Kong, S.-C., Cheung, W. M.-Y., & Zhang, G. (2023). Evaluating an Artificial Intelligence Literacy Programme for Developing University Students' Conceptual Understanding, Literacy, Empowerment and Ethical Awareness. *Educational Technology & Society*, *26*(1), 16-30. https://doi.org/10.30191/ETS.202301_26(1).0002

Evaluating an Artificial Intelligence Literacy Programme for Developing University Students' Conceptual Understanding, Literacy, Empowerment and Ethical Awareness

Siu-Cheung Kong[1,2*], William Man-Yin Cheung[2] and Guo Zhang[2]

[1]Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR // [2]Centre for Learning, Teaching and Technology, The Education University of Hong Kong, Hong Kong SAR // siucheungkong@gmail.com // williamcheung@eduhk.hk // gzhang@friends.eduhk.hk

[*]Corresponding author

(Submitted November 29, 2021; Revised March 26, 2022; Accepted April 25, 2022)

ABSTRACT: Emerging research is highlighting the importance of fostering artificial intelligence (AI) literacy among educated citizens of diverse academic backgrounds. However, what to include in such literacy programmes and how to teach literacy is still under-explored. To fill this gap, this study designed and evaluated an AI literacy programme based on a multi-dimensional conceptual framework, which developed participants' conceptual understanding, literacy, empowerment and ethical awareness. It emphasised conceptual building, highlighted project work in application development and initiated teaching ethics through application development. Thirty-six university students with diverse academic backgrounds joined and completed this programme, which included 7 hours on machine learning, 9 hours on deep learning and 14 hours on application development. Together with the project work, the results of the tests, surveys and reflective writings completed before and after these courses indicate that the programme successfully enhanced participants' conceptual understanding, literacy, empowerment and ethical awareness. The programme will be extended to include more participants, such as senior secondary school students and the general public. This study initiates a pathway to lower the barrier to entry for AI literacy and addresses a public need. It can guide and inspire future empirical and design research on fostering AI literacy among educated citizens of diverse backgrounds.

Keywords: Application development, Artificial intelligence literacy, Conceptual framework, Ethical awareness, University students

1. Introduction

Fostering artificial intelligence (AI) literacy for all citizens has become increasingly crucial, given AI's potential to reshape the competitive landscape and its relevance to individuals' lives and work (Fosso Wamba et al., 2021; JRC & OECD, 2021; WIPO, 2019). However, few studies have comprehensively examined how and what exactly to teach to educate citizens of diverse backgrounds.

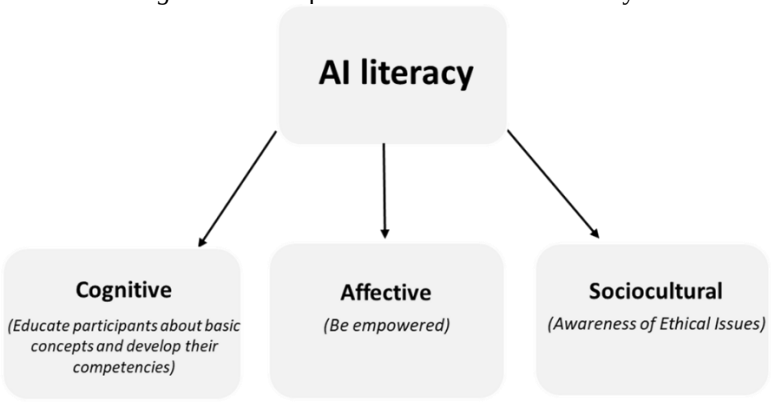
Most studies of conceptual teaching involve mathematical formulae and programming codes, focusing primarily on computer science majors and students with programming knowledge (Green, 2021; Pouly et al., 2019; Stadelmann et al., 2021; Tedre et al., 2021). This approach creates a barrier to literacy amongst the public (Long & Magerko, 2020). While ethical issues related to AI have received increased attention (Ashok et al., 2022; Jobin et al., 2019; Kuipers, 2020; Mehrabi et al., 2021; Prunkl, 2022), ethics thus far have rarely been an explicit component of AI courses (Saltz et al., 2019), and limited information is available on the ethical considerations covered in AI classes (Garrett et al., 2020).

To fill this research gap and to serve social equity and sustainable development goals (Kong et al., 2021b; OECD, 2018a; Vinuesa et al., 2020), this study develops an AI programme that focuses on conceptual understanding, literacy, empowerment and ethical awareness. The literacy development framework presented here focuses on conceptual building, emphasising project work in application development and enhancing participants' awareness of the ethical considerations arising from such work. This study reports the process of designing, implementing and evaluating this AI literacy programme.

2. Background

We follow the conceptual framework of AI literacy from Kong and Zhang (2021) (see Figure 1). This framework is comprised of three dimensions: the cognitive dimension; the affective dimension; and the sociocultural dimension.

Figure 1. Conceptual framework of AI literacy



The cognitive dimension involves teaching major fundamental AI concepts, particularly machine learning and deep learning, and how to use them to evaluate and understand the real world. These concepts have profound societal impacts and are essential to fostering AI literacy (OECD, 2018b; Touretzky et al., 2019; Wong et al., 2020). By understanding these concepts, learners should be able to evaluate AI artefacts in their lives and the impacts of the technology, then apply the concepts to understand the AI-permeated world, and form their attitudes and responses accordingly.

The affective dimension serves to empower participants so they can participate with confidence in the digital world. It contains four components: grasping the value of AI (Thomas & Velthouse, 1990); perceiving the social impact of AI (Frymier et al., 1996); believing in one's ability to produce novel AI ideas and solutions (Paulus & Brown, 2003); and being confident in one's competence in engaging with AI (Bandura, 1982). This four-factor model (meaningfulness, impact, creative self-efficacy and AI self-efficacy) is consistent with the idea of future literacy from the United Nations Educational, Scientific and Cultural Organization (UNESCO), which aims to strengthen learners' imagination and prepare them for change (Yi, 2021). Our initiative aims to develop participants' self-confidence in conducting AI-related activities, educate them about AI's significance and societal impacts, and enhance their digital creativity.

Finally, the sociocultural dimension concerns the ethical use of AI. Our course followed the ethical principles outlined in Kong and Zhang (2021), which was built on those stated in the Belmont Report (NCPHS, 1978): (1) the use of AI should not violate human autonomy; (2) AI's benefits should outweigh its risks; and (3) AI's benefits and risks should be distributed equally. These three principles (autonomy, beneficence/non-maleficence and fairness) have also been covered by recent AI ethical frameworks (Floridi & Cowls, 2019; HLEG, 2019; OECD, 2019). As effective guidelines to follow, they serve as the constructs of the ethical consideration survey detailed in Section 3.4 below.

