# Exploring Effective Strategies and Challenges in Integrating AI Tools into Programming Education

University of Oulu
Information Processing Science
Master's Thesis
Niklas Riekki
2024

# Abstract

This thesis investigates the opportunities and challenges associated with integrating AI tools into the programming classroom. With the mainstream emergence of ChatGPT and similar technologies, it is imperative to understand their impact on student learning. This is particularly relevant for programming students, given the heightened interest and excitement surrounding generative AI in the programming field.

This study utilises qualitative interviews to uncover expert opinions on the current landscape and future challenges and opportunities. It highlights the potential benefits and drawbacks of integrating AI tools into education, emphasising the accessibility they bring to learning, alongside the risks of potential misuse that could undermine students' learning experiences.

Additionally, this thesis advocates for further research into specific avenues using a quantitative approach, building upon the qualitative foundation established. It emphasises the importance of measuring the effects on learning and validating strategies proposed in this thesis for the use of AI tools in education.

*Keywords*
artificial intelligence, AIEd, learning experience, programming education, AI tools

*Supervisor*
PhD, University Lecturer, Elina Annanperä

# Foreword

Thank you to my supervisor PhD, Elina Annanperä for her continued support and to all the interviewees who made this thesis possible.

# Contents

# 1. Introduction

In the landscape of education, the integration of Artificial Intelligence (AI) tools is emerging as a potentially transformative force, presenting unprecedented opportunities for altering the teaching and learning experience. This holds also true in the field of programming education especially with the rise of tools such as ChatGPT, GitHub Copilot, CodeWhisperer and many more. AI itself has also gained increased attention with the advent of the aforementioned tools, as can be seen in Figure 1. This master's thesis studies the realm of programming education, and whether established practices have emerged in combining AI tools with it, while keeping a focus on the perspective of the teacher.

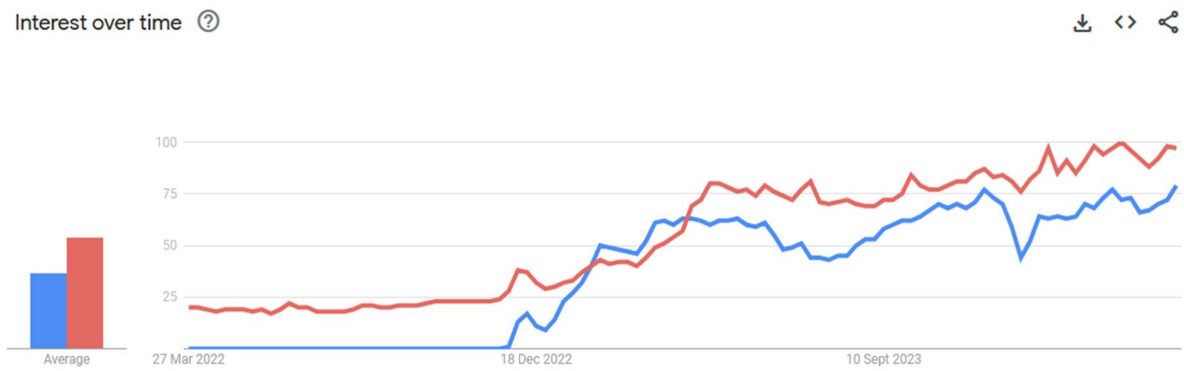


**Figure 1.** Interest over the past years in Google searches with terms "chatgpt" (red line) and "ai" (blue line).

While the affair between AI and education has gained recent attention, e.g. in the form of university guidelines (Guidelines for the Use of Artificial Intelligence in Education, n.d.), there remains a gap in understanding its application specifically from the viewpoint of the programming instructor, or teacher. The perspective of teachers, who are the orchestrators of the learning process, cannot be understated yet is underrepresented in research (Çelik et al. 2022). It is essential to explore the challenges, opportunities, and effective strategies available to teachers in integrating AI tools into programming education to develop a better understanding of the current landscape of teaching and learning programming.

This thesis seeks to investigate whether there already exists crossover between literature on programming education and Artificial Intelligence in Education (AIEd), and whether we can already identify opportunities, challenges, or beneficial practices to use in programming education. When examining the aforementioned matters, a keen focus will be pointed to the role of the teacher, as the aim of this thesis is to map out guidelines for programming teachers regarding the use of AI tools in their teaching. In addition to a literature review, interviews with programming teachers will be held to gauge the opinions of professionals of the field, as well as inquire whether teachers have themselves identified useful practices, obstacles, or anything in between.

The exploration of the above issues will be guided by the following research questions:

> **RQ1: Have effective strategies emerged in using AI tools in teaching programming?**

**RQ2: What are the opportunities for teachers in bringing AI tools to the programming classroom?**

**RQ3: What are the challenges for teachers in for teachers in bringing AI tools to the programming classroom?**

The first research question aims to generate a broad understanding of the current situation of the usage of AI tools with teaching and learning programming. Addressing the question will shed light on the current state of affairs when it comes to the effect of the emergence of AI tools in relation to programming and teaching it. The second and third questions will help bring the focus on the role and viewpoint of the teacher and will hopefully help illuminate the possible evolution of the programming teacher in an era of ubiquitous AI-powered learning assistants in the pockets of each student.

Understanding the effective integration of AI tools in programming education, particularly from the teacher's standpoint, holds implications for pedagogical advancements, curriculum design, and the learning experience of programming students. The findings of this study aim to inform educators and institutions about the potential benefits and challenges associated with incorporating AI tools into programming curricula. While touching on literature from a broader point of view, the primary focus of this research is specifically the teacher's perspective in utilising AI tools for programming education. While the broader context is acknowledged, the reflection and discussion will most closely zoom in on the experiences, practices, and perceptions of educators. This perspective may give rise to certain limitations, such as a reduced consideration of the student viewpoint.

The remainder of this thesis is structured as follows: Chapter 2 provides a comprehensive review of the literature, exploring the current state of programming education, the role of AI in education, and the intersection of AI and programming instruction. Chapter 3 details the research methodology employed to address the research questions, outlining the research design, data collection methods, and data analysis procedures. Chapter 4 presents the findings of the study, organised according to the identified good practices and insights derived from the teacher perspective. Chapter 5 will discuss the key findings and their implications and suggest avenues for future research. Finally, Chapter 6 offers a conclusion that summarises the main points of the thesis.

Notably aligned with the theme of this thesis, AI tools were employed in the writing process of this paper. ChatGPT was utilised for proofreading paragraphs, generating synonyms, and translating interview data from Finnish to English, while DeepL Translate was also employed for synonym assistance and in translating.

# 2. Prior Research

This section will go over related literature, first on programming education, aiming to uncover methods that have proven effective. Next, a section on AIEd which aims to investigate research on both the potential and challenges related to the field. Following them is a section on artificial intelligence in programming education bridging the gap between programming education and AIEd. Closing the chapter is a review on the ethical considerations of artificial intelligence. The goal of this chapter is to identify the potential crossroads where the beneficial outcomes from AI meet the needs of programming educators and students, without forgetting to keep a close eye on the ethical considerations that AI brings with it.

## 2.1 Programming Education

Teaching programming presents a challenge, particularly in catering to the diverse needs of students across varying skill levels. For instance, students with prior experience in programming might find introductory topics redundant, while complete beginners may require more foundational support and encouragement. Thus, educators are encouraged to employ a variety of teaching methods, such as differentiated instruction, where tasks are tailored to the learning needs of students at different levels. The efficacy of teaching programming standards depends on the delivery of personal instruction and the provision of individualised feedback. A one-size-fits-all approach has limitations as not every student will be an identical learner. (Li & Prasad, 2005). Additionally, the findings of Latulipe et al. (2018) underscore the complex relationship between student demographics and pedagogical strategies, as they found that the gender or ethnic background of the student can influence what kind of teaching arrangement will be the most efficient while studying flipped classrooms in an introductory computer science course. As these innate characteristics can determine which learning environment or engagement strategy works best for individuals, the standardised way of teaching programming is brought further into question.

The prevailing methods of teaching programming, as noted by Robins et al. (2003) in their comprehensive review of learning and teaching programming, focus on promoting deep learning among students. However, Robins et al. observed that a significant number of students struggle to achieve a substantial understanding of programming through introductory courses, indicating a potential shortfall in the effectiveness of these teaching methodologies within such a context. Additionally, a comprehensive survey by Pears et al. (2007) underscored the absence of consensus regarding a singular, optimal approach to programming instruction, as underscored by the lack of extensive, systematic studies necessary to address such a complex issue conclusively. This observation suggests a gap in research that warrants further investigation to enhance teaching methodologies in introductory programming courses. The discourse regarding learning in programming has continued in more contemporary educational research. In this context, Rubiano et al. (2015) critically evaluated the conventional priorities of programming course designs, which tend to emphasise the mastery of programming languages and the use of software tools, rather than fostering problem-solving skills. Rubiano et al. argued that such practices overlook the foundational learning challenges encountered by students. Instead, they advocate for a complete paradigm shift in programming education, proposing an approach that prioritises the cultivation of problem-solving competencies, alongside or even before, the technical aspects of programming.

Research in the domain of educational psychology suggests that the prevailing minimal guidance approach in teaching complex subjects - such as programming - lacks empirical support. Kirschner et al. (2006) provide a critical analysis of minimal guidance approaches, as they are predicated on the assumption that learners can effectively construct knowledge from experiences that are only loosely guided by the instructor. However, Kirschner and colleagues argue that such methods can inadvertently foster misconceptions and cognitive overload, significantly impeding the learning process. Kirschner et al. caution against the minimal guidance approach, highlighting its potential to cultivate misconceptions in the minds of students and hinder their progress. Furthermore, Kirschner et al. found that even when students possess a body of prior knowledge, a minimal guidance approach does not outperform strong guidance.

Formative feedback, as defined by Shute (2008), is targeted, informative communication aimed at modifying learners' thinking or behaviour to enhance their learning. It is a process designed not merely to evaluate but to foster growth and improvement. Effective formative feedback is supportive, timely, and specific, focusing on providing learners with actionable insights into their performance in relation to their learning objectives. Shute emphasises that formative feedback's power lies in its ability to provide learners with a clear understanding of where they stand in their learning journey and what steps they need to take to progress further. This involves not just pointing out errors or confirming correctness but offering detailed guidance on how to improve. The timing of feedback is crucial; it needs to be delivered at a juncture where it can be most impactful, which might mean immediately to prevent the reinforcement of misconceptions for some situations, or with a slight delay to promote reflection and deeper understanding for others. Moreover, formative feedback's effectiveness is heightened when it is personalised, taking into account individual differences among learners, such as their prior knowledge, skills, and motivation. This personalised approach ensures that feedback is not a one-size-fits-all solution but a tailored support mechanism that recognizes and addresses the unique needs and goals of each learner. Shute highlighted that despite extensive research on feedback, evidenced by hundreds of articles, the findings regarding specific aspects of feedback have been inconsistent and unreliable, underscoring the necessity for adopting a distinct, formative approach.

Anfurrutia et al. (2018) underscore the significance of providing formative feedback throughout introductory programming modules, particularly emphasising its role in enhancing student engagement and understanding. Their study combined the use of educational robots and visual programming environments with automated code verification and validation feedback mechanisms, aiming to provide a more enriched learning experience. This approach allows students to learn from their mistakes in real-time while supporting the diverse learning needs of students, catering to varying levels of prior knowledge and learning speeds. Their findings reveal a positive impact on student performance and the percentage of passing grades when students receive ongoing formative feedback. Similarly, Hao et al. (2021) corroborate these findings, emphasising the importance of detailed feedback in student success as they investigated the impacts of detailed feedback on student performance and interaction with feedback systems. Utilising an experimental-like design with students receiving various levels of automated feedback, they demonstrate that detailed feedback significantly enhances student performance, engagement, and their ability to correct mistakes. The study further substantiates the need for well-designed feedback mechanisms in educational settings, particularly in programming courses where the complexity of tasks can vary widely. It highlights ways in which feedback can be optimised to improve student learning experiences and outcomes, suggesting that educators should prioritise automatic delivery of constructive feedback.

The Extreme Apprenticeship Method, as detailed by Vihavainen et al. (2011), represents an alternative approach to the pedagogy of introductory programming courses. This method not only addresses the high dropout rates and the challenges novice programmers face, but also introduces a paradigm shift in teaching methodologies. By emphasising a hands-on, practice-oriented learning environment established on continuous, personalised guidance and feedback, the method fosters a more engaging and supportive atmosphere for students. The apprentice-and-master dynamic ensures that learning is actively constructed through the process of doing, significantly enhancing student motivation and success rates. This methodology further underlines the importance of adapting teaching strategies to meet the diverse needs and learning styles of students, promoting not just knowledge building but also the development of problem-solving skills and critical thinking. The Extreme Apprentice Method builds on Cognitive Apprenticeship, which as described by Collins et al. (1989), offers a structured framework for knowledge transfer, where a more experienced master imparts knowledge and skills to a novice through close mentorship and guided practice. By situating learning in the context of real-world problems and tasks, novice programmers are equipped to navigate the complex process of learning programming with greater efficacy. The method promotes an immersive learning environment, where continuous interaction and feedback from the mentor significantly enhance the novice's understanding and application of programming concepts, leading to a more effective learning experience. Additionally, Caspersen & Bennedsen (2007) emphasise the advantages of cognitive apprenticeship in programming education, highlighting how it mitigates the risk of overlooking foundational concepts during early learning stages. This approach ensures that experts, who might consider some basics as trivial due to their familiarity, deliberately address these foundational elements with novices. Such attention to the foundational stages of learning ensures a comprehensive understanding, preventing gaps in knowledge that could hinder the progression of novice programmers.

