Ethics in AI Engineering Education: A Systematic Review of Innovative Curricula and Ethics Frameworks

Mrs. Clelia Perez-Medina, University of Florida

Clelia Perez-Medina is a Ph.D. student in Educational Technology at the University of Florida and a Fulbright Scholar. She holds a Master's degree in Marketing and Advertising, a specialization in Consumer Psychology, and a Bachelor's degree in Graphic Design. She has teaching experience at several accredited universities in Colombia and served (from 2022 to 2024) as the Director of Graduate Programs in Consumer Psychology. Her current research focuses on ethics in artificial intelligence in education, particularly AI ethics frameworks, AI trustworthiness, and their applications within educational settings.

Woorin Hwang, University of Florida

Woorin Hwang is a Ph.D. candidate in Educational Technology at the University of Florida. Her research is focused on integrating Artificial Intelligence (AI) in teaching and learning, with ongoing projects related to AI literacy, Internet of Things (IoT), and engineering education. Prior to joining the University of Florida, Ms. Hwang worked as a teacher and instructional designer in South Korea, where she led several projects in learning content design and the development of educational systems.

Loraine Cisternas-Garcia, University of Florida
Andrea Ramirez-Salgado, University of Florida

Andrea Ramirez-Salgado is an Instructional Assistant Professor in Engineering Education at the University of Florida, where she teaches courses in artificial intelligence and data science. Her research examines how instructional practices shape students' identities in engineering and computer science, and how these identities affect their career goals and persistence. She is particularly committed to creating inclusive, engaging learning environments that support diverse learners.

Dr. Pavlo Antonenko

Pavlo "Pasha" Antonenko is an Associate Professor of Educational Technology at the University of Florida. His interests focus on the design of technology-enhanced learning environments and rigorous mixed-method research on the effective conditions for tec

Ethics in AI for Higher Education: A Systematic Review of Ethics Interventions and Frameworks

Abstract

Artificial Intelligence in Education (AIED) has advanced rapidly, yet scholars and practitioners remain uncertain about how to embed ethical reasoning into AI driven learning contexts. To illuminate current practice, we conducted a systematic review of 25 peer reviewed studies (2021–May 2025) that explicitly taught AI ethics to higher education students. Using PRISMA and Covidence™, we extracted information on intervention types, ethics frameworks employed, evaluation methods, and learning outcomes. Most interventions were course level, delivered online or hybrid, and rooted in engineering or computer science disciplines (32%). Interestingly, only 44% of the studies explicitly tied their curricula to an established ethics framework. The most frequently referenced frameworks in the literature were the UNESCO AI Ethics Guidelines, the EU HLEG Trustworthy AI, and ACM's Code of Ethics, although each appeared in just a few studies. Evaluation methods varied, with mixed methods designs predominant; 20 of the 25 studies (80%) reported substantial improvements in ethical understanding, while 5 studies (20%) revealed gaps in depth, tooling, and institutional support. Our synthesis reveals a disciplinary enthusiasm for ethics in AI Education but a persistent shortage of coherent, philosophically grounded, and consistently applied ethics frameworks. We suggest that a shared, actionable ethics framework—grounded in philosophical traditions—would help inform curriculum design, assessment, and stakeholder engagement in AIED.

Introduction

Artificial Intelligence in Education (AIED was originally envisioned as an evolution of AI computational techniques and concepts [1] and was also established as a term for the studies of technical design and implementation of Intelligent Tutoring Systems (ITSs) research, expected to be a multidisciplinary field by 2010, connecting AI and education [2].

The necessity to incorporate artificial intelligence (AI) into education has led to rapid integration of AI tools and approaches, yet many question whether educators as well as educational researchers are employing AI ethically [3], [4], [5], [6], [7]. For instance, increased attention is devoted to the unintended consequences of AI, responsible use [8], accessibility [9], and equity [10], [11] in AI-enabled education (AIED), prompting research into criteria for designing ethical innovations.

However, these criteria introduce multiple routes that complicate a clear trajectory, making it difficult to adopt a straightforward approach to integrating ethics into AIED. Over the last years, [12] have identified the distrust generated by AI in scenarios such as producing essays containing erroneous references, fictitious authors, anomalous phrases, or disrupted grammatical sequences, as well as in situations where image creation or voice synthesis is intended to be mitigated by an integrity process termed the Interaction Layer in Education by the authors. This layer is proposed as a legitimate solution to ensure process integrity, with the primary goal of cultivating trust, thus ethics, in technology-mediated educational environments. The Interaction Layer is designed to operate without interruption (grounded in submission rules), interoperable (leveraging a functional tag created in Google Docs and submitted to Canvas), open (publicly accessible and stable tags), and independent (rendering the tags resistant to manipulation), which contributes to the ethical processes in technology. In this context, the tags facilitate to honest debates moderated by teachers, preserving the humanity framed by class-room-defined norms, while enhancing an ethical educational setting. This protocol establishes the framework that boosts AI implementation and AIED integration [12].

Some researchers suggest incorporating additional ethics instruction to improve AI education's fairness and privacy outcomes and students' self-reported intentions to act ethically [3]; others underscore the responsibility of institutions as a fundamental, shared obligation with stakeholders, promoting ethical awareness through courses, interactive learning approaches, and embedding ethics in the curriculum rather than as standalone modules [6]. Another line of inquiry establishes five key areas: stakeholders' AI education to develop shared understanding, increase effective research in both human-centric design and human-machine interaction, guaranteeing data transparency, tracking ethical concerns in inputs and outputs, and focusing on human-AI training experience use [8].