This multidimensional conceptual framework informs our design, development and evaluation of this literacy programme. Using this framework, this study focused on the three research questions: (1) Can the AI literacy programme address AI concepts and literacy? (2) Will participants feel empowered after completing the AI literacy programme? and (3) Can the AI literacy programme foster participants' ethical awareness?

3. Methodology

3.1. Course participants

We launched a literacy programme at a Hong Kong university for convenience sampling. A total of 36 university students from diverse backgrounds joined the programme. Twenty-three were female and thirteen were male. Seventy-five per cent of the participants were enrolled in bachelor's degree programmes, including students in their first, second, third and fourth years of study. The remaining participants were from postgraduate or higher diploma programmes. As shown in Table 1, the participants came from a wide range of academic backgrounds, namely Mathematics, Information and Communication Technology, Health Education, Chinese Language Studies, Psychology, the Sciences (Natural Science & STEM Education), English Language Studies, General Studies, Music, History, Global and Environmental Studies and Global and Hong Kong Studies.

Table 1. Distribution of programme participants' academic backgrounds

Academic background	Number (percentage)	Academic background	Number (percentage)
Mathematics	8 (22.22%)	English Language Studies	2 (5.56%)
Information and Communication Technology	5 (13.89%)	General Studies	2 (5.56%)
Health Education	4 (11.11%)	Music	2 (5.56%)
Chinese Language Studies	4 (11.11%)	History	1 (2.78%)
Psychology	3 (8.33%)	Global and Environmental Studies	1 (2.78%)
The Sciences (Natural Science & STEM Education)	3 (8.33%)	Global and Hong Kong Studies	1 (2.78%)
Total			36 (100%)

3.2. Curriculum

The programme consisted of three courses: Machine Learning, Deep Learning and Developing Artificial Intelligence Applications. The first two courses develop conceptual understanding of two important AI areas (Kong & Zhang, 2021; Kong et al., 2021b), thus fostering AI literacy in the cognitive dimension. The third course further develops AI literacy through applying acquired concepts to project work. This project work in turn serves as a concrete example to reflect on ethical issues, thus covering the sociocultural dimension. The affective domain is also enhanced as participants can feel more empowered with more understanding of AI throughout all three courses.

3.2.1. Course 1: Machine learning

Course 1 introduced the concepts and some related algorithms in both supervised and unsupervised learning. An overview of AI's development was first provided, followed by concepts of strong and weak AI. The participants were encouraged to voice their thoughts on AI's impact on society.

With this foundation, the participants then discussed the "five steps of machine learning," together with hands-on experience using these steps to perform image recognition on an online platform. Afterwards, the participants learned about two instances of supervised learning, "regression" and "classification," through examples and hands-on experience. Finally, this course covered the concept and working principles of unsupervised learning by applying k-means clustering in a series of case studies (Kong et al., 2021b).

In teaching these concepts, we emphasised conceptual building from the beginning: we used analogies and real-life scenarios rather than programme codes and mathematical formulae to foster students' conceptual understanding (Kong et al., 2021b). This allowed the course participants to understand the fundamental concepts of AI and the rationale that underlie them, thereby simplifying the learning process while deepening their conceptual understanding.

3.2.2. Course 2: Deep learning

In the same vein, Course 2 developed the participants' conceptual understanding of deep learning. The course covered several topics, including data cleaning, data augmentation, neural networks, computer vision, deep learning and convolution neural networks. Through reviewing the application of the five steps of machine learning in case studies, the course presented the ideas of data cleaning and data augmentation. The concept of neural networks was introduced by explaining the ideas of perception, input layers, hidden layers, output layers and weights, among others. The participants' understanding was deepened through a lab session of training neural networks to learn to distinguish different data points within various data sets. The concept of computer vision was then discussed, as it is commonly applied in neural networks; related applications were shared with the course participants to provide first-hand experience. The participants were also introduced to convolution neural networks through a lab session and various discussions. Finally, the participants were given the opportunity to experience more machine learning tools.

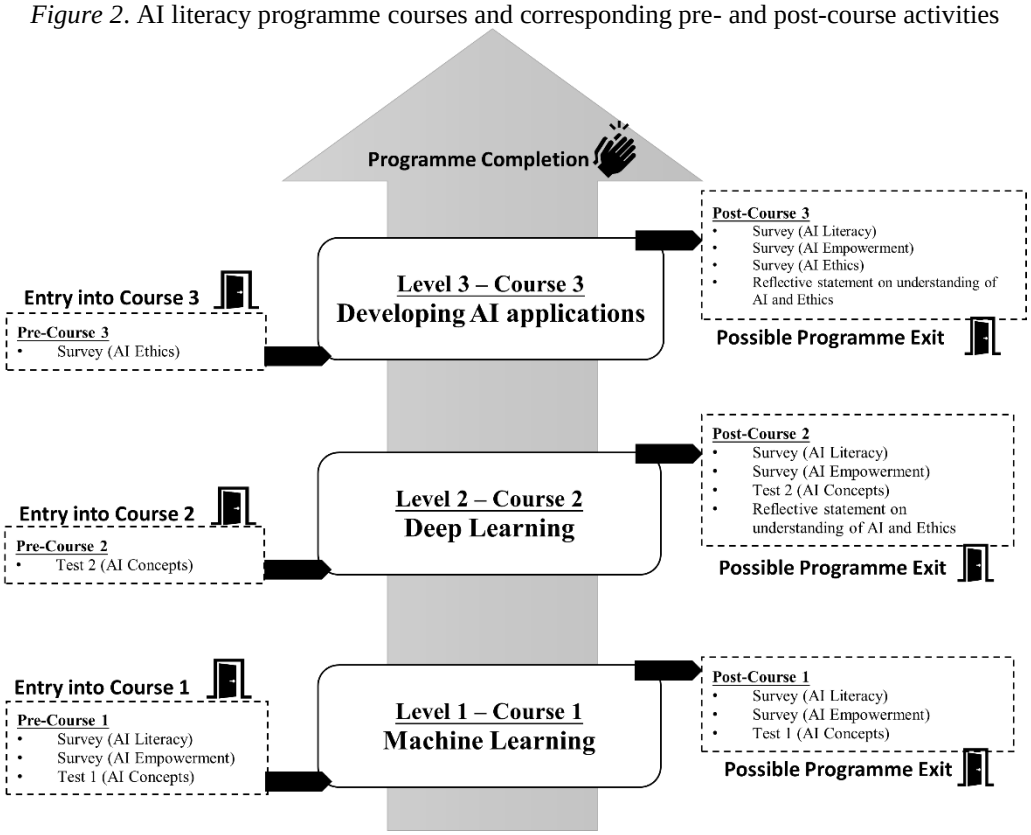
3.2.3. Course 3: Developing AI applications