Cognitive load theory pertains to how we utilise cognitive resources when solving problems or learning new information. The theory suggests that problem-solving activities can heavily tax our cognitive capacity, limiting our ability to process and retain new information effectively. This is particularly relevant in learning environments where students are faced with complex tasks, as students will then have limited capacity in their working memory to absorb new information beyond the immediate task at hand. Conversely, learning through working examples imposes lower cognitive load and facilitates easier processing, allowing students to focus their attention on the learning process itself. As the use of cognitive resources is optimised, this approach facilitates a more effective learning experience, promoting deeper understanding and retention of new information (Chandler & Sweller, 1991).

Rubin's (2013) research into the effectiveness of live coding in introductory programming education underscores its potential as a learning method in the field. By utilising live examples, educators can shed light on complex programming concepts, making them more accessible and relatable to students. This approach not only aids in contextualising abstract concepts but also significantly enhances students' problem-solving skills. Engaging students with real-world programming scenarios allows them to discern meaningful patterns and conceptual structures more effectively, thereby fostering a deeper understanding and facilitating an interactive and dynamic learning environment. Through this method, Rubin advocates for a shift towards a more engaging and practical approach to programming education, highlighting the role of live examples in facilitating a comprehensive understanding of programming concepts. Additionally, Caspersen & Bennedsen (2007) also observed that by exposing students to working examples, educators can effectively scaffold learning, allowing students to identify and understand meaningful information and the underlying conceptual structure of programming concepts. This pedagogical strategy not only engages students more deeply but also builds a more intuitive grasp of programming structures and problem-solving techniques among the students, jumpstarting their progress from novices to proficient programmers by making abstract concepts more tangible and relatable.

The building blocks of teaching programming successfully are cognitive load, cognitive apprenticeship and working examples (Caspersen & Bennedsen, 2007). The researchers underscore the importance of cognitive load management and the integration of cognitive apprenticeship principles into pedagogical practices. Through working examples, students can mitigate cognitive overload and direct their attention towards active learning. This notion is reinforced by the work of Morrison et al. (2014), as their work expands on understanding cognitive load in educational settings, particularly in the realm of computer science education. Their research into developing an instrument for measuring cognitive load highlights the impact of pedagogical approaches on cognitive load levels experienced by students. By differentiating between text-based materials and those incorporating visual and auditory elements, their findings provide empirical evidence of the need for diverse instructional strategies. These strategies not only reduce cognitive load but also enhance the learning process by catering to the varied cognitive capacities of learners. Jia & Hermans (2022) emphasised the importance of clarity in programming instruction, arguing that clear instructions are crucial for engaging students' cognitive processes efficiently. Their research supports the notion that well-structured and clear instructions can significantly impact learning outcomes in programming education by making complex concepts more accessible and understandable. This approach not only aids in reducing cognitive load but also in enhancing student engagement and motivation.

Wilson & Shrock (2001) provided insightful evidence on the importance of creating a comfortable learning environment in programming classes. They found that students' comfort levels significantly impact their performance, suggesting that an atmosphere where students feel free to engage, ask questions, and express themselves leads to better learning outcomes. Such an environment encourages psychological safety and engagement, critical for student success. Furthermore, the study highlights the negative effects of attributing success to external factors like luck, which can undermine students' belief in their abilities and effort, emphasising the need for educators to promote a sense of agency in their students.

Bergin et al. (2005) highlight the significance of intrinsic motivation and task value in fostering cognitive and metacognitive strategies among students. Cognitive strategies, such as creating analogies, help students in connecting new information with old information. Metacognitive strategies are used to make comprehension easier and can help manage resources more efficiently. Examples of metacognitive strategies are, for example, setting goals to aid in planning or ways of managing time and environment. By cultivating an environment that promotes enjoyment and perceived task significance, educators can enhance student engagement and self-regulated learning practices.

Ultimately, an understanding of the cognitive processes involved in programming is crucial for effective instruction. Robins et al. (2003) noted the cognitive challenges that novice programmers face in understanding code execution principles. They emphasised that beginners often struggle to form mental models in the way experienced programmers do, which impedes their ability to comprehend how computers interpret code. This gap in understanding underscores the need for educational approaches that promote the development of these mental models, enabling novices to bridge their conceptual understanding with the logical structures that underpin programming languages and their execution by computers. Furthermore, Ahmadzadeh et al. (2005) brought into light the disparity between novice and professional programmers, particularly in the realm of debugging. Despite achieving good grades in introductory programming courses, novice students often struggle with debugging, a critical skill in programming. The findings of Ahmadzadeh et al. suggest an imperative for curriculum designs to integrate debugging practices more thoroughly, ensuring students not only understand programming languages but can also effectively troubleshoot problems in their code, thereby bridging the gap between theoretical knowledge and practical application.

## 2.2 AIEd

In recent years, education paradigms have shifted from viewing knowledge as a static product to emphasising learning as a dynamic process. Roll & Wylie (2016) highlight how the shift in educational paradigms toward dynamic learning processes opens new avenues for AI in personalising education. By harnessing AI, educators can tailor learning environments to individual needs, thereby enhancing engagement and efficacy. Through AI, education can become more inclusive and adaptive, accommodating diverse learning styles and needs, thus revolutionising traditional educational models. Chen et al. (2020) explored the evolution of AI in education, emphasising its transition from traditional computer-based tools to sophisticated web-driven solutions, including chatbots as learning assistants. This progression signifies a shift towards more interactive, responsive educational environments, where AI technologies play a pivotal role in personalising learning experiences and enhancing student engagement. Through the deployment of AI, including chatbots, education systems can offer more tailored support and feedback, addressing individual student needs and promoting more effective learning pathways. However, while there is optimism surrounding AIEd, Zawacki-Richter et al. (2019) highlight the critical need for integrating theoretical and ethical considerations into the application of AIEd, especially in higher education contexts. They argue for a deeper exploration of the ethical and educational frameworks that guide AIEd, emphasising that such an approach is necessary to navigate the challenges and risks associated with implementing AI technologies. As the theoretical foundations of AIEd are strengthened and made both educationally beneficial and ethically sound, AI's adoption in educational settings will be more informed and reflective.

Chassignol et al. (2018) examined the transformative role of AI in education, highlighting its capacity to facilitate massive open online courses (MOOCs) and tailor learning experiences to individual student needs, even within large-scale educational settings. Their analysis underscores the potential of AI to upheave traditional teaching methods by providing personalised learning pathways, thereby enhancing both accessibility and effectiveness of education across diverse learning environments. This advancement signifies a step towards more inclusive and adaptive educational practices, leveraging technology to meet the varied demands of the global student population. Wang et al. (2023) explored the impact of AI on curricula development for the growing population of international students, emphasising AI's potential to customise educational experiences. Their work suggests that AI can address the diverse and evolving educational needs of international students, improving access, engagement, and outcomes in higher education. This approach represents a path towards leveraging AI technology to support the unique challenges faced by international students, once again facilitating a more inclusive and adaptive learning environment while also aligning with the broader shift in educational paradigms towards dynamic and learner-centred processes as underlined by Devedžić (2004). Devedžić advocates for the significance of personalised learning approaches, stressing that adaptability and learner comfort are central to the principles of AIEd. By emphasising personalised learning, Devedžić underscores the importance of creating educational experiences that are tailored to the unique needs and preferences of each student, thereby enhancing engagement and efficacy in learning outcomes. Tapalova & Zhiyenbayeva (2022) also investigated AIEd's potential to customise education for an increasingly diverse student population. Their work underscores the importance of adaptive learning systems that cater to the unique characteristics, backgrounds, and talents of each student, reinforcing the potential of educational technologies that offer personalised learning experiences. This approach not only aligns with the evolving demands of global education but also emphasises the role of AIEd in enhancing student engagement and success by providing learning environments that are responsive to individual learner profiles.

Timms (2016) envisioned a future where AIEd, through educational robots and smart classrooms, significantly enhances personalised learning experiences. By integrating AI with advances in robotics and sensor technology, educational environments can become more interactive and supportive, offering tailored assistance, and fostering deeper engagement in learning. This forward-looking perspective underlines the potential of AI not only to augment traditional teaching methods but also to create more dynamic and responsive educational settings tailored to individual student needs and preferences. Guan et al. (2020) underscored the pivotal shift in AIEd research towards student profiling models and learning analytics. Their analysis sheds light on the emerging focus on leveraging data to understand and enhance the learning process, emphasising the role of continuous assessment in modern educational paradigms. This transition marks an evolution from traditional educational methodologies, proposing a data-driven approach that supports personalised learning environments and informed instructional strategies. Luckin & Holmes (2016) also noted the transformative potential of AIEd, particularly in enhancing continuous assessment and moving away from traditional testing methods. They highlight the shift in educational assessment models, where AI enables ongoing evaluation without interrupting the learning process, thereby promoting a more dynamic and responsive educational environment. This approach would not only support a more nuanced understanding of student progress, but also align with contemporary educational needs for adaptability and personalised learning experiences.

The concept of scaffolding, as elaborated by Wood et al. (1976), encapsulates a teaching methodology aimed at enhancing independent problem-solving skills. The approach emphasises the supplying of hints rather than direct answers, thus encouraging learners to engage deeply with the task at hand and develop their problem-solving capabilities. By controlling task elements that are beyond the learner's current capacity, scaffolding allows for focused attention on manageable aspects, facilitating a more effective learning experience. Lim et al. (2023) explored the efficacy of AI-assisted personalised scaffolding in self-learning scenarios, aiming to enhance student performance. While their study did not find significant improvements in performance outcomes, it did reveal an increase in the use of metacognitive strategies by students. This suggests that while direct performance metrics may not have shown improvement, the implementation of personalised scaffolding has potential to bring about more self-regulated learning behaviours in students as their learning process becomes more strategic. Providing one potential explanation in their study, Huang et al. (2023) explored the significant role of motivation in the context of AI-powered personalised learning environments. Their research emphasises that while AI tools are increasingly utilised to enhance learning experiences through personalised recommendations, the intrinsic motivation of learners remains a critical factor in determining the effectiveness of these technologies. This suggests that technology alone is not a cure-all, rather, its success in improving educational outcomes is closely tied to how it engages and motivates students, reinforcing the importance of considering learner motivation in the design and implementation of AI in education.

Nevertheless, AIEd is not without its challenges. Pedró et al. (2019) in their comprehensive examination of the integration of AI in education, pinpointed various challenges in the realm of AIEd. These challenges include issues related to inclusivity and equity, underscoring the need for an education system that caters to all demographics. Teacher preparation emerges as another critical aspect, emphasising the necessity for educators to be well-versed in AI technologies and their pedagogical applications. The accuracy of AI in educational contexts, its impact on the social dynamics of learning environments, and the evolving nature of student-teacher relationships due to AI intervention are also highlighted. Additionally, the discussion extends to the overarching policy frameworks that govern the ethical and transparent use of AI in education, advocating for robust theoretical foundations to guide its implementation and ensure that AI serves as a beneficial educational tool rather than a disruptive force. Ng et al. (2023) studied the complexity of integrating AI into educational settings, focusing on the challenges educators face in fully utilising AI tools due to a gap in their AI competencies. This gap not only hinders the effective use of AI in enhancing learning environments but also reflects the broader challenge of aligning teacher education with rapid technological advancements. The study highlights the crucial need for professional development and training programs that equip teachers with the necessary digital and AI-related competencies, ensuring they can confidently navigate and implement AI technologies to support student learning.

Çelik et al. (2022) underscored the need for AIEd research to more thoroughly consider the perspectives and experiences of teachers, who are fundamental to the successful integration and adoption of AI in educational settings. Despite their critical role, there exists a notable gap in understanding how teachers interact with, perceive, and can effectively use AI tools within their pedagogical practices. This oversight in the literature suggests a pressing need for a more teacher-centric approach in AIEd research, focusing on professional development, ethical use, and practical application of AI technologies to support and enhance teaching methodologies. Crompton & Burke (2023) emphasised the observation that AIEd research has predominantly centred on student perspectives, accounting for over 70% of studies, potentially skewing the comprehensive understanding of AI's impact across all facets of educational practices. This oversight suggests a significant gap in recognizing and incorporating the views and experiences of teachers, who play a crucial role in the practical application and integration of AI technologies in educational settings.

Chounta et al. (2021) studied the perceptions of teachers towards the integration of AI in their pedagogical practices, highlighting a general appreciation for AI as a supportive tool. Despite acknowledging its potential benefits, their research uncovers a notable scepticism among educators concerning AI’s accuracy and its impact on the social aspects of learning, particularly the dynamics of student-teacher relationships. This ambivalence reveals the potential need for addressing both the technical and relational dimensions of AI in educational settings, emphasising the importance of maintaining the human element in teaching and learning processes. Touching upon the same vein, Vivar & García-Peñalvo (2023) discussed the integration of AI in education and its potential to redefine the role of teachers, emphasising a shift towards prioritising the social aspects of teaching. This evolution allows tasks like assessment to be managed by AI, suggesting a significant transformation in educational practices and the teacher-student dynamic, encouraging a focus on enhancing interpersonal interactions and emotional support within the learning environment.

Schiff (2021) critically examines the lack of focus on the ethical dimensions of AIEd within global AI policy strategies, observing that policies primarily frame AIEd as a tool for creating an AI-literate workforce rather than exploring its broader educational benefits. This analysis reveals a significant gap in policy discussions around the ethical implications and transformative potential of AIEd, suggesting a need for more comprehensive policy frameworks that recognize and address these critical aspects. Flores-Viva & García-Peñalvo (2023) highlighted the imperative for AI developers to integrate ethical considerations throughout the development process of AI systems. This ethical foresight is deemed crucial for ensuring AI technologies can effectively contribute to equalising educational opportunities for students from diverse backgrounds, thereby addressing inclusivity and equity within the learning environment. Their emphasis on the ethical dimension underscores the broader responsibility of AI development to contribute positively to societal goals, particularly in education.

