them into the lesson design process, as exemplified by Chiu et al. (2021). This latter approach makes the most convincing case in terms of best practices, as the curriculum benefits from the experience of the educators while the educators benefit from the expertise of the AI researchers.

## 6. Conclusion and Future works

This systematic literature review sought to investigate what best practices were emerging in experimental pilots of STEM curriculums utilizing and/or focused on AI technology. A sample of seven publications were assessed through keyword analysis, offering 11 keywords that yielded three important themes in best practices. These were: the use of AI to teach about itself, the use of hands-on AI activities that simulate real world situations, and a focus on the expertise of educators teaching the units.

The current research effort was limited by the early stage at which AI integration into educational curricula currently stands. These seven publications represent early efforts at adopting the current wave of AI technology. As the field progresses, more publications will emerge and trends in best practice suggestions will further develop. Future research into what best practices have come from educational research may use the conclusions of this research as a point of comparison. If possible, it is recommended that future research efforts make a division between teacher and student AI training. Many of the publications reviewed here noted that educator familiarity with AI could be a major factor in student success, yet only a few too formal efforts to train educators. A dedicated review of research into educator AI training would be valuable research effort in terms of discerning best practices for training educators specifically.

## References

[1] Anghileri, J. (2006). Scaffolding practices that enhance mathematics learning. *Journal of Mathematics Teacher Education*, *9*, 33-52.

[2] Bacca Acosta, J. L., Baldiris Navarro, S. M., Fabregat Gesa, R., & Graf, S. (2014). Augmented reality trends in education: a systematic review of research and applications. *Journal of Educational Technology and Society, 2014, vol. 17, núm. 4, p. 133-149*.

[3] Beigman Klebanov, B., Burstein, J., Harackiewicz, J. M., Priniski, S. J., & Mulholland, M. (2017). Reflective writing about the utility value of science as a tool for increasing STEM motivation and retention–can AI help scale up?. *International Journal of Artificial Intelligence in Education*, *27*, 791-818.

[4] Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, *10*(1), 29.

[5] Chiu, T. K., Meng, H., Chai, C. S., King, I., Wong, S., & Yam, Y. (2021). Creation and evaluation of a pretertiary artificial intelligence (AI) curriculum. *IEEE Transactions on Education*, *65*(1), 30-39.

[6] Chng, E., Tan, A. L., & Tan, S. C. (2023). Examining the use of emerging technologies in schools: A review of artificial intelligence and immersive technologies in STEM education. *Journal for STEM Education Research*, *6*(3), 385-407.

[7] Cope, B., & Kalantzis, M. (2023). A little history of e-learning: finding new ways to learn in the PLATO computer education system, 1959–1976. *History of Education*, *52*(6), 905-936.

[8] Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative psychological and behavioral science*, *42*, 266-290.

[9] Gibellini, G., Fabretti, V., & Schiavo, G. (2023). Ai education from the educator's perspective: Best practices for an inclusive ai curriculum for middle school. *Extended*

*Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*.

[10] Ghosh, A., Chakraborty, D., & Law, A. (2018). Artificial intelligence in Internet of things. *CAAI Transactions on Intelligence Technology*, *3*(4), 208-218.

[11] Gonzalez, H. B., & Kuenzi, J. J. (2012). *Science, technology, engineering, and mathematics (STEM) education: A primer*. Washington, DC: Congressional Research Service, Library of Congress.

[12] Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542-570.

[13] Jang, J., Jeon, J., & Jung, S. K. (2022). Development of STEM-based AI education program for sustainable improvement of elementary learners. *Sustainability*, *14*(22), 15178.

[14] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255-260.

[15] Karampelas, A. (2021). Artificial Intelligence and Machine Learning in the STEAM classroom: Analysis of performance data and reflections of international high school students. *Hellenic Journal of STEM Education*, *1*(2), 59-66.

[16] Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology*, *51*(1), 7-15.

[17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

[18] Lee, I., & Perret, B. (2022). Preparing high school teachers to integrate AI methods into STEM classrooms. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 36, No. 11, pp. 12783-12791).

[19] Lin, C. H., Yu, C. C., Shih, P. K., & Wu, L. Y. (2021). STEM based artificial intelligence learning in general education for non-engineering undergraduate students. *Educational Technology & Society*, *24*(3), 224-237.

[20] Luzano, J. (2024). An Integrative Review of AI-Powered STEM Education. *International Journal of Academic Pedagogical Research*, *8*(4), 113-118.

[21] Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education*, *103*(4), 799-822.

[22] Oliveira, M., Bitencourt, C., Teixeira, E., & Santos, A. C. (2013, July). Thematic content analysis: Is there a difference between the support provided by the MAXQDA® and NVivo® software packages. In *Proceedings of the 12th European Conference on Research Methods for Business and Management Studies* (pp. 304-314).

[23] Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. (2012). Qualitative analysis techniques for the review of the literature. *Qualitative Report*, *17*, 56.

[24] Page, J., Bain, M., & Mukhlish, F. (2018, August). The risks of low level narrow artificial intelligence. In *2018 IEEE international conference on intelligence and safety for robotics (ISR)* (pp. 1-6). IEEE.

[25] Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M., & Yang, D. (2023). Is ChatGPT a general-purpose natural language processing task solver?. *arXiv preprint arXiv:2302.06476*.

[26] Russell, H. R., Wutich, A., & Ryan, G. W. (2016). *Analyzing qualitative data: Systematic approaches*. SAGE publications.

[27] Sakulkueakulsuk, B., Witoon, S., Ngarmkajornwiwat, P., Pataranutaporn, P., Surareungchai, W., Pataranutaporn, P., & Subsoontorn, P. (2018, December). Kids making AI: Integrating machine learning, gamification, and social context in STEM education. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 1005-1010). IEEE.

[28] Sharma, R., & Mohan, M. (2022). Artificial Intelligence and E-Learning: Best Practices Across the Globe. In *Adoption and Implementation of AI in Customer Relationship Management* (pp. 1-25). IGI Global.

[29] Sharma, S., Rawal, R., & Shah, D. (2023). Addressing the challenges of AI-based telemedicine: Best practices and lessons learned. *Journal of education and health promotion*, (1), 338.

[30] Torres-Carrión, P. V., González-González, C. S., Aciar, S., & Rodríguez-Morales, G. (2018, April). Methodology for systematic literature review applied to engineering and education. In *2018 IEEE Global engineering education conference (EDUCON)* (pp. 1364-1373). IEEE.

[31] Van Tuijl, C., & van der Molen, J. H. W. (2016). Study choice and career development in STEM fields: An overview and integration of the research. *International journal of technology and design education*, *26*(2), 159-183.

[32] Wahono, B., Lin, P. L., & Chang, C. Y. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *International Journal of STEM Education*, *7*(1), 36.

[33] Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *International Journal of STEM Education*, *9*(1), 59.

[34] Yannier, N., Hudson, S. E., & Koedinger, K. R. (2020). Active learning is about more than hands-on: A mixed-reality AI system to support STEM education. *International Journal of Artificial Intelligence in Education*, *30*, 74-96.

**Taha Mansor Khawaji** received the B.E. from Teacher College in Jazan in 2003, and M.E. degree from UOW, Australia in 2009. He received the PhD from QUT, Australia in 2016. He is an assistant professor (from 2016) in the Al-Qunfudhah University College & Dept. of Curriculum and Instruction, Umm Al-Qura University, Saudi Arabia. His research interests include educational technology, e-learning, ICT in education & training, Technology Enhanced Learning, STEM