Course 3 developed the participants' ethical awareness through project work in application development. Before the course, the participants were provided with self-directed reading materials concerning ethical dilemmas and principles of ethics related to AI. Ethical principles on designing and using AI were introduced in the lectures to deepen participants' understanding through real-life examples and discussions. Three additional examples of AI applications were then presented, and the participants discussed the ethical issues involved in these cases in groups. Afterwards, the participants brainstormed their project ideas with preliminary feedback from the course tutors, who later organised individual consultations to provide more detailed instructions on project developments. Sessions introducing different project development platforms and tools, such as Microsoft Azure Machine Learning Studio, Google Teachable Machine and Microsoft Azure QnA Maker, were also offered. With this foundation, the participants began collaborative work on their projects. Each group presented their work in the final session with peer evaluation and discussion on the ethical issues involved in each project, which provided them with more opportunities for reflection, further fostering their ethical awareness. Each group's work was assessed using a rubric that included the discussion of ethical considerations as an important component (see Table 7 and Table 11).

3.3. Course administration

Figure 2 shows the flow of courses in the programme and indicates the corresponding pre- and post-course evaluation activities. These surveys and tests were conducted both before and after the courses to study the participants' progress. Because the content of Course 1 differed from that of Course 2, the AI concepts on tests 1 and 2 were designed according to the relevant content. The participants were asked to write a reflective piece either in English or Chinese on their understanding of AI and related ethical issues. The participants were able to exit the programme at the end of each course.

Figure 2. AI literacy programme courses and corresponding pre- and post-course activities



The number of participants in Courses 1, 2 and 3 were 120, 82 and 36, respectively. In this article, we report the findings related to the 36 participants who completed the whole programme. The number of participants for the three courses decreased as the programme progressed. This could be attributed to the increasing difficulty of each course, leading to the withdrawal of participants less confident with the material. This echoes the rationale that participants who are more empowered are more likely to begin or to continue working on the task at hand and make more effort in AI-related projects (Paulus & Brown, 2003; Kong et al., 2021b).

3.4. Instrument design and use

Quantitative and qualitative data were collected through tests, surveys and reflective writings designed to explore participants' development of AI literacy and to encourage self-reflection on ethical issues. In this study, we present the analyses of participants' responses to the following instruments. (1) AI concepts tests: These tests assessed how an AI literacy programme can develop participants' concepts. (2) AI literacy survey: A survey assessed participants' perceptions of their own AI literacy. (3) AI empowerment survey: This survey evaluated participants' empowerment (self-efficacy, meaningfulness, impact, creativity) after completing the AI literacy programme. (4) Survey on ethical considerations in developing AI applications: This survey was used to assess participants' awareness of ethical issues around AI applications. (5) Focus group interview questions on the AI literacy course: The participants were interviewed about their views on AI literacy. (6) Self-reflections on understanding AI ethics. The participants were asked to write 100 to 200 words in either English or Chinese on their understanding of ethical issues related to AI.

Table 2. Bilingual taxonomy of keywords on AI and ethics used for text mining from course participants' self-reflections

Real-world Examples
dilemmas
dilemma / debate / important questions / 困境 / 辯論 / 重要問題
autonomous car / 自動駕駛汽車
copyrights / rights / careful attention / without their knowledge or consent / remuneration / remunerate / reward / 版權 / 權利 / 知識產權 / 謹慎關注 / 報酬 / 未經他們的知情或同意
decision-making / make decisions / make right decisions / decide / moral decision / choose / right answer / judgement / judge / 作出決定 / 作出正確的決定 / 決定 / 道德決策 / 選擇 / 抉擇 / 正確答案 / 判斷
ethics / ethical consideration / ethical issue / ethical problem / ethical reflection / ethical conundrum / ethical solution / 倫理 / 道德 / 道德考量 / 道德問題 / 道德議題 / 道德反思 / 道德難題 / 符合倫理的解決方案
threats
harm / harmful / risk / risky / safety / safe / safely / data security / consequence / bad / threat / 危害 / 有害 / 風險 / 有風險的 / 安全 / 後果 / 不良 / 威脅
replace / 取代 / 代替
piracy / plagiarism / exploit / personal data / privacy / private information / personal information / 盜版 / 抄襲 / 利用 / 個人數據 / 私隱 / 個人信息 / 個人資料
misuse / misusing / abuse / 濫用
discrimination / discriminatory 歧視
bias / biased / stereotypical representations / prejudice / stereotype / 偏見 / 刻板印象
negative / non-transparent / unexplainable / unjustifiable outcomes / lack of explainability / trouble / problematic / problem / inequity / unfairness / unfair / lack of clarity / 負面 / 不透明 / 無法解釋 / 不合理的結果 / 缺乏可解釋性 / 麻煩 / 問題 / 不公平 / 不夠清晰 / 不可靠
isolation / disintegration / reduction of human-to-human interaction / polarise social relationships / damage the wellbeing of individuals / public welfare / 隔離 / 解體 / 減少人與人之間的互動 / 分化社會關係 / 損害個人福祉 / 公共福利
Principles
guide / principle / framework / legal / 指引 / 原則 / 框架 / 合法的
beneficence / nonmaleficence / benefits should outweigh harm / advantages outweigh disadvantages / 為善 / 毋損害 / 利益應大於傷害 / 優點大於缺點
justice / fairness / fair / equity / 正義 / 公道 / 公義 / 公平 / 合理
accuracy / accurate / reliability / reliable / soundness / sound / good reasons / reasonableness / reasonable / 準確 / 可靠 / 健全 / 充分理由
accountability / autonomy / sustainability / transparency / integrity / accountable / sustainable / transparent / responsible / responsibility / responsibilities / accountable / 問責 / 自主 / 可持續 / 透明 / 完整性 / 負責 / 責任 regulation / regulate / 規管 / 監管 / 規範

For the AI concepts tests, the participants were asked to answer multiple choice questions about AI concepts. The tests were designed and guided by the learning progression set forth in Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001). The AI literacy survey addressed the following themes: "AI concepts," "using AI concepts for evaluation" and "using AI concepts to understand the real world" (Kong & Zhang, 2021). The survey items were designed to evaluate the participants' understanding of concepts and related competencies.