## 2.3 Artificial Intelligence in Programming Education

Nwana (1990) outlines the innovative potential of AI in programming education, particularly through the development of intelligent tutoring systems (ITSs). These systems are distinguished by their ability to "know what they teach, who they teach, and how to teach it". By leveraging AI, ITSs aim to replicate effective human teaching behaviours, embodying good teaching practices in a digital format. This capability suggests a significant leap forward in educational technology, offering personalised learning experiences that can adjust to the individual needs of students, thereby enhancing the learning process and outcomes in programming education. Anderson et al. (1985) also underscore AI's potential in programming education through ITSs, as they provide personalised feedback and tailored learning experiences, simulating the instructional capabilities of human tutors, adapting to the unique learning needs of each student. Eilermann et al. (2023) highlight the significant benefits of employing ITSs within programming education, especially in automation. Their work presents ITSs as both complementary tools to traditional teaching methods and standalone resources for self-learning. By personalising the learning experience, ITSs can adjust to the individual needs of students, offering customised tasks and feedback. This adaptability not only enhances the learning process but also addresses the unique challenges of teaching programming within the specialised field of automation, showcasing the versatility and potential of ITS in modern educational environments. However, Crow et al. (2018) uncovered a significant gap in ITSs for programming education, focusing on the disconnect between the domain knowledge embedded within these systems and the specific content needs of learners. They identify that while ITSs possess considerable programming knowledge, they frequently lack comprehensive, programming-specific reference material. This shortfall can frustrate students who struggle to access relevant information, as the system's knowledge base, crucial for enhancing learning, remains hazy and inaccessible to them, underlining the importance of integrating comprehensive reference materials into ITS design.

Anderson et al. (1985) brought attention to the capability of ITSs to utilise adaptive pedagogical models, which allow for the creation of customised teaching strategies that cater to the varied needs of students. This approach not only delivers personalised feedback directly to learners but also equips educators with insights into the learning styles and progress of their students, thereby enhancing the overall pedagogical framework and effectiveness of educational practices. Nonetheless, Figueiredo & García-Peñalvo (2020) stressed the indispensable role of teachers within the framework of ITSs, underscoring that despite the advanced capabilities of these systems, teacher guidance and the relational dynamics in the classroom remain crucial to the learning process. The expertise of teachers guides the ITS to better serve student needs and ensures that the technology complements, rather than replaces the human aspects of teaching and learning.

Real-Fernández et al. (2019) discuss the importance of designing instructional strategies that allow intelligent systems to autonomously provide meaningful learning experiences to students. They emphasise that teachers must carefully create assignments and adjust instructional strategies to enable these systems to deliver personalised content that aligns with the desired learning outcomes. This approach ensures that the technology serves as an effective tool in the learning process, highlighting the teacher's role in guiding and enhancing the use of intelligent systems for educational purposes. Banić et al. (2023) investigated the integration of AI tools like ChatGPT into programming education, noting their potential to boost student motivation and cognitive strategies. This inclusion signifies a promising enhancement to traditional teaching methodologies, offering personalised learning experiences. However, the study also cautions educators about the implications of increased workload, suggesting the need for careful consideration in balancing technological integration with pedagogical objectives to maximise the benefits for both students and teachers. Ismail and Ade-Ibijola (2019) emphasise the importance of AI assistants in providing both academic and emotional support to students. These AI tools offer a low-threshold, non-intimidating way for students to seek help, potentially reducing the emotional pressure associated with asking for assistance directly from human educators. This approach not only supports students in their learning journey but also acts as a valuable resource for educators, enhancing the overall educational experience by fostering an environment where students feel more comfortable and supported. This assertion is corroborated by Peng & Wan (2023), who explored the nuanced relationship between AI teaching assistants (TAs) and students, particularly those experiencing social anxiety. Their study revealed a preference for AI assistance over human interaction in such students, attributed to the reduced emotional pressure when engaging with AI. However, as the complexity of problems increases, students' preference for AI assistance decreases, indicating a nuanced relationship between the nature of the task and the choice of assistant. This research underscores the importance of considering both psychological factors and the inherent limitations of AI in educational settings.

Andersen et al. (2022) discussed the significant benefit of integrating AI tools in education, particularly their ability to provide immediate feedback, which can greatly alleviate the workload of teachers. AI tools were found to not only aid in streamlining the educational process but also enhance the learning experience by offering timely and personalised feedback to students, thereby facilitating a more efficient and effective learning environment. Saari et al. (2023) emphasise the evolving role of educators, as it extends to guiding students on the effective use of AI tools within programming courses. As AI becomes more integrated into educational environments, particularly in programming, teachers play a vital role in enabling students to utilise these technologies responsibly and effectively. This necessitates a shift towards a more facilitative teaching approach, focusing on the responsible use of AI tools, to harness their potential in enhancing the learning experience while addressing the challenges and expectations associated with their use. Puryear and Sprint (2022) also discussed the need for teachers to distinguish between human and AI-generated work. This task requires educators to adjust teaching and assessment methods to uphold academic integrity while making the most of AI's potential to aid learning. The shift emphasises practical challenges, such as ensuring students learn critical thinking and coding fundamentals despite the ease offered by AI. It also highlights the need for a balance, ensuring AI tools serve as aids rather than replacements for learning, and preparing students to understand and ethically use such technology. However, Martín Núñez & Diaz Lantada (2020) underscore the significant gap in traditional teacher training programs regarding the incorporation and understanding of AI technologies. This discrepancy arises as AI begins to transform engineering education, where its integration into pedagogy could significantly enhance the teaching-learning process. The challenge lies not only in the adaptation of current educational frameworks to include AI but also in ensuring that educators are equipped with the necessary knowledge and skills to effectively utilise AI tools. The situation potentially calls for an adjustment in teacher education, as AI literacy becomes an increasingly important component of teacher training programs of the future.

Recent advancements in AI, exemplified by tools such as ChatGPT, have catalysed pedagogical shifts in programming curricula, prioritising the programming process over mere output. This paradigm shift is mirrored in educators' growing preference for supervised assessments over traditional homework assignments. Prather et al. (2023a) explore this transition, highlighting an approach towards assessment that values the process of problem-solving and comprehension, over the final code, fostering a deeper learning experience. This change could also be reflective of a broader educational shift towards active learning strategies and critical thinking. However, Hellas et al. (2023) investigated the necessity of programming knowledge in effectively utilising AI tools such as ChatGPT for educational feedback and found prior programming knowledge critical in their use. They examined the challenge educators and learners face in distinguishing between valid feedback and incorrect or “hallucinated” problems generated by these tools and found that it is important to understand programming concepts beforehand for students to be able to filter potential misinformation provided by the tools and enhance learning outcomes. Ouh et al. (2023) echoed similar sentiments while investigating ChatGPT's efficacy in passing an introductory Java course. Their study reveals ChatGPT's ability to generate accurate and structurally sound Java solutions, highlighting its potential to offer alternative, efficient coding methods. However, the research underlines the importance of critical evaluation of ChatGPT's responses by students, turning the evaluation process into a learning opportunity itself. Nevertheless, this approach has the potential to lead students to deepen their understanding and critical thinking skills, underlining the tool's value as an educational ally when used judiciously. Kazemitabaar et al. (2023) conducted a study to examine the impact of AI code generators on the learning outcomes of novice programmers. Their findings reveal that novices using AI code generators like OpenAI Codex completed programming tasks at a significantly higher rate compared to those not using such tools. This suggests that AI code generators can be effective educational aids, enhancing task completion efficiency and potentially accelerating the learning process. However, Prather et al. (2023b) uncovered potential unintended consequences of using AI code generation tools like GitHub CoPilot among novice programmers, noting a tendency to over-rely on these tools due to perceptions of infallibility. This reliance potentially impedes critical learning processes, as novices often accept large code blocks without fully understanding them, diverting valuable cognitive resources from learning to deciphering. The study suggests that while such tools can speed up task completion, they might also encumber the learning experience by making it more difficult for novices to differentiate between correct and incorrect suggestions, thereby hindering their development of problem-solving skills and understanding of programming concepts.

Considering the broader context, Van Slyke et al. (2023) advocated for the strategic integration of AI tools like ChatGPT into educational settings, recognizing their inevitable use by students. This perspective emphasises the importance of guiding both educators and students on the effective use of these technologies to enhance learning outcomes. By educating on their application, the potential of AI tools can be maximised, transforming them from mere technological advancements into significant educational assets. This approach not only acknowledges the reality of AI tool utilisation but also leverages it to enrich the educational experience. Lau & Guo (2023) found that while programming educators are actively discussing AI tools like ChatGPT among peers, they hesitate to integrate these tools into their curriculum. This reluctance stems from a lack of comprehensive insight into the extent of student engagement with these AI resources. Despite the prevalent conversations about AI's role in education, there is a gap in applying these discussions toward practical curriculum development, indicating a cautious approach towards embracing AI in educational settings.

## 2.4 Ethical implications in AIEd

As briefly touched upon earlier, the ethical implications of implementing artificial intelligence in educational settings is an important aspect to regard, yet it is underrepresented in the research landscape. Mouta et al. (2023) conducted a comprehensive systematic literature review to explore the ethical dimensions of artificial intelligence (AI) technologies in education, revealing both the potential and the challenges of integrating AI into educational practices. Their investigation underscores a critical gap in the current research landscape, particularly in terms of addressing ethical considerations, cultural inclusion, and emotional engagement within AI-enabled educational environments, as the research of the ethical dimensions of AIEd is largely focused on western people and the realm of STEM. Despite the growing enthusiasm for AI's capabilities to enhance learning experiences and administrative efficiency, the review highlights a significant oversight concerning the implications for privacy, equity, and the dehumanisation of educational interactions. Moreover, the authors point out the need for ethical frameworks tailored to the educational sector to ensure that AI technologies are implemented in ways that are both equitable and conducive to fostering inclusive learning experiences.

In a related study, Holmes et al. (2021) delved into the need for an ethical framework tailored to the integration of AI in educational settings. The study gathered insights from leading researchers in the field, revealing a consensus on the necessity of addressing issues such as fairness, accountability, transparency, bias, autonomy, agency, and inclusion within AI-enabled educational technologies. Despite the widespread acknowledgment of these ethical imperatives, the paper highlights a gap in existing research and practice, with most AIEd researchers lacking the necessary training to navigate these ethical complexities. Through a comprehensive analysis of survey responses, Holmes et al. propose the development of an actionable ethical framework that not only addresses the specific ethical challenges posed by AI in education but also promotes a responsible, inclusive, and equitable approach to its application. This framework aims to guide educators, policymakers, and technology developers towards ethical decision-making and implementation, ensuring that AI's integration into education maximises benefits for all stakeholders while mitigating potential harms.

However, guidelines for the ethical integration of AI into education have emerged recently. The European Commission's "Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators" (2022) outlines a comprehensive framework aimed at fostering the responsible integration of AI technologies within educational settings. Developed with the support of an expert group, the guidelines emphasise the importance of addressing ethical challenges and leveraging AI's potential to enhance educational practices. Key considerations such as human agency, fairness, humanity, and justified choice underpin these guidelines, ensuring that the deployment of AI technologies respects human dignity and promotes inclusivity and equity in education. The document offers practical guidance for educators, including key requirements for trustworthy AI, guiding questions for reflective practice, and strategies for effective implementation. It advocates for ongoing professional engagement, digital resource management, and the empowerment of learners through ethical AI use, aiming to equip educators with the knowledge and tools needed to navigate the complexities of AI in education ethically and effectively.

Within the narrower context of generative AI, UNESCO has also published guidelines in their effective use. Their "Guidance for Generative AI in Education and Research" (Miao & Holmes, 2023) addresses the rapid evolution and integration of generative AI tools in educational settings, emphasising the urgent need for a human-centred approach to their use. With the proliferation of generative AI technologies like ChatGPT and the pace at which they are evolving, existing national regulatory frameworks struggle to keep up, leaving user data privacy unprotected and educational institutions unprepared. The guidance aims to assist countries in implementing actions, formulating long-term policies, and developing human capacities to navigate the challenges posed by GenAI. It highlights potential risks to core humanistic values such as human agency, inclusion, equity, gender equality, and linguistic and cultural diversities. UNESCO proposes steps for regulatory bodies to mandate data privacy protection and considers age limits for generative AI tool usage. It also underscores the importance of validating generative AI systems for ethical and pedagogical appropriateness in educational contexts.

In their paper, Adams et al. (2023) delved into the ethical considerations necessary for integrating AI into K-12 education (from kindergarten to 12th grade), offering a detailed exploration of the landscape shaped by AI's proliferation in educational settings. Their investigation illuminated the gradual but significant shift in ethical guidelines, driven by both global discourse and the unique demands of ensuring children's welfare and teachers' well-being within the educational ecosystem. By analysing recent and globally relevant AI ethics guideline statements through the lens of previously identified AI ethics principles, Adams et al. not only affirm the applicability of core ethical principles such as transparency, justice, and non-maleficence in educational contexts but also introduce novel principles tailored to K-12 education, including pedagogical appropriateness, children’s rights, AI literacy, and teacher well-being. Moreover, their discussion advocates for a relational and humanised approach to AI integration that respects the relationship between humans and technologies.