The AI empowerment survey included four components: "AI self-efficacy," "meaningfulness," "impact" and "creative self-efficacy" (Kong & Zhang, 2021; Kong et al., 2021b). The ethical consideration survey evaluated the participants' awareness of ethical issues related to AI applications by employing three components: autonomy, beneficence and fairness (Kong & Zhang, 2021). The survey consisted of 12 questions, with four questions concerning each component. For all of the surveys, the participants were asked to indicate their level of agreement with each item on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). Regarding the reliability of the instruments, the Cronbach's alpha coefficient for the post-test results of the two AI concepts tests and the AI literacy, empowerment and ethical consideration surveys were 0.66, 0.67, 0.89, 0.93 and 0.76, respectively.

Focus group interviews were held to solicit in-depth views from the participants on the content of the courses and their perceptions of AI's relevance to society. The participants were asked to write a reflective piece on their understanding of AI and related ethical issues before and after joining the courses. Besides employing the survey, the study evaluated participants' ethical awareness by analysing their self-reflective writings. The text was analysed by a bilingual text-mining system using a keyword framework on AI and ethics (Kong et al., 2018; Kong, 2021; Kong et al., 2021a). To identify the level of participants' ethical awareness, they were asked to write a self-reflective essay on 'understanding of AI and ethics' before and after Course 3, using the Moodle discussion forum. The participants were allowed to write in English or Chinese. Accordingly, a bilingual taxonomy of keywords with synonyms in both English and Chinese (see Table 2) was designed. To do so a course instructor and a research staff independently went through contents of the courses and the reflective essays of the participants to identify keywords, followed by discussions to arrive at the final version. We then used the bilingual text-mining system to count the number of keywords in the participants' self-reflective writing.

4. Results and discussion

This section reports the results of the two AI concepts tests and the AI literacy, empowerment and ethical consideration surveys both before and after attending the courses. The key findings from the self-reflective writings are also included below. The discussions here complement those in two related publications: Kong and Zhang (2021) discuss thoroughly the establishment of the AI literacy framework outlined in Section 2, whereas Kong et al. (2021b) reports findings about Course 1. The current discussion, in contrast, offers a longitudinal investigation of students completing the entire programme.

4.1. Developing AI concepts and literacy

This section reports the development of the participants' conceptual understanding and AI literacy. The results of the AI concepts tests and AI literacy survey show that the three courses successfully enhanced the participants' conceptual understanding and literacy. Tables 3 and 4 show the means, standard deviations and paired *t*-test scores of the first and second concepts tests, respectively. The findings show that the increase in learning achievements in both concepts tests was statistically significant. This indicates that course participants from diverse backgrounds experienced significant progress in grasping AI-related concepts. This also implies that Course 1 and Course 2 provided participants with conceptual readiness for developing AI applications, which also built a framework for discussing ethics.

Table 3. Statistical results on the AI concepts Test 1 before and after Course 1

Concept	Before Course 1 (max. mark = 14)		After Course 1 (max. mark = 14)		Paired *t*-test
	Mean	*SD*	Mean	*SD*	
Machine learning	6.87	2.00	10.75	2.20	9.49^{***}

Note. $N = 36$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Table 4. Statistical results on the AI concepts Test 2 before and after Course 2

Concept	Before Course 2 (max. mark = 14)		After Course 2 (max. mark = 14)		Paired *t*-test
	Mean	*SD*	Mean	*SD*	
Deep learning	6.72	2.50	9.19	2.75	4.68^{***}

Note. $N = 36$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Participants also reported whether they had programming knowledge. Table 5 further compares how participants with and without programming knowledge performed in the concepts' tests. The results show that the two groups of participants did not exhibit statistically significant difference before Course 1, after Course 1 and before Course 2. However, after Course 2 participants without programming knowledge demonstrated even better performance which is statistically significant. This analysis supports that our courses, while not involving programming, are suitable for participants from diverse backgrounds to develop AI concepts.

Table 5. Comparing results of AI concepts tests by participants with and without programming knowledge

Concepts Test 1 (max. mark = 14)	Without programming knowledge (N = 14)		With programming knowledge (N =22)		Paired t-test
	Mean	*SD*	Mean	*SD*	
Before Course 1	7.29	1.98	6.59	2.02	1.02
After Course 1	11.36	2.37	10.36	2.04	1.34
Concepts Test 2 (max. mark = 14)					
Before Course 2	7.14	2.03	6.45	2.77	0.80
After Course 2	10.57	2.47	8.32	2.61	2.58*

Note. N = 36; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

The participants' perceptions of their own AI literacy (see Table 6) increased significantly and stabilised after Course 1. This may be attributable to the ceiling effect, as a similar phenomenon was witnessed by Lee et al. (2021). The mean scores remained high after Course 1 (with the mean score above 4.10 out of 5).

Table 6. Statistical results on the AI literacy survey before and after the courses

	Before Course 1 Mean (SD)	After Course 1 Mean (SD)	After Course 2 Mean (SD)	After Course 3 Mean (SD)	F-value	p-value	Partial eta squared	Pairwise comparison
AI literacy (max. mark = 5)	2.74 (0.72)	4.10 (0.41)	4.07 (0.46)	4.19 (0.41)	46.91	$< .001^{***}$	0.81	Before Course 1 < After Course 1; Before Course 1 < After Course 2; Before Course 1 < After Course 3;

Note. N = 36; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Table 7. Selected quotes from participants' interviews and self-reflective writing after Course 3 on the usefulness of the project work in developing AI concepts

In terms of learning the concepts, my understanding of AI was not clear until conducting the project work. Now I understand that the study of AI focuses on computers' abilities to learn and identity objects. In our project ("Mask detection, home security and reminder systems"), the application identified whether the targets wore masks and if they were family members. This application development process helped me grasp the key elements of AI, which furthered my conceptual understanding (interview; S5).
The project work in Class 3 was indeed practical. We had opportunities to exchange ideas with others. The project work allowed us to present in class, interact with each other and discuss the topics and brainstorm how to implement the AI programme. These interactions and hands-on experiences contributed significantly to my understanding of AI concepts (interview; S14).
I used the "teachable machine" concept in my project. Through the project work, I deeply understood the underlying rationale of the algorithm. My conceptual understanding was strengthened so that in my daily life, I am now aware of the working principles of the algorithms behind some common related AI artefacts (interview; S31).
After Course 3, I have a better understanding of AI concepts, principles and applications. Although the trial phase was full of challenges, from trial and observation, we have learned about AI practices in society. This makes me more eager to apply the principles of AI in my work, and at the same time I am eager to improve the application of AI. (reflection translated from Chinese; S1).

Beyond the significant improvement demonstrated in the test and survey, participants' responses in focus group interviews and their reflective pieces (see Table 7) after Course 3 validated the usefulness of the project work for their understanding of AI concepts. For instance, some of the participants mentioned that they better understood AI concepts after completing their projects. They understood that AI involves computer learning and object

identification, an analytical process undertaken before making decisions. They acquired knowledge of the underlying principles of algorithms more thoroughly, which strengthened their conceptual understanding. Thematic analysis of interview transcripts and reflective writings showed the project works deepened participants' knowledge of machine learning and various platforms for developing AI applications (see Figure 3). This validated the feasibility of teaching concepts through project work, transforming their conceptual understanding from simple knowledge acquisition to novel applications of that knowledge (Roth, 1990; Schleicher, 2018). This in turn contributed to deeper conceptual understanding.

The development of participants' AI literacy was also evidenced by the results of their projects. Table 8 provides an overview of the 10 projects with project names and the total score and sub-score for ethical considerations each group received. In the project work, each group defined a real-life problem and created an application using the platforms demonstrated in class to solve that problem. In this process, they also evaluated their application by considering the possible ethical issues involved, gaining a better understanding of the permeation of AI in the real world (Kong & Zhang, 2021).

Table 8. Overview of group projects on "Developing Artificial Intelligence Applications"

Project no.	Project name	Number of participants	Total score (max.: 12)	Score for discussion of ethical considerations (max.: 3)
1	"Mind-RoadBot": A mindfulness chatbot	4	12.00	3.00
2	"Perfect Letter": A tool to help toddlers learn good English handwriting	2	10.50	2.36
3	Grade prediction model for evaluating students' learning outcomes	4	10.00	2.40
4	"Play with Sol-Fa Names": Recognising hand signals for musical notes	3	10.00	2.00
5	Mask detection, home security and reminder systems	3	8.67	2.17
6	Predicting stock price to make a long-short portfolio	4	8.67	1.67
7	Garbage classification: Distinguishing different recyclable wastes from photos	5	8.25	2.17
8	"General Education Helper": A Q&A bot for choosing general education courses	4	7.67	1.50
9	"Healthy Life Helper": A chatbot to share health-related tips	3	7.50	1.84
10	A chatbot for song recommendations	4	6.50	1.88

Note. The total score is an average based on the marks from the instructors and participants.

4.2. Developing AI empowerment

Table 9 shows the results of the AI empowerment survey. As indicated, the mean scores remained high throughout the programme (above 4 marks a maximum score of 5) but increased significantly. This shows that participants felt empowered by learning the concepts and acquiring experience in developing AI applications.

Tracking the changes in participants' perceptions of their level of AI empowerment by course suggests that application development empowered the participants. The mean scores after Course 1 remained at a high level with no significant increase. This could be caused by the already high pre-course score (4.02 of 5), which signals that most participants in this group felt highly empowered even before starting the programme. Lee et al. (2021) reported a similar phenomenon in a study of AI literacy education and noted that this could be attributable to the ceiling effect. Although a significant decrease in the average score was observed before and after Course 2 (from 4.20 to 4.06), the mean score remained high. This may be due to the greater demands placed on participants in Course 2 as compared to Course 1. Student S4, for example, expressed in her reflective writing after Course 2 that "…*I found that the knowledge behind is actually quite complicated*…" and "…*we have to handle different technical issues such as the overfitting problem.*" This fluctuation in AI empowerment deviates from the results of the AI concepts tests which increased over both courses. One reason might be that the 36 students who ultimately completed the programme had high expectations for themselves and did not feel more empowered through the course, especially when their knowledge had yet to be applied. This also indicates the necessity of project work in application development. Tissenbaum et al. (2019) similarly argued that digital empowerment

involves instilling in learners the belief they can move beyond learning into meaningful action, suggesting project work as a means to achieve this. In this vein, our study engaged the participants in meaningful projects to empower them. This also explains the significant increase witnessed in the mean score before and after Course 3, which peaked at 4.22. The project work provided the participants with a valuable experience to help them creatively apply their knowledge to novel situations (Schleicher, 2018), further empowering the participants and actualising their digital creativity (Lee & Chen, 2015). These results are in line with the goal of cultivating AI-empowered, proactive citizens (Pemberton et al., 2019), who can leverage the benefits of AI and contribute to society more generally (JRC & OECD, 2021). In future offering of our programme, the level of difficulty of Course 2 is to be adjusted to suit the participants' needs better for enhancing AI empowerment. Considering the importance of project work, more reference to real-life applications can also be added to Course 2.

Table 9. Statistical results on the AI empowerment survey before and after the courses

	Before Course 1 Mean (SD)	After Course 1 Mean (SD)	After Course 2 Mean (SD)	After Course 3 Mean (SD)	*F*-value	*p*-value	Partial squared eta	Pairwise comparison
AI empowerment (max. mark = 5)	4.02 (0.49)	4.20 (0.49)	4.06 (0.50)	4.22 (0.40)	3.96	$< .05^{*}$	0.27	After Course 1 > After Course 2; After Course 2 < After Course 3; Before Course 1< After Course 3

Note. $N = 36$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

4.3. Developing ethical awareness around AI

As the focus of the programme, the participants' development in ethical awareness around AI was demonstrated both through the surveys and their self-reflective writing. Their performance in the project work further validated the growth in their awareness. Table 10 shows the means, standard deviations and paired t-test scores of the ethical consideration survey before and after Course 3, demonstrating a statistically significant increase. This shows that the participants' perceived level of their own ethical awareness was enhanced.

Table 10. Statistical results on the ethical consideration survey before and after Course 3

Ethical consideration	Before Course 3 (max. mark = 5)		After Course 3 (max. mark = 5)		Paired *t*-test
	Mean	*SD*	Mean	*SD*	
	4.07	0.36	4.22	0.37	3.09^{**}

Note. $N = 36$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

The statistical results for counting the matched keywords related to AI ethics are shown in Table 11. The statistically significant increase in the mean scores demonstrates that the participants made significant progress in understanding real-world AI examples and principles.

Table 11. Statistical results for counting matched keywords related to AI ethics based on participants' self-reflections on AI ethics before and after Course 3

Keyword category	Before Course 3 reflection on AI ethics		Before Course 3 reflection on AI ethics		Paired *t*-test
	Mean	*SD*	Mean	*SD*	
Real-world examples	0.33	0.72	3.28	2.24	7.88^{***}
Principles	0.22	0.49	0.89	1.39	2.58^{**}
All keywords	0.56	0.88	4.17	2.57	8.10^{***}

Note. $N = 36$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

The increase in participants' ethical awareness was also reflected in their project work. The ethical considerations involved in each project are listed by project in Table 12. The participants considered the possible ethical implications emerging from the design, deployment and use of the application at the initial design stage.