Similarly, Akgün & Greenhow (2021) underlined the importance of addressing the ethical implications of AI deployment in their paper on the ethical challenges in K-12 settings, including issues related to privacy, surveillance, autonomy, bias, and discrimination. Akgün and Greenhow argue for the necessity of educating both teachers and students about these ethical challenges, proposing the adoption of AI and ethics-focused curricula. Nguyen et al. (2022) also addressed the interplay between technological innovation and ethical considerations in the realm of AIEd. Their research analysed international policies and guidelines to forge a set of ethical principles that stakeholders should follow when integrating AI into educational settings. Recognizing AI's potential to transform learning while acknowledging associated ethical risks, such as issues of data privacy, learner autonomy, and bias, the study underscores the need for a balanced approach. By proposing a comprehensive ethical framework, which underlines principles in governance and stewardship, transparency and accountability, sustainability and proportionality, privacy, security and safety, inclusiveness and human-centeredness, the authors aim to guide educators, developers, and policymakers in deploying AI technologies responsibly.

Swist et al. (2024) presented a pioneering study on the co-production of ethical principles for AI-enabled Educational Technology (AAI-EdTech) under the framework of Deliberative democracy (DD) within a university setting during the COVID-19 lockdown. Deliberative democracy refers to “an association whose affairs are governed by the public deliberation of its members" (Cohen, 1989), which in the context of this study refers to students, teachers, and institutional leadership. The empirical investigation of Swist et al. showcased an innovative approach to engaging a diverse group of stakeholders - students, educators, and leaders - in the collaborative development of ethical guidelines. By foregrounding DD, the study underscores the potential of such methodologies to foster a reflective, consensus-driven, and contextually informed set of ethical principles. This not only advances the discourse on AAI-EdTech - and more generally AIEd - ethics but also models a process that values stakeholder engagement, diverse perspectives, and collective decision-making. Through this, Swist et al. contributed to the understanding of ethical co-production in education, advocating for a more deliberative, inclusive, and sustainable approach to integrating AI technologies in educational contexts.

Milana et al. (2024) explored the impact of AI conversational agents and generative AI on adult education practice and research. Their paper highlighted both the transformative potential and the ethical challenges inherent in leveraging these technologies for adult learning. By examining the advancements in AI technologies, such as Machine Learning (ML), Generative AI, and Large Language Models (LLMs), the authors advocate for the development of AI literacy and capabilities among adults, emphasising the need for competencies that enable critical engagement with AI technologies across various life domains. The paper calls for a nuanced understanding of AI's benefits and limitations, urging adult educators and learners to navigate the AI era with informed scepticism and ethical vigilance.

The subject of ethics in AI has garnered attention in a broader context as well, and at the turn of the decade the European Union conducted the report "The Ethics of Artificial Intelligence: Issues and Initiatives," compiled by Bird et al. (2020). It scrutinised the ethical dilemmas and moral questions that AI presents, covering a broad spectrum of concerns from societal impacts, such as employment and inequality, to privacy issues, bias, and the potential for increased surveillance. The report emphasised the uneven effects of AI across different demographics, highlighting the risk of exacerbating existing inequalities and introducing new forms of bias. Furthermore, it explores the implications of AI on human rights, particularly concerning privacy and freedom of expression, and the potential for AI technologies to be used for surveillance and control by governments or other powerful entities. The document not only identifies problems but also reviews a wide array of international ethical guidelines and frameworks designed to address these challenges. It presents a comparative analysis of these frameworks, revealing gaps in addressing issues like fair benefit-sharing, responsibility allocation, worker exploitation, and the environmental impact of AI technologies. The report calls for thoughtful government regulation, the development of norms for acceptable uses of AI, and emphasises the necessity for AI systems to be transparent, accountable, fair, and under human control. It advocates for a proactive approach in shaping AI policies to ensure that the benefits of AI technologies are equitably shared and that their deployment respects human dignity and rights, ultimately contributing positively to society and democracy.

The bias of AI systems has already been experienced in the real world. For instance, in their article, Gurupur and Wan (2020) examined the inherent bias in AI-based healthcare decision support systems, attributing it to information bias, selection bias, and flaws in experimental design. They argue that these biases compromise the effectiveness of AI systems by affecting the accuracy of healthcare recommendations. The paper emphasises the importance of reliable information and careful methodological approaches in AI system development to mitigate these biases. It highlights the role of hindsight bias and the dangers of unreliable information sources in skewing AI predictions. Additionally, Leavy (2018) highlighted the issue of gender bias in AI systems, emphasising the risk of perpetuating societal inequalities through technology predominantly designed by men. These biases materialise, for example, as text corpora that the AI systems use for training are overrepresented by men or unequal in their descriptions of the genders. Leavy argues that to combat gender bias in AI, it is essential to incorporate diverse perspectives and gender theory into machine learning, especially algorithms trained on textual data that inherently carry gender ideologies. Highlighting the importance of diversity among AI developers, Leavy points out that female researchers are often at the forefront of recognizing and addressing bias in AI, suggesting that increased gender diversity in the field could help identify and mitigate biases more effectively. The paper reviews several ways gender bias manifests in language and suggests computational methods to identify and remove such biases from training data, advocating for a proactive approach to ensuring AI technologies do not reinforce outdated gender stereotypes.

DeCamp and Lindvall (2020) discuss the challenge of latent biases in AI algorithms in healthcare, emphasising the risks these biases pose to patient safety and the equity of care. They outline three main concerns: AI models developing biases over time, the influence of pre-existing clinical biases on AI decision-making, and the selection of outcomes or goals that may not align with patient or community interests. To mitigate these risks, the authors suggest proactive bias monitoring, regulatory oversight that includes bias detection, and stakeholder engagement in AI development and implementation. They advocate for viewing latent biases as a patient safety issue, necessitating pre-emptive action to prevent bias-related harm in clinical settings. The suggested actions against biases can also be utilised in other settings, such as education, as the same demographics exist on the societal level. In a similarly broad point of view, Trotta et al. (2023) conducted an examination on the field of AI ethics, underscoring the imperative for a multidisciplinary framework that integrates ethics, philosophy, and sociology to guide the development and application of AI technologies. Through a comprehensive analysis of various contributions, the authors highlight the ethical challenges and opportunities presented by AI's increasing pervasion in societal domains including healthcare, education, and employment. They advocate for the incorporation of ethical principles that prioritise transparency, accountability, and human values in AI systems to mitigate potential biases and ensure equitable benefits. Trotta et al. emphasise the necessity of global cooperation and public interest-driven approaches in shaping AI advancements, proposing future research directions that focus on fairness, transparency, and the avoidance of unintended consequences. The work of Trotta et al. mirrors the findings within the field of AIEd discussed earlier in this chapter.

# 3. Methodology

This thesis employs a qualitative approach, integrating a literature review with qualitative interviews. By combining these methods, the study endeavours to address the research questions through a betrothal of theoretical insights and practical perspectives. While the literature review establishes a theoretical foundation by exploring the potential opportunities of AI tools in programming education, the qualitative interviews offer firsthand insights from educators. These interviews serve to corroborate, challenge, or expand upon the findings of the literature review, enriching the analysis with real-world experiences and perspectives.

## 3.1 Research Problem and Questions

The research problem addressed in this thesis is grounded in contemporary trends. With the rapid rise of ChatGPT and its contemporaries into the mainstream, there has been significant interest in AI and the renowned tools that harness it. However, the impact of these tools on learning remains largely undocumented, leaving their effects on education unclear. This thesis specifically investigates how these newly popular AI tools influence students' learning, particularly in the context of programming education. Thus, the research problem is refined into how AI tools affect learning in the realm of teaching programming, which the three research questions aim to shed light on.

To recap, the research questions are as follows:

> **RQ1: Have effective strategies emerged in using AI tools in teaching programming?**
>
> **RQ2: What are the opportunities for teachers in bringing AI tools to the programming classroom?**
>
> **RQ3: What are the challenges for teachers in for teachers in bringing AI tools to the programming classroom?**

The first research question seeks to determine whether established strategies for integrating AI tools into teaching programming already exist. This investigation involves reviewing existing literature to identify any published research on this topic, given the relatively early stage of the field. The interviews with teachers serve as a supplementary source to explore this question further, providing insights into potential successful strategies that have been implemented in integrating AI tools into their courses. Through both literature review and teacher interviews, this research aims to ascertain the current landscape of proven strategies for incorporating AI tools in programming education.

The second research question places greater emphasis on teacher perspectives, drawing from their expert opinions and daily observations in teaching. Through analysis of interview data, this research aims to identify trends in teacher perspectives on the opportunities presented by AI tools. Complementing the interview data is information gathered from the literature review, which serves to either corroborate or challenge the insights obtained from interviews. Similarly, the third research question shifts focus to identifying challenges associated with the integration of AI tools in programming education. This analysis again relies on interview data to capture teacher perspectives on these challenges, which are then compared and contrasted with findings from the literature review. Together, the interviews and literature review provide a comprehensive understanding of both the opportunities and challenges related to the use of AI tools in teaching programming.

## 3.2 Literature Review

The literature review was conducted using the work of Kitchenham et al. (2009) as a guideline, although a notable difference to their work is that only the author participated in selecting the material. The search for papers from different conference proceedings and scientific journals was done manually, using the databases listed in Table 1.

**Table 1.** Selected Online Databases for Literature Retrieval

| Database |
|---|
| ACM Digital Library |
| ERIC (Ebsco) |
| Google Scholar |
| IEEE Xplore |
| Scopus |

The full search strings that were used can be found in Appendix A, while the most commonly occurring individual words can be seen in the word cloud of Figure 2. The search was conducted following the three separate themes that can also be observed in the Prior Research section of the thesis: Programming Education, AIEd, Artificial Intelligence in Programming Education and Ethical Implications in AIEd. The papers were then included or excluded based on an analysis of their title and abstract, and whether they either matched one of the three themes and offered recent insight (from the 19th century) or were crucial in explaining a core element of the theme in question.

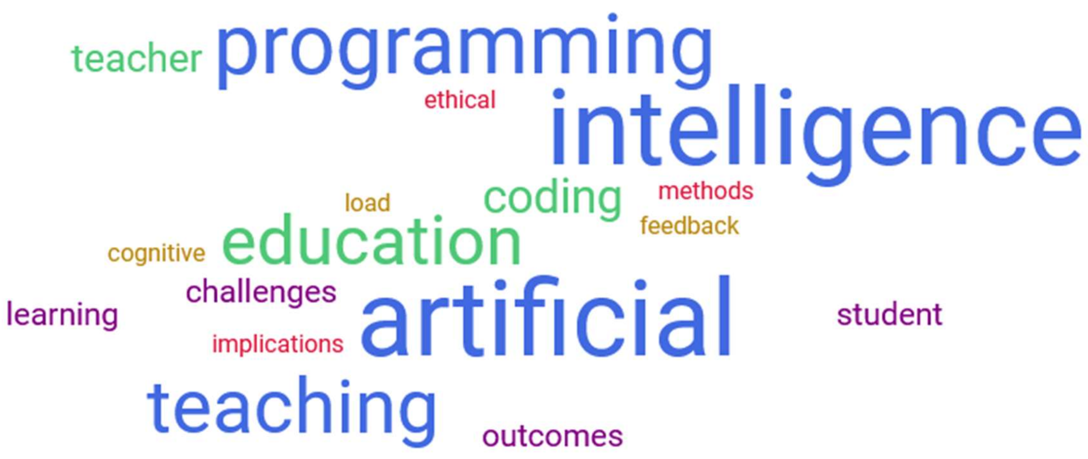


**Figure 2.** Key Search Strings Used in Literature Review

Additionally, for all papers, whether excluded or included, “snowballing” was applied. The term refers to going through the references of references - or in the case of this thesis, references of papers that would not quite make it to references - in an effort to find more useful references (Greenhalgh & Peacock, 2005). The initial inclusion criteria for the first three themes required papers to be published in scientific journals, as part of conference proceedings or a part of a relevant book. The selected papers were then examined to determine if they addressed topics related to teaching programming, AIEd, the use of AI tools or ITSs in programming education or closely related subjects, such as cognitive load. If they did not meet these criteria, they were excluded. Notably, mapping studies and literature reviews were extensively utilised in this thesis, following the rationale that they offer valuable insights into a field without necessitating a complete investigation from scratch, as advocated by Kitchenham et al. (2011). This approach was particularly beneficial given the novelty of AI tools in education, providing a comprehensive overview of the field from multiple different vantage points. Additionally, the fourth theme included scientific publications from organizations such as the EU or UNESCO, as the subject matter extends beyond academia and has broader societal implications. Exclusion criteria for all four themes was studies not written in English.

The literature review aims to answer the first research question by finding optimal practices for programming education and AIEd, and where these two subjects have crossover, if any. The second research question is also kept in mind, as even though there might not be much evidence yet in the way of proven practices for implementing AI in programming education, there might be theories on the subject already. Lastly, the challenges will be examined through the lens of ethics in AIEd settings.

## 3.3 Interviews

The qualitative interviews serve as a method for addressing the research questions, aiming to gather firsthand perspectives and experiences from programming educators regarding the integration of AI tools into their teaching practices, including any identified opportunities and challenges. The main focus of the interviews was to answer the second and third research questions, and to find out the current perspective of programming teachers on the challenges and opportunities that come with AI in teaching programming. However, if data arose to answer the first research question, bringing forth effective practices that have already emerged with utilising AI in teaching programming, they were welcomed. These interviews involved six teachers affiliated with University of Oulu, and one affiliated with University of Turku, who instruct classes covering diverse levels and technologies within the programming domain. The selection of interviewees focused on teachers actively engaged in programming courses or those teaching subjects closely related to programming, such as software security, information systems, or virtualisation. For the Oulu teachers, the selections were further guided by the Study Guide of the Information Processing Science programme of the Faculty of Information Technology and Electrical Engineering, in addition to recommendations from the Programme Director of the same programme. In the case of the teacher from University of Turku, the interviewee was selected based on the recommendations from a HR specialist and a professor of software engineering.