Table 12. Ethical considerations in group projects on "Developing Artificial Intelligence Applications"

Project no.	Ethical considerations
1	• Detect and remove hate speech and discriminatory, embarrassing or biased answers. • Protect the privacy of users' personal data; avoid asking about users' private matters. • Avoid asking users questions containing marketing messages.
2	• Require consent to collect images of English letters from others. • Accuracy of the results affects toddlers' engagement and teaching effectiveness.
3	• Some people may not possess the knowledge to check the validity of AI algorithms. • Be cautious of how to interpret the results generated by AI; regulate the parties accountable for making decisions with AI. • Require the consent of students and parents for data sharing and data exploration.
4	• Require consent to collect images of hand signals from others. • Accuracy of the system influences children's learning motivation.
5	• Privacy issues may arise in labelling people other than family members as friends or strangers. • Ensure data safety. • Safeguard home users' privacy without sacrificing too much freedom.
6	• Some people may not possess the knowledge to check the validity of AI algorithms. Information inequality arises. • There is uncertainty over who is held accountable for algorithm-based investment recommendations.
7	• Collect authorised photos to train the AI model. • Implement sufficient data protection policies; avoid sharing AI inference results for commercial purposes.
8	• Protect the data privacy of users' conversation. • Be aware of the potential discrimination portrayed in the chatbot's responses.
9	• Protect the privacy of user's body measurement data; keep user informed of how personal data are processed.
10	• Avoid potential bias in the music database, e.g., whether to include commercial music. • Protect the data privacy of users' conversation.

Table 13. Selected quotes from participants' interviews and self-reflections after Course 3 on the usefulness of the project work in developing their ethical awareness

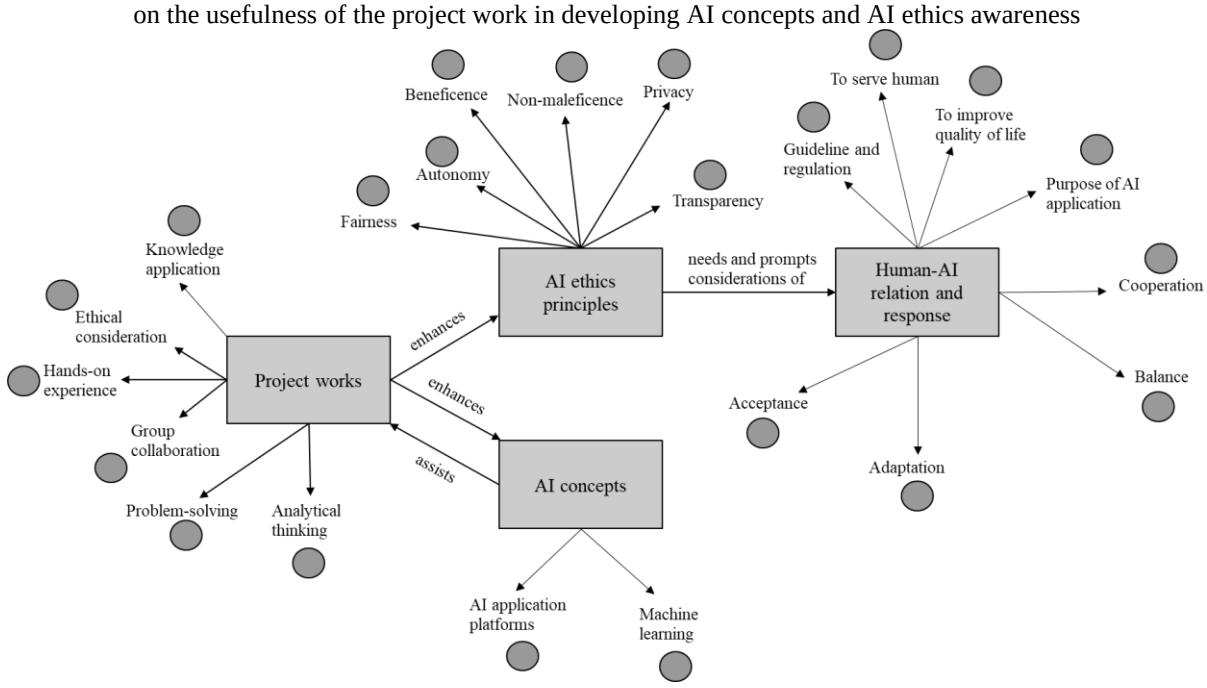
Course 3 enhanced my critical thinking skills and awareness of AI ethics. While collecting data for our project, I also reflected on the ethical debate concerning AI and gained a deeper understanding of whether the advantages outweighed the disadvantages or vice versa. This process increased my critical thinking and ethical awareness (interview; S7).
My ethical awareness was enhanced and developing such an awareness interests me a lot. As a novice language teacher, developing the general analytical skills and thinking through the ethics of AI are even more important than developing an AI model. It is important to integrate technologies in teaching, but what matters even more is ethical awareness. This enhanced my analytical skills in deciding which data to use, considering copyright and other ethical issues, which all benefit our individual development. Today, being aware of how to use AI - knowing how to cope with societal change - really matters for both teachers and students. I benefited a lot from the project work (interview; S11).
The project work in Course 3 prompted me to reflect on ethics. Before attending the courses, I thought through application development with no concern for ethics. The two cases studies in the beginning of Course 3 served as effective examples of ethical issues in application development. This course reminded me to consider the possible ethical implications involved (interview; S18).
Many machine learning models generate their results by operating on high-dimensional correlations beyond the interpretive capabilities of human-scale reasoning. In these cases, the rationale of algorithmically produced outcomes that directly affect decision subjects remains opaque to those subjects. In some applications, the processed data could cause discrimination, bias, inequity or unfairness. The opaqueness of the model may be deeply problematic. Therefore, people should pay attention to those issues before applying AI techniques (reflection; S14).

Figure 3 reports the thematic analysis of participants' reflective pieces and focus group interviews. Four themes were identified, namely "project works," "AI concepts," "AI ethics principles" and "human-AI relation and response." In the thematic map, squares represent the themes, with the codes, represented by circles, emanating

from the corresponding themes via arrows. Table 13 shows sample quotations and Appendix 1 shows the operational definition of each code.

Participants reported that the AI application development projects enabled them to apply their knowledge, have hands-on experience, consider ethical issues and train their analytical thinking, which collectively improved their understanding and awareness of AI ethical principles. Not only did participants consider principles explored in the course (namely autonomy, beneficence/non-maleficence and fairness), they also mentioned elements of AI ethics beyond the curriculum, such as transparency and privacy. Furthermore, participants were able to provide suggestions and considerations on human-AI relations and on humans' responses to the societal changes caused by AI. The finding fits with our goal of a holistic cultivation of AI literacy, which includes the ability to evaluate and reflect on AI in real-world scenarios. It also validates the feasibility and success of the novel approach to teaching AI ethics by integrating ethical considerations into project work. This approach differed from the delivery of abstract principles of AI ethics (Borenstein & Howard, 2021) as it effectively guided participants to reflect on the complex ethical concerns emerging from the design, deployment and use of AI technologies.

Figure 3. Thematic analysis of the text from participants' interviews and self-reflective writings after Course 3 on the usefulness of the project work in developing AI concepts and AI ethics awareness



5. Conclusions and implications

This study presented an evaluation of an AI literacy programme which developed university students' concepts, literacy, empowerment and ethical awareness. We conducted a 30-hour programme and piloted it with 36 university students in Hong Kong with diverse backgrounds. Our surveys, tests and self-reflective writing assignments demonstrated that the course participants felt empowered and made significant gains in their understanding of major AI concepts, literacy and ethical awareness.

One limitation of our programme is the decline in participant number from 120 in Course 1 to 82 in Course 2 and to 36 in Course 3, which may be attributed to it being non-credit-bearing and participants' other credit-bearing activities (Oakley et al., 2011). Despite this decline, participants' evaluation of each course showed their satisfaction, even when including outgoing participants, suggesting the dropout was unrelated to course quality. Another limitation is the online teaching mode under COVID-19, which potentially affected participants' learning and participation (Salas-Pilco et al., 2022).

Despite the limitations, the results of our course have several important implications. First, the study refocuses AI literacy programmes on conceptual building instead of first emphasising mathematical formulae and programming codes (Kong et al., 2021b). Teaching concepts in this way can lower the barrier and ensure equal access to AI literacy for people from all walks of life (Long & Magerko, 2020), which is a great leap in promoting AI literacy among educated citizens of diverse backgrounds.

Second, the study highlights the importance of project work in an AI literacy programme. The survey results demonstrated that Course 3 played a significant role in empowering participants. The project work unleashed their digital creativity (Lee & Chen, 2015) and deepened their conceptual learning by enabling them to creatively apply their knowledge to new contexts (see Section 4.1).

Finally, the study initiates and validates the method of teaching AI ethics through project work in application development. This approach differed from the delivery of abstract principles of AI ethics in an after-the-fact manner, and instead emphasised their importance at every stage of learning about AI (Borenstein & Howard, 2021). It effectively guided the participants to reflect on the complex ethical concerns emerging from the design, deployment and use of AI technologies.

The significance of the study lies in its validation of a pathway to develop AI literacy among educated citizens from diverse academic backgrounds. It not only contributes to the demystification of AI among the public by fostering conceptual understanding, but also cultivates AI-empowered, proactive and ethically informed citizens who can leverage the benefits of AI to contribute to society more generally.

Acknowledgement

The work described in this paper was substantially supported by a grant from the Research Grants Council, University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. EdUHK CB302), and received funding support from the Li Ka Shing Foundation.

References

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for learning, teaching, and assessing: A Revision of Bloom's taxonomy of educational objectives*. Longman.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and digital technologies. *International Journal of Information Management*, *62*, 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147. https://doi.org/10.1037/0003-066X.37.2.122

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

European Commission, Joint Research Centre (JRC) & Organisation for Economic Co-operation and Development (OECD). (2021). *AI watch, national strategies on artificial intelligence: A European perspective.* Publications Office of the European Union. https://doi.org/10.2760/069178

Fosso Wamba, S., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A Bibliometric review and research agenda. *Technological Forecasting and Social Change*, *164*, 120482. https://doi.org/10.1016/j.techfore.2020.120482

Frymier, A. B., Shulman, G. M., & Houser, M. (1996). The Development of a learner empowerment measure. *Communication Education*, *45*(3), 181-199. https://doi.org/10.1080/03634529609379048

Floridi, L., & Cowls, J. (2019). A Unified framework of five principles for AI in society. *Harvard Data Science Review*, *1*(1), https://doi.org/10.1162/99608f92.8cd550d1

Garrett, N., Beard, N., & Fiesler, C. (2020). More than "If time allows": The Role of ethics in AI education. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 272-278). https://doi.org/10.1145/3375627.3375868

Green, N. (2021). An AI ethics course highlighting explicit ethical agents. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 519-524). https://doi.org/10.1145/3461702.3462552

High Level Expert Group on Artificial Intelligence (HLEG). (2019). *Ethics guidelines for trustworthy AI*. Publications Office of the European Union. https://doi.org/10.2759/177365

Jobin, A., Ienca, M., & Vayena, E. (2019). The Global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kong, S. C., Li, P., & Song, Y. (2018). Evaluating a bilingual text-mining system with a taxonomy of key words and hierarchical visualization for understanding learner-generated text. *Journal of Educational Computing Research, 56*(3), 369-395. https://doi.org/10.1177/0735633117707991

Kong, S. C. (2021). Delivery and evaluation of an e-learning framework through computer-aided analysis of learners' reflection text in a teacher development course. *Research and Practice in Technology Enhanced Learning*, *16*, Article 28. https://doi.org/10.1186/s41039-021-00172-w

Kong, S. C., Kwok, W.Y., & Poon, C.W. (2021a). Evaluating a learning trail for academic integrity development in higher education using bilingual text mining. *Technology, Pedagogy and Education, 30*(2), 305-322. https://doi.org/10.1080/1475939X.2021.1899041

Kong, S. C., Cheung, W. M. Y., & Zhang, G. (2021b). Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. *Computers and Education: Artificial Intelligence, 2*, 100026. https://doi.org/10.1016/j.caeai.2021.100026

Kong, S. C. & Zhang, G. (2021). A Conceptual framework for designing artificial intelligence literacy programmes for educated citizens. In S. C. Kong, Q. Wang, R. Huang, Y. Li, & T. C. Hsu (Eds.), *Conference Proceedings (English Track) of the 25th Global Chinese Conference on Computers in Education, GCCCE 2021* (pp. 11-15). The Education University of Hong Kong.