The intention was to ensure representation from educators with direct involvement in programming-related materials. Importantly, participants were chosen irrespective of their prior enthusiasm or knowledge regarding the utilisation of AI tools in programming education. This approach aimed to capture a broad spectrum of perspectives and experiences among programming educators.

The interviews were conducted following the methodological framework outlined by Myers and Newman (2007), particularly the principles of semi-structured interviews, as a semi-structured approach will ensure flexibility in exploring the topics of the thesis while maintaining a systematic approach to data collection. The structured portion of the interviews, consisting of prepared questions, drew inspiration from the approach employed by Wardat et al. (2024). Their utilisation of a questionnaire to assess mathematics teachers' perspectives on AI provided valuable insights and established a solid foundation for the formulation of interview questions in this study. By leveraging their methodology, this study aimed to build upon existing research while tailoring the questions to the specific context of programming education and the integration of AI tools.

The interview questions found in Appendix B were designed to elicit detailed responses about the opportunities and challenges faced by teachers, while also leaving room for more open-ended discussion. The collected qualitative data from the interviews then underwent thematic analysis. The analysis was both motivated and guided by the work done by Braun & Clarke (2006), as they noted that thematic analysis can help produce "a rich and detailed, yet complex, account of data" while providing flexibility as a tool. This process involved coding segments of the interview transcripts based on common topics or issues raised by the participants.

The smaller sample size of interviewees (n=7) has both advantages and drawbacks. As Marshall (1996) pointed out, a smaller sample size, or a "key informant technique," can yield quality data in a short amount of time, although it may not be as universally representative. As the sample size was small, quantitative analysis was not done on the data. The initial themes were challenges in integrating AI tools to programming education and opportunities in integrating AI tools to programming education. However, if more themes emerged that had not been considered prior, they were identified and coded accordingly.

The coding of the interview data involved examining the transcripts from the interviews. The process drew inspiration from the approach outlined by Cruzes and Dybå (2011), as the transcripts were first divided into quote objects, followed by labelling the objects and translating them into a categorical theme. Each quote object within the transcripts was categorised into one of the two predefined categories based on the research questions. If a quote object did not align with either category, a new category was identified, or the text was excluded from the analysis.

All participants took part in the interviews under informed consent, as all participants were briefed on their rights and the study's procedures. Detailed information regarding participant rights and data processing protocols can be found in Appendix C.

# 4. Results

This chapter offers an exploration of the perspectives and opinions gained from interviews conducted with programming teachers. As outlined in the methodology chapter, the primary focus was on uncovering insights regarding the opportunities and challenges associated with integrating AI tools into programming education. However, beyond these predetermined themes, any additional emerging themes discovered during the analysis are also presented and categorized herein. Following an examination of the interview transcripts, both the predetermined and newly discovered themes are listed below.

**Table 2.** All identified categorical themes within the interview data.

| Theme | Occurrences |
|---|---|
| Benefits of AI tools in teaching programming | 45 |
| Challenges of AI tools in teaching programming | 42 |
| Teacher Perspectives on the Effective Integration of AI Tools | 28 |
| Impact of AI tools on learning | 12 |
| Effective general learning practices | 10 |
| Other Benefits of AI tools | 8 |
| Other Challenges of AI tools | 4 |

As seen in the Table 2, several additional themes surfaced alongside the two primary ones initially defined prior to the interviews.

## 4.1 Opportunities with AI tools

This chapter goes over the data points from the interviews that were categorised to fall under potential benefits of integrating AI tools in programming education.

A significant emerging theme within the "opportunities" category was accessibility highlighted by the interviewees regarding the potential of AI tools as learning support resources. This motif was identified in twenty-one out of the forty-five instances within the overarching theme. Accessibility manifests in various forms, with a recurring example being the availability of AI tools like ChatGPT round the clock for students. This accessibility ensures that students are not reliant on the availability of TAs to address challenges they encounter, empowering them to tackle obstacles independently at any hour of the day.

> *"When students know how to use these tools properly, they can have let's say a teacher 24/7 so they can ask questions and and that too can*

> *provide… I mean I don't think that right now or the current status they are going to substitute teachers. In the end there are issues that are a little bit more complicated that perhaps these tools they are not going to handle correctly. But I think that what the students can get is a kind of TAs 24/7 for them."*

Another avenue of accessibility that was brought up multiple times was the dynamic way that, for example, generative AI tools can relay information to their user. For instance, if a student were to have trouble understanding an assignment, they could refer to an AI assistant to try and help the student understand the assignment, by way of rewording or even visualising the assignment. This sentiment is embodied by the following quote from one of the interviews:

> *"So in that sense, you could, like, I took the scenario as an example, that the student doesn't ask the AI to give me the binary search but explain to me how it works, so, like, this kind of interactive dialogue with the AI, if the AI can, like, in a way that the student can understand, explain the matter, the student can request, for example, a picture or animation or, like, some other kind of, like, different explanations and perspectives on how the thing works." (Translated from Finnish)*

Illustrating concepts and even snippets of code were identified as potential learning enhancers as students gain the ability to tailor the material into not only a style of their liking, but also to the benefit of their own understanding. This was also identified as potentially useful for beginner programmers, as error messages can be quite cryptic to novices who do not have experience in reading stack traces and such.

On the other hand, while opportunities for beginner programmers were identified, some responses underlined that the best outcomes would come to those who already have experience with programming, who possess a foundational understanding of the underlying concepts. This was identified in six occurrences in the interview data. These individuals can effectively delegate mechanical, repetitive tasks to AI tools, allowing them to concentrate on problem-solving aspects.

> *"But of course, then if we use AI-type tools, well, then the task itself becomes more complex and there's more like this kind of reflection and creativity required in the assignment, it's not just mechanical anymore, and that's actually good, because, like, a human coder's work is rarely very mechanical anymore, it's more problem-solving and, and like optimization, so it's perhaps not so much anymore about writing those loops by hand. It better prepares for the work life." (Translated from Finnish)*

In the above comment, the interviewee added some reflection on how programming assignments of the future might need to evolve to focus more on creativity and reflection. It should, however, also be mentioned that while most interviewees who identified the positive yields of AI tools in the hands of more experienced students as substantial, some speculated the effect - even when positive - to be minor at best.

Another advantageous application of AI tools identified in the interviews was the ability to generate content. The ability to generate content was found be of use to both students and teachers, although in different methods of application. For students, a generative AI tool could serve as an external source of requirements for a programming project, thus guiding a process that more closely mirrors real-life projects, alternative to the scenario where students devise their own requirements. The interviewees noted several different applications for teachers, such as generating model answers and creating variations on assignments.

> *"But yeah, in many ways, one thing I've personally found, and what we sort of collectively acknowledged, where they're really useful in terms of education, is in making task variations, like in programming. Now, I notice that if I have, say, three specific tasks in an exam, and I need to quickly come up with a new exam that measures the same thing, well, like, a basic ChatGPT or CoPilot or something easily generates variations of those tasks, asking the same thing, but with a different theme." (Translated from Finnish)*

As illustrated by the above comment, teachers can make variations on an assignment without altering the learning objectives, meaning that students can all have their own assignment, and on the other hand, students will not be able to share their solutions to each other and pass by cheating.

Three interviewees also pondered on the potential of AI to enable a more personalised learning experience for students, ranging from individually crafted paths through university education to helping report individual students' progress or difficulties to the teacher. Personalised learning experiences are delineated as a distinct category from the accessibility aspect discussed earlier. In the context of the results of this thesis, personalised learning refers to a comprehensive approach involving customised assignments tailored to individual students' learning styles and needs, aiming to guide each student towards the specific learning objectives of each course in the most efficient manner possible, taking into account the unique characteristics of each student. With these answers, however, there seemed to exist a veiled notion that the AI tools of today are probably not yet quite at the level of providing personalised learning experiences for students.

> *"Yeah, I don't know if we are close to that or not, but I think yes, that if this AI is able to identify the needs of a student, so to compare for example the current status of the student, skills and the goals of a course, it would be easy to find."*

During a small portion of the interviews, ITS-type applications of AI were also acknowledged, with four occurrences related to this theme. The potential for real-time analytics of student progress, coupled with AI providing helpful suggestions, was recognised. However, it was noted that achieving this might be challenging due to potential issues with data protection and student perceptions regarding constant monitoring required by such a system.

> *"Some kind of data collection and analysis, if it could be realized that someone is falling behind or becoming passive, or something similar, so if there was a tool that could somehow collect this data, then it would be like bringing it up." (Translated from Finnish)*

An underlying understanding that was ever-present even with the answers regarding the opportunities of AI tools in teaching programming was the importance of ensuring that students do not merely outsource their entire learning experience to AI. Rather, it is crucial for students to ensure that these tools genuinely support their learning process. This topic will be further examined in the subsequent chapter.

## 4.2 Challenges with AI tools

The predominant challenge associated with the integration of AI tools into programming education, as highlighted by the interviewees, revolved around the concern that relying on such tools for support might lead to an excessive delegation of the learning process, resulting in students not genuinely learning core programming concepts. This subcategory was identified in twenty out of the forty-two instances within the broader theme of challenges. Particularly for introductory programming courses, where language models and tools like ChatGPT excel in generating functional solutions for simple assignments, there was considerable concern that students could potentially pass an entire introductory course by solely depending on an AI tool. Following are three quotes from three different interviews on the topic.

> *"I'm worried about those who blindly trust it and think that now that I've done this task with the help of ChatGPT, now I know it because I got through the task, and the point has completely missed them. And most likely, the code doesn't even work." (Translated from Finnish)*

> *"So, in a way, there might be a risk of relying on the fact that this AI does these tasks for me every time. But then if you have a specific enough problem for which there is no data available, for example, where the AI has been trained, suddenly you don't have a solution." (Translated from Finnish)*

> *"So yeah, sure, you might get through the courses with these tools or something, but you're not really learning anything, it will catch up with you at some point, and fortunately, it probably happens quite often at some point in your studies, not just in the workplace." (Translated from Finnish)*

Another prevalent issue within the challenges theme, found in ten occurrences, was the occasional fallibility of AI tools, such as chatbots based on large language models, leading to errors in their responses. It is crucial for programming students to recognize the possibility of inaccuracies and other mistakes that these tools can generate, rather than viewing them as infallible. Given that the solutions produced by language models can appear convincing even to experienced programmers despite being incorrect, novice programmers, such as students, must remain particularly mindful of the pitfalls of AI tools.

> *"And even if it... even if it gave you a good answer - I experienced this just a couple of weeks ago last time that... I was trying to write a piece of code*

> *that was pretty clear in my mind, and I thought I'd ask ChatGPT like how to implement this thing, give me some starting code and so on. It looked good, it worked to some extent, and then there was a bug, and I tried to hunt it down, and it ended up being in the code provided by ChatGPT, but it sounded so good that it took me 2 days before I found it." (Translated from Finnish)*

Plagiarism surfaced as a concern in five instances during the interviews. It was noted that programming teachers may also find themselves tasked with distinguishing between AI-generated code and student-generated code, highlighting the additional time and resources this demands.

> *"And if we now consider the effort, if we go back to this question, whether it's more laborious or not, then it's kind of an additional aspect to the teacher's work. When you're pondering whether this thing is self-made or not, we're wondering if it's stolen from friends, from Stack Overflow, from some website, or from AI, so it's kind of one more consideration added to it when you suspect that this can't be a self-made solution." (Translated from Finnish)*

Some interviewees brought up the point that as the AI tools are capable of producing believable deliverables for programming tasks, perhaps the evaluation of students should evolve with the tools. The potential need for a new way of evaluating programming assignments was identified in three occurrences.

> *"I mean when you give a test to the students, you need to be very explicit on what they need to do. So, this is also very valuable information for these type of language models. So, they are going to produce a good solution. So perhaps what we need to try to think is what we are assessing perhaps doesn't make sense if we want to know if the students have basic knowledge on programming."*

Additionally, isolated instances raised other concerns, including the prospective role of future programming teachers needing to incorporate quality assurance measures if materials were generated using an AI tool. Furthermore, there were reflections on students not being adequately trained to effectively utilize AI in a constructive manner or the potential decrease in motivation upon observing an AI tool outperforming students in simple programming tasks.

## 4.3 Teacher perspectives on the effective integration AI tools

The majority of instances within this category revolved around mitigating overreliance on and the delegation of actual learning to AI tools. Out of the twenty-eight instances under the overarching theme, nineteen data points were related to comprehending the limitations of the tools in use and being fully aware of their functionality and purpose. This includes recognising their weaknesses, such as the aforementioned hallucinations, and understanding that even when they sound convincing, they might produce absolute nonsense. Four quotes on the topic from four different interviews

> *"As long as one understands what it is capable of, what can be done with it, and then knows how to act in various situations where it then starts to produce something nonsensical or does something downright crazy, so*

*that you know how to start solving the problem from there." (Translated from Finnish)*

*"Perhaps it's more about those who use it needing to understand its limitations and potential areas where it can be helpful, and to what extent. So, those kinds of boundaries." (Translated from Finnish)*

*"At some point, perhaps more emphasis could be placed on students being critical of the answers they receive, not accepting them at face value." (Translated from Finnish)*

*"I think that the important thing is to teach students how to use correctly and of course you need to come up with because I mean in my opinion it's impossible to forbid them."*

Another eight instances emphasized the importance of students recognizing that the tools are intended to support learning, rather than completely appropriating the learning process. Additionally, the personal responsibility of students and the importance of taking pride in acquired knowledge were highlighted. Three examples below, all from different interviews.