Kuipers, B. (2020). Perspectives on ethics of AI: Computer science. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford Handbook of Ethics of AI* (pp. 419-441). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.013.27

Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021). Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 191-197). https://doi.org/10.1145/3408877.3432513

Lee, M. R., & Chen, T. T. (2015). Digital creativity: Research themes and framework. *Computers in Human Behavior, 42*, 12-19. https://doi.org/10.1016/j.chb.2014.04.001

Long, D. & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16), Honolulu, United States. https://doi.org/10.1145/3313831.3376727

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on bias and fairness in machine learning. *ACM Computing Surveys*, *54*(6), 1-35. https://doi.org/10.1145/3457607

National Commission for the Protection of Human Subjects (NCPHS). (1978). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.

Oakley, G., Lock, G., Budgen, F., & Hamlett, B. (2011). Pre-service teachers' attendance at lectures and tutorials: Why don't they turn up? *Australian Journal of Teacher Education*, *36*(5), 31-47. http://dx.doi.org/10.14221/ajte.2011v36n5.3

Organisation for Economic Co-operation and Development (OECD). (2018a). *Bridging the digital gender divide: Include, upskill, innovate*. OECD Publishing. http://www.oecd.org/digital/bridging-the-digital-gender-divide.pdf

Organisation for Economic Co-operation and Development (OECD). (2018b). *Future of education and skills 2030: Conceptual learning framework.* OECD Publishing. https://www.oecd.org/education/2030/Education-and-AI-preparing-for-the-future-AI-Attitudes-and-Values.pdf

Organisation for Economic Co-operation and Development (OECD). (2019). *Recommendation of the Council on Artificial Intelligence*. OECD Publishing. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Paulus, P. B., & Brown, V. R. (2003). Enhancing ideational creativity in groups: Lessons from research on brainstorming. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 110–136). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195147308.003.0006

Pemberton, D., Lai, Z., Li, L., Shen, S., Wang, J., & Hammer, J. (2019). AI or Nay-I? Making moral complexity more accessible. In *Extended abstracts of the Annual Symposium on Computer–Human Interaction in Play Companion* (pp. 281–286). https://doi.org/10.1145/3341215.3358248

Pouly, M., Koller, T., & Arnold, R. (2019). A Game-centric approach to teaching artificial intelligence. In *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)* (pp. 398–404). https://doi.org/10.5220/0007745203980404

Prunkl, C. (2022). Human autonomy in the age of artificial intelligence. *Nature Machine Intelligence*, *4*(2), 99-101. https://doi.org/10.1038/s42256-022-00449-9

Roth, K. (1990). Developing meaningful conceptual understanding in science. In B. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 139–175). Lawrence Erlbaum Associates, Inc.

Salas-Pilco, S. Z., Yang, Y., & Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A Systematic review. *British Journal of Educational Technology, 53*(3), 593-619. https://doi.org/10.1111/bjet.13190

Saltz, J., Skirpan, M., Fiesler, C., Gorelick, M., Yeh, T., Heckman, R., Dewar, N., & Beard, N. (2019). Integrating ethics within machine learning courses. *ACM Transactions on Computing Education (TOCE), 19*(4), 1-26. https://doi.org/10.1145/3341164

Schleicher, A. (2018). *World class: How to build a 21st-century school system: Strong performers and successful reformers in education*. OECD Publishing. https://doi.org/10.1787/9789264300002-en

Stadelmann, T., Keuzenkamp, J., Grabner, H., & Würsch, C. (2021). The AI-Atlas: Didactics for teaching AI and machine learning on-site, online, and hybrid. *Education Sciences, 11*(7), 318. https://doi.org/10.3390/educsci11070318

Tedre, M., Toivonen, T., Kahila, J., Vartiainen, H., Valtonen, T., Jormanainen, I. & Pears, A. (2021). Teaching machine learning in K–12 classroom: Pedagogical and technological trajectories for artificial intelligence education. *IEEE Access, 9*, 110558-110572. https://doi.org/10.1109/ACCESS.2021.3097962

Thomas, K. W., & Velthouse, B. A. (1990). Cognitive elements of empowerment: an "interpretive" model of intrinsic task motivation. *Academy of Management Review, 15*(4), 666-681. https://doi.org/10.2307/258687

Tissenbaum, M., Sheldon, J., & Abelson, H. (2019). From computational thinking to computational action. *Communications of the ACM, 62*(3), 34-36. https://doi.org/10.1145/3265747

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? In *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Nerini, F. F. (2020). The Role of artificial intelligence in achieving the sustainable development goals. *Nature Communications*, *11*(1), 233. https://doi.org/10.1038/s41467-019-14108-y

Wong, G. K. W., Ma, X., Dillenbourg, P., & Huan, J. (2020). Broadening artificial intelligence education in K-12: Where to start? *ACM Inroads*, *11*(1), 20-29. https://doi.org/10.1145/3381884

World Intellectual Property Organization (WIPO). (2019). *WIPO technology trends 2019 – Artificial Intelligence*. World Intellectual Property Organization. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf

Yi, Y. (2021). Establishing the concept of AI literacy: Focusing on competence and purpose. *Jahr - European Journal of Bioethics, 12*(2), 353-368. https://doi.org/10.21860/j.12.2.8

Appendix 1. Codes and definitions for thematic analysis

Theme	Code	Operational definition
AI Ethics Principles	Fairness	AI applications do not have prejudice, bias or stereotypes to any individuals, and their benefits and harms are equally distributed to everyone
	Beneficence	AI applications actively promote humanity's safety and well-being
	Non-maleficence	AI applications do not harm humans
	Transparency	The usage, merits and drawbacks of AI applications are clearly stated
	Autonomy	AI applications are not used to manipulate people. Humans are the ones to make decisions and be accountable
	Privacy	AI applications protect the security of people's data and do not infringe on people's privacy
Human-AI Relation and Response	Adaptation	People adapt to a world permeated with AI
	Acceptance	People accept the changes brought about by AI
	Balance	People balance the risks and benefits of AI before adopting it
	Guideline and regulation	People regulate the usage of AI and establish guidelines to manage it
	To improve quality of life	AI improves people's quality of life
	To serve human	AI serves humans as a tool
	Cooperation	AI cooperates with people and complements people's shortcomings
	Purpose of AI application	People think about the purpose of using AI before adopting it
Projects	Knowledge application	Participants apply their knowledge acquired previously
	Hands-on experience	Participants practice their knowledge hands-on
	Group collaboration	Participants collaborate and interact with groupmates
	Ethical consideration	Participants consider the ethics of their AI applications
	Problem-solving	Participants solve real-world problems with their AI applications
	Analytical thinking	Participants think analytically and critically when evaluating their AI applications
AI Concepts	AI application platforms	Participants use platforms (such as Google Teachable Machine and Microsoft Azure QnA Maker) to develop their AI applications
	Machine learning	Participants apply their knowledge of machine learning learnt in the first 2 courses