*"It is, however, what you will certainly use in your future work, but at the same time, you must ensure that self-learning takes place, that you understand the ideas behind it, such as what the courses are meant to teach, so that you are not just satisfied with passing a course but have gained something from it. Let's try to be ambitious for ourselves, and it will definitely carry us far in life." (Translated from Finnish)*

*"So, one must have problem-solving skills oneself, and understanding whether the solution is really what was sought. And that's why I still emphasize that in those studies, you really need to internalize what you're doing, even if in practice you mostly use that AI tool in everyday life." (Translated from Finnish)*

*"Ask a fellow student, ask ChatGPT, ask some other LLM, but try to acquire the knowledge and understanding with the help of a reliable resource." (Translated from Finnish)*

In addition to the aforementioned instances, one interviewee provided a concrete example of an effective strategy in using AI tools, namely, offering a link to the complete exchange of prompts and AI-generated responses that the student and the tool had undergone.

*"My guidance is such that if an AI tool has been used, then a link is provided to the conversations held with it, as you can get these links directly from the tools, so then the reviewer can directly check how it has been used." (Translated from Finnish)*

This transparent approach can assist the student in providing evidence of responsible use of the AI tool and can also aid the teacher in rectifying any erroneous usage of the tools, should it occur.

## 4.4 Impact of AI tools on learning

This chapter will outline the interviewee views on how AI tools have already affected student learning, as well as how they envisioned the tools affecting student learning in the future.

This theme did not have as many data points as the previous ones, but the instances are of significant importance as it is one of the central themes of the thesis. A consistent viewpoint evident across many responses from various interviews was that there has not been much of a noticeable impact on student learning as of now, or at most, a slightly negative one.

> *"Yeah, I think they have worsened it at this stage. Because again, we, I think that the big mistake that we have done is not tell students how to use it correctly and efficiently"*

However, since there is not all that much data available yet, the instances of this category remain, as admitted by multiple interviewees, largely speculative.

> *"Still, there isn't data, but I would guess that they have slightly weakened [the learning outcomes]." (Translated from Finnish)*

## 4.5 Effective general learning practices

This chapter highlights several instances of good practices that manifest at a broader level than the examples from Chapter 4.3 and can be applied to almost any course, not limited to programming courses.

The data points in this category largely revisited familiar subjects from earlier chapters. However, reiterated and reinforced here is the notion that despite the seductive ease of prompting a generative AI for answers, students must still learn the material covered in the courses themselves. The interviewees brought up the need for a personal interest and pride in your studies, as they too can be powerful weapons in a student's studies.

> *"It should be accepted that learning is also partly a painful process in this sense, that it requires effort and struggle, that you learn, and if you avoid it by using any tool that is intended only to avoid that pain, then you won't learn. So, learning just doesn't succeed without a certain kind of effort, it shouldn't be avoided." (Translated from Finnish)*

> *"Because, in a way, going through and repeating things time and again builds a profound expertise and understanding, and if you kind of outsource that, then how can that profound personal expertise and understanding ever grow?" (Translated from Finnish)*

Another point that was highlighted in this category, although to a lesser extent, was the significance of support for students from the teachers' side, and the notion that students should make more use of assistance from this direction. Although resources may be limited, students should keep in mind that the teachers are there to assist their learning.

> *"Firstly, in collaboration with those other means, usually it's the teacher, it could be a textbook, it could be a good website, anything, but it's like source criticism and in a certain sense, reaching out to knowledgeable*

*individuals who are interested in your own learning - that is, teachers." (Translated from Finnish)*

*"In my opinion, it's important that there is kind of a lot of help available." (Translated from Finnish)*

## 4.6 Other Benefits of AI tools

This chapter will go over the benefits identified in the interview data that do not necessarily relate to programming education but can be applied to teaching and studying on a more general level.

Half of the instances in this category were associated with the potential of translating and spellchecking texts written in a language not native to the interviewee. Given that universities often comprise multinational environments within themselves, not to mention their collaboration with other institutions, having a tool to enhance interlingual communication was recognised as an asset.

*"It's saving me a lot of time and it's allowing me to do things that that I cannot do otherwise. For example, I'm using quite a lot ChatGPT for translating into Finnish. Although it's not perfect, I think that the results are quite good. So many times if I have, I have some students that they prefer Finnish and I'm not very fluent at Finnish, so I use this ChatGPT, for instance, to translate them the message into Finnish. So in that sense, I think that it has improved the communication with the students."*

*"Just some guidelines, for instance, when writing yourself, often I proofread them using ChatGPT, and it's able to correct those grammar errors quite well and sometimes gives some ideas on how to clarify, so I do use it, yes." (Translated from Finnish)*

Other advantageous applications of AI tools that surfaced in this category involve employing an AI tool as a creative assistant. Exchanging ideas back and forth with a tool can facilitate discovering things faster than one would alone.

*"In that kind of idea bouncing and getting additional ideas, which is what I also encourage, I think it's good, especially on that side, like when you're working on a thesis, for example, you have to narrow down a large topic area, so something like a ChatGPT-style tool fits very well for that." (Translated from Finnish)*

## 4.7 Other Challenges of AI tools

As the previous chapter went over the more general benefits, it is deemed necessary to also present the more general challenges that emerged in the interviews, as few as they might be.

All four instances in this category pertained to plagiarism, whether in programming related assignments or the traditional essays, and the future challenge of discerning between AI-generated text and student-generated text. Interviewees expressed concern about the additional burden this could place on teachers, particularly in courses with a high number of participants, where resources may be insufficient to thoroughly investigate every submitted assignment.

> *"And if it becomes widespread, then I think the problem will be that if you can't really trust that the work has been properly researched and sourced and referenced when it comes to sources, so with the number of students we have, the teacher can't do that work anymore." (Translated from Finnish)*

## 4.8 Summary of Results

Out of the 167 quote objects analysed and coded, the majority (58.4%) were associated with the initial two themes that were pre-defined and anticipated before the data analysis. Figure 3 illustrates that the newly identified themes comprised the minority, accounting for 41.6% of the data.

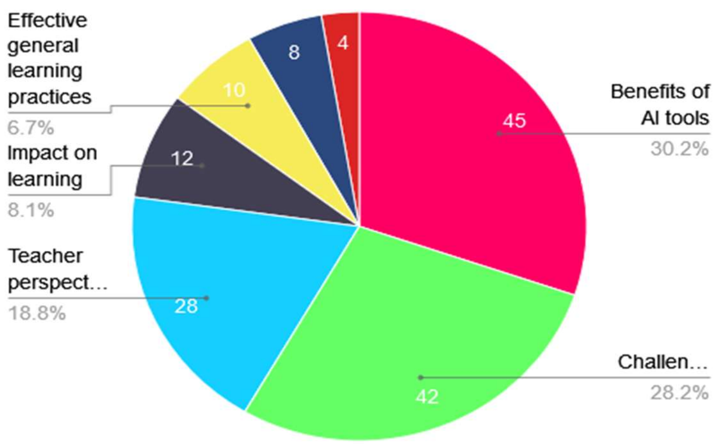


**Figure 3.** Distribution of Themes in Interview Data Analysis (created with https://piechartmaker.co)

Although the newly identified themes constituted the minority of the coded data, their consistent appearance across multiple interviews suggests the need for discussion. The similarity of data points within these categories across interviews indicates a trend in opinions, making it important to include them in the discussion section.

# 5. Discussion

This chapter will discuss the results of the interviews in relation to the research questions while mirroring them to the central findings of the literature review. Due to the limited sample size, this section will primarily highlight prominent trends observed in the interview data, which were recurrent across multiple interviews and exhibited consistency among participants. These findings will be corroborated by aligning them with the conclusions drawn from the literature review, thereby enhancing the credibility of the shared insights. This approach aims to underscore the recurrent themes that emerged from the interviews, validating their significance, and demonstrating their resonance with broader scholarly discourse in the field. In addition, the limitations of this thesis will be addressed and points for future work suggested.

## 5.1 RQ1

**RQ1: "Have effective strategies emerged in using AI tools in teaching programming?"**

The prominent strategy present in the interviews was – perhaps a more metacognitive strategy than a concrete how-to practice – the mindfulness of the strengths and weaknesses of artificial intelligence. This is realised by knowing what exactly it is what the AI tools do and understanding that while they can appear efficient in solving programming-related problems, it is imperative to keep in mind that at least as of yet, no AI tool is infallible. A similar sentiment was highlighted by both Hellas et al. (2023) and Van Slyke et al. (2023), as they underlined the need for educated utilisation of AI and its strengths and weaknesses, strengthening the importance of this prominent strategy found in the interviews.

Another crucial consideration that emerged in the interviews regarding this research question, though not exclusively tied to AI tools, is the notion that learning is an event that requires some level of effort, which means that students must resist the temptation to offload their work entirely to AI tools. This concept was reinforced not only by the interview data but also by findings in the literature. As mentioned by Caspersen & Bennedsen (2007), cognitive load is one of the key building blocks of learning, which implies that without utilising any of the student's own cognitive processing power, the actual learning is diminished, if not outright supplanted by the tool. Therefore, even if students can employ the tool in a master-apprentice dynamic akin to the effective learning methods described by Collins et al. (1989) and Vihavainen et al. (2011) (albeit with human mentors), it is imperative that the student themself remains the main character of the story in their learning experience, while the AI tool remains in a supporting role.

Another noteworthy point, though arising from only one interview, was a concrete strategy involving generative AI tools to provide a link to the dialogue between the student and the tool that led to the solution submitted by the student. This transparent approach enables students to demonstrate they are not engaging in plagiarism and offers teachers an opportunity to guide students in the proper use of the tool. Such an approach aligns with the vision of Saari et al. (2023) regarding the evolving role of programming teachers in the future. It also addresses concerns raised by Real-Fernández et al. (2019) and Van Slyke et al. (2023), who emphasize the importance of AI tools serving the learning process and the necessity of guiding students in their effective utilization. Furthermore, this approach would also quell the teacher concerns brought up by Lau &

Guo (2023), in which teachers remain in the dark as to how their students actually utilise AI tools.

## 5.2 RQ2

**RQ2: What are the opportunities for teachers in bringing AI tools to the programming classroom?**

The most frequently acknowledged advantage of AI tools in programming education in the interviews was the accessibility they offer to students. The interviewees recognised the potential for an AI tool to serve as a valuable resource in students' learning arsenal, akin to a teaching assistant available round-the-clock. Andersen et al. (2022) proposed this opportunity to potentially lighten the workload of teachers, although the interviewees pointed out that actual teachers should still be consulted as the primary source of help whenever they are available. Accessibility also surfaced in the provision of alternative presentations for materials, such as visualising data or assignment text. This utilisation of AI tools aligns with the shift towards a more interactive and responsive learning environment, as highlighted by Chen et al. (2020).

On a related note, Anfurrutia et al. (2018) highlighted the significance of formative feedback for students, suggesting that AI chatbots functioning as 24/7 teaching assistants could potentially offer considerable benefits. However, given that AI tools are still evolving, while they may be capable of providing feedback to students regarding their progress and issues, it might take some time before they can offer truly formative feedback. As noted by Hao et al. (2021), detailed and continuous feedback can be a potent asset in teaching programming to students, indicating a potentially fruitful avenue for the future within the context of the already existing chatbots.

Another beneficial opportunity that emerged in the interviews was the potential for generating material, whether it be teachers generating content for their classes or students using it as a creative tool in their studies. Interestingly, this subject was not directly addressed in the literature review. However, it was paralleled by the discussions of Andersen et al. (2022) and Saari et al. (2023) regarding the future roles of teachers and how they might evolve with these tools. It is worth noting that as this specific benefit of utilising AI tools was not identified before the interviews, it was also not actively searched for, potentially leaving a gap in the knowledge presented by this thesis.

The relatively limited data on the personalisation of learning from the interviews was somewhat unexpected, given the extensive literature highlighting the benefits of AI tools in delivering personalised learning experiences to students (Chassignol et al., 2018; Devedžić, 2004; Eilermann et al., 2023; Guan et al., 2020; Lim et al., 2023; Roll & Wylie, 2016; Tapalova & Zhiyenbayeva, 2022; Timms, 2016; Wang et al., 2023). However, this might also stem from the decision in this thesis to treat accessibility as a distinct thematic category from personalized learning, as the distinction between the two is not always straightforward and might vary from researcher to researcher. However, in the instances where personalised learning did come up, the notions of the interviewees were similar to those that arose from the literature. The interviewees noted the potential of AI tools being able to reflect student needs on an individual level, needs which Latulipe et al. (2018) pointed out can vary based on a number of immutable characteristics, differing with each student.

Similarly to the personalised learning category, ITSs were mentioned in a surprisingly small frequence. There could be a couple of potential reasons for this. First, the questionnaire did not directly prompt interviewees to discuss ITSs, instead allowing them to mention relevant topics organically. Secondly, the hot contemporary topic is the generative chatbot, meaning that it is, with a high likelihood, the first thing that comes to mind for most people when confronted with the term “AI tool”. Nonetheless, despite the limited number of occurrences in the interview data, the benefits discussed aligned with those identified in the literature review. These included the potential for personalised learning facilitated by automated, ongoing support resembling a TA-like system tailored to the specific requirements of each student (Anderson et al., 1985; Eilermann et al., 2023).

However, some interviewees suggested that the opportunities presented by AI tools would primarily benefit students who are already proficient in programming, indicating that novice programmers might be less likely to use these tools effectively. This aligns with the observations made by Hellas et al. (2023), who noted that a certain level of programming knowledge is necessary to discern useful output from, for example, a generative AI tool.

## 5.3 RQ3

**RQ3: What are the challenges for teachers in for teachers in bringing AI tools to the programming classroom?**

The prevalent concern identified in the interviews was that by using AI tools, students may delegate too much of the problem-solving process to the tool, thereby neglecting to learn the core concepts. Though it should be noted that as of yet, the interviewees did not perceive a significant reduction in results attributable to AI tools, as there is not much data, or even metrics, on the subject as of yet. Nevertheless, the interview findings reiterate the findings of Prather et al. (2023b), as their research pointed out that especially novice programmers are susceptible to delegating too much work to the tool and missing critical learning processes. The problem is exacerbated by the findings of Kazemitabaar et al. (2023) and Ouh et al. (2023), who noted that AI code generators are already proficient enough to solve introductory programming assignments. Adding to the challenge is the fact that novice programmers often struggle to develop mental models related to programming, unlike more experienced individuals (Robins et al., 2003), which means that the already complex process becomes even more difficult, with crucial knowledge building blocks being missed entirely. Additionally, the degree of adherence to the learning process can vary depending on the student's level of motivation and comfort in their environment, as emphasized by both Bergin et al. (2005) and Wilson & Shrock (2001).

Another commonly recurring concern within the interviews was the simple fact that AI tools are fallible and may produce factually incorrect results. This underscores the necessity for students to grasp the core concepts of programming to discern between the AI tool's hallucinations and useful output. This same necessity was identified by Hellas et al. (2023) and Prather et al. (2023b), as they too emphasised that understanding programming is crucial for distinguishing between useful and useless outputs produced by AI tools. Van Slyke et al. (2023) proposed a potential solution for this challenge by advocating for the active integration of these tools into educational settings, providing teachers with an opportunity to educate students on their efficient and correct use. However, one other potential challenge this task of differentiation brings is the fact that it adds to its own cognitive load. As Chandler & Sweller (1991) pointed out, students have a limited cognitive load capacity, which means that increasing it by deciphering the output of AI tools might be counterproductive, especially considering that the prevalent minimal guidance approaches of teaching programming (Robins et al., 2003) can already be cognitively demanding (Kirschner et al., 2006).

There was also concern about plagiarism, as some interviewees worried about the scenario of students presenting completely AI-generated code as their own. This concern was compounded by the worry that teachers would have to differentiate between AI-generated code and students' original code. Puryear and Sprint (2022) also highlighted this issue and proposed that new assessment methods would need to emerge to address this challenge. However, in addition to plagiarism-related challenges, ethical concerns were largely absent in the interview data. Once again, ethical considerations were not prompted by the questionnaire questions, which may explain this absence. Regardless, as Mouta et al. (2023) found, the ethical dimension is one largely unexplored in the educational sector when it comes to the implementations and frameworks of AIEd, which could help explain why it was absent in the interviews as well.

Additionally, one concern highlighted in the literature review but not identified during the interviews is the lack of focus on the teacher perspective in AIEd research (Çelik et al., 2022; Crompton & Burke, 2023). Furthermore, Ng et al. (2023) raised the issue that integrating AI tools into teaching can be complex due to potential lack of AI competency among teachers, hindering efficient and responsible utilisation of these tools.

## 5.4 Limitations

A potential limitation in the research is that some of the studies, even those only two years old, may provide perspectives that are now outdated. Given the rapid – especially recently, with the advent of ChatGPT - pace of advancements in AI technology, views, and understandings from just a few years ago may not fully reflect the current landscape. In a similar vein, some of the research focusing on prevalent teaching practices dated back almost 20 years, meaning that while the data might have been historically accurate, the latest practices might have flown under the radar of this thesis.

Another set of limitations stems from the interview methodology. With only seven participants, the interviews offer a limited perspective, not fully representing either a Finnish or global viewpoint comprehensively. Moreover, since the coding process was conducted by a single individual, inherent biases may have influenced the results, potentially leading to interpretations that differ from those of other researchers. This concern is compounded by the challenge of distinguishing between certain categories, such as the subcategories of accessibility and personalised learning within the broader "Opportunities with AI tools" category. These distinctions can be subjective and open to interpretation, resulting in varying categorization decisions across different coders.

Additionally, since the translations were not performed by professional translators but rather by the author of the thesis with the aid of AI tools ChatGPT and DeepL Translator, there is a possibility that some contextual nuances were lost in translation, failing to accurately and completely communicate the nuances of the Finnish-language interviews.

## 5.5 Future Work

In this thesis, various potential opportunities and challenges related to the effective integration of AI tools into programming education were explored. However, the findings primarily rely on theoretical expert opinions from teachers, lacking quantitative data to support them. Consequently, there is a need for a quantitative approach to address the research questions posed in this thesis, likely involving longitudinal measurement of student performance with the support of AI tools over an extended period. In addition, the interviews highlighted a noteworthy aspect worth exploring further: the potential for AI tools to generate content, a capability recognised for both teachers and students. However, it is important to note that the literature review in this study did not extensively cover this particular aspect, which means that it is entirely possible that there is already existing research on this topic.

Another aspect to consider is how to effectively educate students on the responsible use of AI tools, and where the responsibility for this education lies— with the teacher or the educational institution. Specifically, there is the question of whether to dedicate an entire course to the subject, include a class on it within existing courses, or provide students with basic instructions and let them navigate on their own. Given the nascent nature of the subject, these methods are still in the experimental phase, and no standardised practice has yet been established.

As highlighted in the literature review and further emphasised in the discussion of the interview findings, the ethical aspects of AIEd are unfortunately not receiving adequate attention, with both research and teachers focusing on other aspects of AI tools. It is essential that the ethical dimensions of AI in education be given consideration, at the very least in the development of guidelines if not in further research, to ensure a balanced approach that promotes fair and equitable education alongside efficient teaching delivery.

# 6. Conclusion

This thesis presents qualitative research on the challenges and opportunities associated with integrating AI tools in the programming classroom, using a method of semi-structured interviews with programming teachers. The research questions of the thesis aimed to answer whether proven strategies with AI tools have already emerged and whether teachers perceive notable opportunities and challenges in making them into one of the tools in the arsenal of programming students. The central finding from the interviews underscores the significance of students employing AI tools with a thoughtful and educated approach, which manifested in numerous ways. Firstly, the metacognitive strategy of mindful tool utilisation emerged as pivotal. Secondly, the most notable challenge identified in the interviews regarded students failing to engage with the tools mindfully and responsibly, risking a lack of substantial learning by relying solely on the answers of the tool, or simply not understanding them. Opportunity-wise, the most important finding was the ability to make teaching more accessible, as AI tools can act as a substitute for TAs at any time or place, who can help students in figuring out assignments and – in a perfect world – even assist with the learning process.

The research method chosen for this thesis was motivated by the broad nature of the research questions. Given the expansive scope of the research questions, it was necessary to comprehensively explore the significant opportunities and challenges posed by AI tools in programming education. Would the thesis have focused on specific aspects without this prior knowledge, they would have to have been selected at random. More extensive interviews across multiple institutions could have enriched the study; however, practical constraints such as time, space, and a graduation deadline limited the scope. The limited scope was also considered, as the key findings presented in this study align with the prevalent themes that emerged from the interviews, prioritising common trends over isolated or infrequent data points.

Further research is necessary to accumulate a qualitative corpus of data on the topics addressed in this thesis, particularly because the impact of AI tools on learning is a new area of study, and comprehensive measurements of their effects have not yet been undertaken. Additionally, future research could extend the findings of this thesis by identifying specific performance-related metrics that could offer a different perspective on addressing the research problem. This approach would deepen our understanding of how AI tools influence student learning outcomes and inform more targeted educational strategies. Furthermore, this thesis identified areas where expert opinions were not as focused, which could create a knowledge gap if left unaddressed. Specifically, the ethical implications of AIEd deserve close attention to ensure responsible and effective implementation of these technologies.

This thesis endeavoured to investigate how AI tools impact the learning experience within the realm of teaching programming. The contributions of this study provide insights for more focused research in the future by uncovering expert opinions from programming teacher interviews. Through these interviews, the study identified significant challenges and opportunities that warrant further investigation. Additionally, the strategies in utilising AI tools in programming studies that emerged from this research should be tested and validated in real-world educational settings.

# References

Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Computers and Education. Artificial Intelligence*, *4*, 100131. https://doi.org/10.1016/j.caeai.2023.100131

Ahmadzadeh, M., Elliman, D., & Higgins, C. (2005). *An analysis of patterns of debugging among novice computer science students*. https://doi.org/10.1145/1067445.1067472

Akgün, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI And Ethics*, *2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Andersen, R., Gjølstad, E., & Mørch, A. (2022). Integrating Human-Centered Artificial Intelligence in Programming Practices to Reduce Teachers' Workload. *CEUR Workshop Proceedings*, 3136, 30–35.

Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. *Science*, *228*(4698), 456–462. https://doi.org/10.1126/science.228.4698.456

Anfurrutia, F. I., Álvarez, A., & López-Gil, J. (2018). Integrating formative feedback in introductory programming modules. *IEEE-RITA*, *13*(1), 3–10. https://doi.org/10.1109/rita.2018.2801898

Banić, B., Konecki, M., & Konecki, M. (2023). *Pair Programming Education Aided by ChatGPT*. https://doi.org/10.23919/mipro57284.2023.10159727

Bergin, S., Reilly, R. G., & Traynor, D. (2005). *Examining the role of self-regulated learning on introductory programming performance*. https://doi.org/10.1145/1089786.1089794

Bird, E., Fox-Skelly, J., Larbey, R., Jenner, N., Weitkamp, E., & Winfield, A. (2020). The ethics of artificial intelligence: Issues and initiatives. European Parliament. https://data.europa.eu/doi/10.2861/6644

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Caspersen, M. E., & Bennedsen, J. (2007). *Instructional design of a programming course*. https://doi.org/10.1145/1288580.1288595

Çelik, İ., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The Promises and Challenges of Artificial Intelligence for Teachers: a Systematic Review of Research. *TechTrends*, *66*(4), 616–630. https://doi.org/10.1007/s11528-022-00715-y

Chandler, P., & Sweller, J. (1991). Cognitive load Theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

Chassignol, M., Хорошавин, A. B., Klímová, 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, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: a review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/access.2020.2988510

Chounta, I., Bardone, E., Raudsep, A., & Pedaste, M. (2021). Exploring Teachers' Perceptions of Artificial Intelligence as a Tool to Support their Practice in Estonian K-12 Education. *International Journal of Artificial Intelligence in Education*, *32*(3), 725–755. https://doi.org/10.1007/s40593-021-00243-5

Cohen, J. (1989). *Deliberation and Democratic Legitimacy. The Good Polity.*

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Lawrence Erlbaum Associates, Inc.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1). https://doi.org/10.1186/s41239-023-00392-8

Crow, T., Luxton-Reilly, A., & Wünsche, B. C. (2018). *Intelligent tutoring systems for programming education*. https://doi.org/10.1145/3160489.3160492

Cruzes, D. S., & Dybå, T. (2011). *Recommended Steps for Thematic Synthesis in Software Engineering*. https://doi.org/10.1109/esem.2011.36

DeCamp, M., & Lindvall, C. (2020). Latent bias and the implementation of artificial intelligence in medicine. *Journal of the American Medical Informatics Association*, *27*(12), 2020–2023. https://doi.org/10.1093/jamia/ocaa094

Devedžić, V. (2004). Web Intelligence and Artificial Intelligence in Education. *Journal of Educational Technology & Society*, *7*(4), 29–39. http://www.jstor.org/stable/jeductechsoci.7.4.29

Eilermann, S., Wehmeier, L., Niggemann, O., & Deuter, A. (2023). *KIAAA: An AI Assistant for Teaching Programming in the Field of Automation*. https://doi.org/10.1109/indin51400.2023.10218157

European Commission, Directorate-General for Education, Youth, Sport and Culture, (2022*). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*, Publications Office of the European Union. https://data.europa.eu/doi/10.2766/153756

Figueiredo, J., & García-Peñalvo, F. J. (2020). *Intelligent Tutoring Systems approach to Introductory Programming Courses*. https://doi.org/10.1145/3434780.3436614

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ. British Medical Journal*, *331*(7524), 1064–1065. https://doi.org/10.1136/bmj.38636.593461.68

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

*Guidelines for the use of Artificial intelligence in Education*. (n.d.). University of Oulu. https://www.oulu.fi/en/for-students/studying-university/guidelines-use-artificial-intelligence-education

Gurupur, V. P., & Wan, T. T. H. (2020). Inherent bias in Artificial Intelligence-Based Decision support systems for healthcare. *Medicina*, *56*(3), 141. https://doi.org/10.3390/medicina56030141

Hao, Q., Smith, D. H., Ding, L., Ko, A. J., Ottaway, C., Wilson, J., Arakawa, K., Turcan, A., Poehlman, T., & Greer, T. (2021). Towards understanding the effective design of automated formative feedback for programming assignments. *Computer Science Education*, *32*(1), 105–127. https://doi.org/10.1080/08993408.2020.1860408

Hellas, A., Leinonen, J., Sarsa, S., Koutcheme, C., Kujanpää, L., & Sorva, J. (2023). *Exploring the Responses of Large Language Models to Beginner Programmers' Help Requests*. https://doi.org/10.1145/3568813.3600139

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T. T., Shum, S. B., Santos, O. C., Rodrigo, M. M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in Education: Towards a Community-Wide framework. *International Journal of Artificial Intelligence in Education*, *32*(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

Huang, A. Y., Lu, O. H., & Yang, S. J. (2023). Effects of artificial Intelligence–Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers and Education/Computers & Education*, *194*, 104684. https://doi.org/10.1016/j.compedu.2022.104684

Ismail, M., & Ade-Ibijola, A. (2019). *Lecturer's Apprentice: A Chatbot for Assisting Novice Programmers*. https://doi.org/10.1109/imitec45504.2019.9015857

Jia, X., & Hermans, F. (2022). *Teaching Quality in Programming Education:* https://doi.org/10.1145/3501385.3543962

Kazemitabaar, M., Chow, J., To, C. K., MA, Ericson, B., Weintrop, D., & Grossman, T. (2023). *Studying the effect of AI Code Generators on Supporting Novice Learners in Introductory Programming*. https://doi.org/10.1145/3544548.3580919

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, Problem-Based, experiential, and Inquiry-Based teaching. *Educational Psychologist :/Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

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. https://doi.org/10.1016/j.infsof.2008.09.009

Kitchenham, B., Budgen, D., & Brereton, O. P. (2011). Using mapping studies as the basis for further research – A participant-observer case study. *Information and Software Technology, 53*(6), 638–651. https://doi.org/10.1016/j.infsof.2010.12.011

Latulipe, C., Rorrer, A., & Long, B. E. (2018). *Longitudinal Data on Flipped Class Effects on Performance in CS1 and Retention after CS1*. https://doi.org/10.1145/3159450.3159518

Lau, S., & Guo, P. J. (2023). *From "Ban It Till We Understand It" to "Resistance is Futile": How University Programming Instructors Plan to Adapt as More Students Use AI Code Generation and Explanation Tools such as ChatGPT and GitHub Copilot*. https://doi.org/10.1145/3568813.3600138

Leavy, S. (2018). *Gender Bias in Artificial Intelligence: The Need for Diversity and Gender Theory in Machine Learning*. https://ieeexplore.ieee.org/documeant/8452744

Li, X., & Prasad, C. (2005). *Effectively teaching coding standards in programming*. https://doi.org/10.1145/1095714.1095770

Lim, L., Bannert, M., Van Der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Raković, M., Molenaar, I., Moore, J. D., & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, *139*, 107547. https://doi.org/10.1016/j.chb.2022.107547

Luckin, R., & Holmes, W. H. (2016). *Intelligence Unleashed: An argument for AI in Education*. https://discovery.ucl.ac.uk/id/eprint/1475756/

Marshall, MN. (1996). The key informant technique. *Family Practice*, *13*(1), 92–97. https://doi.org/10.1093/fampra/13.1.92

Martín Núñez, J. L., & Diaz Lantada, A. (2020). Artificial Intelligence Aided Engineering Education: State of the Art, Potentials and Challenges. *International Journal of Engineering Education*, *36*(6), 1740–1751.

Miao, F., Holmes, W. (2023) Guidance for generative AI in education and research. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Milana, M., Brandi, U., Hodge, S., & Hoggan-Kloubert, T. (2024). Artificial intelligence (AI), conversational agents, and generative AI: implications for adult education practice and research. *International Journal of Lifelong Education*, *43*(1), 1–7. https://doi.org/10.1080/02601370.2024.2310448

Morrison, B. B., Dorn, B., & Guzdial, M. (2014). *Measuring cognitive load in introductory CS*. https://doi.org/10.1145/2632320.2632348

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2023). Uncovering Blind Spots in Education Ethics: Insights from a Systematic Literature Review on Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00384-9

Myers, M., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, *17*(1), 2–26. https://doi.org/10.1016/j.infoandorg.2006.11.001

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, *71*(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. T. (2022). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, *28*(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Nwana, H. S. (1990). Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, *4*(4). https://doi.org/10.1007/bf00168958

Ouh, E. L., Gan, B. K. S., Gan, B. K. S., & Wlodkowski, S. (2023). *ChatGPT, Can You Generate Solutions for my Coding Exercises? An Evaluation on its Effectiveness in an undergraduate Java Programming Course.* https://doi.org/10.1145/3587102.3588794

Pears, A., Seidman, S. B., Malmi, L., Mannila, L., Adams, E., Bennedsen, J., Devlin, M., & Paterson, J. H. (2007). A survey of literature on the teaching of introductory programming. *SIGCSE Bulletin*, *39*(4), 204–223. https://doi.org/10.1145/1345375.1345441

Pedró, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education : challenges and opportunities for sustainable development. *MINISTERIO DE EDUCACIÓN*. http://repositorio.minedu.gob.pe/handle/20.500.12799/6533

Peng, Z., & Wan, Y. (2023). Human vs. AI: Exploring students' preferences between human and AI TA and the effect of social anxiety and problem complexity. *Education and Information Technologies*, *29*(1), 1217–1246. https://doi.org/10.1007/s10639-023-12374-4

Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Craig, M., Keuning, H., Kiesler, N., Kohn, T., Luxton-Reilly, A., MacNeil, S., Petersen, A., Pettit, R., Reeves, B., & Šavelka, J. (2023a). *The Robots Are Here: Navigating the Generative AI Revolution in Computing Education*. https://doi.org/10.1145/3623762.3633499

Prather, J. E., Reeves, B., Denny, P., Becker, B. A., Leinonen, J., Luxton-Reilly, A., Powell, G., Finnie-Ansley, J., & Santos, E. A. (2023b). "It's Weird That it Knows What I Want": Usability and Interactions with Copilot for Novice Programmers. *ACM Transactions on Computer-human Interaction*, *31*(1), 1–31. https://doi.org/10.1145/3617367

Puryear, B., & Sprint, G. (2022). Github copilot in the classroom: Learning to code with AI assistance. *Journal of Computing Sciences in Colleges, 38*(1), 37–47.

Real-Fernández, A., Molina-Carmona, R., & Largo, F. L. (2019). *Instructional Strategies for a Smart Learning System*. https://doi.org/10.1145/3362789.3362915

Robins, A., Rountree, J., & Rountree, N. (2003). Learning and Teaching Programming: A review and discussion. *Computer Science Education*, *13*(2), 137–172. https://doi.org/10.1076/csed.13.2.137.14200

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

Rubiano, S. M. M., López-Cruz, O., & Soto, E. G. (2015). *Teaching computer programming: Practices, difficulties and opportunities*. https://doi.org/10.1109/fie.2015.7344184

Rubin, M. J. (2013). *The effectiveness of live-coding to teach introductory programming*. https://doi.org/10.1145/2445196.2445388

Saari, M. Y., Rantanen, P., Nurminen, M., Kilamo, T., Systä, K., & Abrahamsson, P. (2023). Survey of AI Tool Usage in Programming Course: Early Observations. In *Lecture notes in business information processing* (pp. 182–191). https://doi.org/10.1007/978-3-031-48550-3_18

Schiff, D. (2021). Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. *International Journal of Artificial Intelligence in Education*, *32*(3), 527–563. https://doi.org/10.1007/s40593-021-00270-2

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Swist, T., Shum, S. B., & Gulson, K. N. (2024). Co-producing AIED Ethics Under Lockdown: an Empirical Study of Deliberative Democracy in Action. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00380-z

Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial Intelligence in Education: AIED for Personalised Learning Pathways. *Electronic Journal of e-Learning*, *20*(5), 639–653. https://doi.org/10.34190/ejel.20.5.2597

Timms, M. J. (2016). Letting artificial intelligence in education out of the box: educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education*, *26*(2), 701–712. https://doi.org/10.1007/s40593-016-0095-y

Van Slyke, C., Johnson, R. T., & Sarabadani, J. (2023). Generative Artificial intelligence in Information Systems Education: Challenges, consequences, and responses. *Communications of the Association for Information Systems*, *53*(1), 1–21. https://doi.org/10.17705/1cais.05301

Vihavainen, A., Paksula, M., & Luukkainen, M. (2011). *Extreme apprenticeship method in teaching programming for beginners*. https://doi.org/10.1145/1953163.1953196

Vivar, J. M. F., & García-Peñalvo, F. J. (2023). Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4). *Comunicar Digital/Comunicar*, *31*(74), 37–47. https://doi.org/10.3916/c74-2023-03

Wang, T., Lund, B., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: generative AI, chatbots, analytics, and international student success. *Applied Sciences*, *13*(11), 6716. https://doi.org/10.3390/app13116716

Wardat, Y., Tashtoush, M. A., Alali, R., & Saleh, S. (2024). Artificial Intelligence in Education: Mathematics Teachers' perspectives, Practices and challenges. *Iraqi Journal for Computer Science and Mathematics*, *5*(1), 60–77. https://doi.org/10.52866/ijcsm.2024.05.01.004

Wilson, B. C., & Shrock, S. (2001). Contributing to success in an introductory computer science course. *SIGCSE Bulletin*, *33*(1), 184–188. https://doi.org/10.1145/366413.364581

Wood, D., Bruner, J. S., & Ross, G. (1976). THE ROLE OF TUTORING IN PROBLEM SOLVING*. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

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), 39. https://doi.org/10.1186/s41239-019-0171-0

# Appendix A. Search strings used in the literature review

ACM Digital Library:

- [All: teaching methods] AND [All: programming education]
- [All: cognitive load] AND [All: programming education]
- [All: artificial intelligence] AND [All: programming teacher]
- [All: artificial intelligence] AND [All: student learning outcomes]
- [All: intelligent tutoring systems] AND [All: programming education]

ERIC (Ebsco)

- ethical implications AND artificial intelligence education
- ChatGPT and programming education
- AI integration and programming curriculum

Google Scholar

- teaching coding
- "artificial intelligence" and "teaching coding"
- "artificial intelligence" and teaching coding
- "artificial intelligence" and "teaching programming"

IEEE Xplore

- ("All Metadata":feedback) AND ("All Metadata":programming education)
- ("All Metadata":artificial intelligence) AND ("All Metadata":education challenges)
- ("All Metadata":AI-driven development environments) AND ("All Metadata":programming education)
- ("All Metadata":ChatGPT) AND ("All Metadata":programming education)
- ("All Metadata":AI integration) AND ("All Metadata":programming curriculum)

Scopus:

- TITLE-ABS-KEY ( artificial AND intelligence AND education )
- TITLE-ABS-KEY ( artificial AND intelligence AND programming AND teacher )
- TITLE-ABS-KEY ( artificial AND intelligence ) AND programming AND teacher
- TITLE-ABS-KEY ( teaching AND programming ) OR ( teaching AND coding )
- TITLE-ABS-KEY ( teaching AND programming )
- TITLE-ABS-KEY ((artificial intelligence) AND higher AND education)
- TITLE-ABS-KEY ((artificial intelligence) AND education)
- TITLE-ABS-KEY (artificial intelligence) AND (organization change)
- TITLE-ABS-KEY ( artificial AND intelligence ) AND ( student AND learning AND outcomes )
- TITLE-ABS-KEY ( artificial intelligence ) AND ( education challenges )

- TITLE-ABS-KEY ( intelligent AND tutoring AND systems ) AND ( programming AND education )

# Appendix B. Interview questions

- What course(s) are you teaching?
- How many years have you been teaching?
- How experienced are you in using AI tools?
- How much do you think that students are using AI tools right now?
- How much have you heard your colleagues discussing these AI tools? (And in what settings?)
- Are you proficient in using AI tools for teaching programming?
- Is it required to know a lot about AI to use AI tools for teaching programming?
- Does using AI tools in programming education require more effort from teachers compared to traditional teaching methods?
- Is a challenge faced by programming teachers the lack of availability of specialists in AI technologies for guidance and support?
- Is there a lack of training and awareness on AI tools that makes it hard to utilise them successfully?
- Do you think implementing AI tools in the programming classroom presents challenges for student learning?
- Do you actively seek opportunities to integrate AI tools into your programming curriculum?
- Have you used AI tools for teaching programming? Have they made some aspect of teaching programming easier or better?
- Do you think implementing AI tools in the programming classroom presents significant opportunities for enhancing student learning?
- Do you think AI tools have enhanced or worsened the learning experience for programming students?
- Does incorporating AI tools have an effect on motivation among programming students?
- Do AI tools provide flexibility in delivering programming education?
- In your opinion, do AI tools effectively cater to the diverse learning needs of programming students?
- Anything else?

# Appendix C. Participant consent form

You are invited to participate in a research study conducted as part of a Master's thesis conducted by Niklas Riekki, a student at Oulu University. Your participation in this study is entirely voluntary, and you have the right to withdraw at any time without any negative consequences. Before deciding whether to participate, please read the following information carefully.

This study aims to understand the current situation regarding opportunities and challenges with artificial intelligence tools in the context of teaching programming.

By agreeing to participate in this study, you acknowledge the following rights:
● The right to voluntarily participate and to refuse from participating.
● The right to stop participating at any point without any negative consequences to you.
● The right to withdraw consent to participating at any time.
● The right to receive information about the contents of the research, the handling of personal information, the practical implementation of the research, and details on how the collected data is handled and stored.
● The right to receive an understandable and truthful overview of the goals of the research and of the possible harms and risks.
● The right to know when the researcher and the participants have a dynamic outside of just researcher and participant (i.e. employer and employee).

All the collected data will be stored for up to six months or until the thesis is complete, whichever is earlier. No personal data from which the participants can be recognized from will be stored.

Contact info:
Researcher: Niklas Riekki, niklas.riekki@student.oulu.fi
Supervisor: Elina Annanperä, elina.annanpera@oulu.fi