Additionally, scholars have delineated distinct avenues for an ethics personalized AIED pathways, targeting three principal stakeholder groups: students, institutions, and teachers [9]. For students, the focus is on personalized, self-paced enrichment, access to 21st-century skills, and immersive tools that boost outcomes and career readiness; for institutions, the emphasis lies on system-level gains—rapid, large-scale data analysis, streamlined teaching processes, and curricula that adapt to each student's profile; and for teachers, the benefits include data-driven personalization, engaging virtual content, and real-time feedback that allows them to tailor instruction to individual strengths and track skill mastery [9]. Furthermore, [10] advocate integrating AI with other emergent educational technologies and expanding the scope, and highlighting the relevance of AI ethics engineering curricula. [11] establish a bidirectional connection between AIED principles and Diversity, Equity and Inclusion (DEI) research, approaching it through a rigorous analysis of learners, settings, and learning outcomes. According to these authors, incorporating AIED within DEI frameworks for STEM introductory courses enhances the comprehension of DEI constructs and helps to overcome associated challenges, thereby improving performance, active learning, collaboration, and assessment. Conversely, DEI considerations in AIED assessment entail person-centered data, mitigation of biases, honest feedback, and a focus on the AI's user.

A number of significant ethical challenges have surfaced in early AIED research. [13, p. 66] identified the challenge of AI systems deciding on "the value of reason and truth"; [14] emphasized the responsible reflection of AI research toward society in both individual and professional practice; [15, p. 277] cautioned about the inclusion of moral values into the machine as "AI has lost its philosophical flavor", underscoring respect for human intelligence and concerns about classroom implications and central learning paradigms [1]; and inaccurate tutoring systems and poor conceptualizations of learning [2].

By 2000, conversations about ethical practices in AIED implementation began to develop the need to establish principles for AIED. These principles refer to the moral commitments to contribute to society, avoid harm to others, honesty, trustworthiness, fairness, property rights, respect, and confidentiality, and specific responsibilities such as dignity in both professional processes and products, evaluations of computer systems and their impacts, among others outlined in the Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct [16]. Building on this foundation, later proposals include a code of ethics for software agents inspired by scientific fiction writer Isaac Asimov's Three Laws of Robotics, particularly because a robot may not injure humanity [17].

During the 2010s, academic researchers have moved towards a more structured approach to AIED ethics trying to determine and elaborate more structured frameworks. Drachsler and

Greller developed a useful tool for implementing learning analytics (LA) framed into AI ethics parameters, namely the DELICATE checklist, an acronym for Determination, Explain, Legitimate, Involve, Consent, Anonymize, Technical, and External, for educators [18]. Some authors examine ethical problems emerged by AIs' such as social unintended impacts, job displacement, risks of superintelligence, and ethical treatment of AIs, incorporating theoretical ethics perspectives [19]. AIED research emphasizes the importance of practical ethics in Learning Analytics (LA) [20] and highlights concerns regarding ownership and responsibility among researchers and developers [21]. However, some authors described ethics in AIED as a "moral vacuum" [22, p. 1], indicating the absence of ethics focus in recent AIED publications.

During this evolution, several definitions of AIED ethics have emerged. One of them focuses on a fair understanding of aspects, both positive and negative, in the development and use of AI in different formal and informal settings [23]. [5] defines the ethics of AIED as a process of identifying and confronting any ethical consequence arising during research, development, deployment, and the engagement of communities in the integration of AI into the learning process.

By 2022, the challenges associated with AIED difficulties have been narrowed to achieving educational goals, forms of assessment, workload, learner equity, and non-discrimination, preserving autonomy, privacy, transparency, accountability, informed participation, and the design of ethically impactful activities [11]. In AIED research, trust must be built between researchers and participants, as well as between AI users and AI systems; therefore, AIED systems must be designed with consideration of the actions taken during the intervention, respect for human and non-human knowledge, democratic values, and transparent methods and critiques that contribute benefits to research [24].

In 2023, the increasing interest and significance of the topic led to the development of more than 200 governance policies and ethics guidelines for AI, the majority (65%) of which were published between 2017 and 2019. This "AI ethics boom" [25, p. 9] arises from incremental issues triggered by technological adoption, such as the need to update educational systems, broaden access to education both inside and outside academia, and reduce gender disparity in STEM fields. Only 12.5% of these guidelines addressed academic institutions, with variations by country of publication [25]. Some AIED solutions to the many ethical dilemmas have been proposed, emphasizing stakeholder collaboration, risk detection, explainability, and democratic oversight [26]. At the same time, the eXplainable AI (XAI) frameworks aim to enhance trust in AI systems through Fairness, Accountability, Transparency, and Ethics (FATE). Notably, XAI-ED focuses on developing human-centered educational AI tools [27]. By 2025, several AI ethics frameworks have been developed by large corporations, including Amazon (AWS), Anthropic, Apple, Google (Gemini), IBM, Meta (Facebook), Microsoft, Mistral, OpenAI, and xAI, as well as by professional organizations such as the EU's Ethics Guidelines for Trustworthy AI and the World Economic Forum: Advancing Responsible AI Innovation [28].

In this burgeoning body of AI-ethics frameworks, the philosophical foundations of ethics merit particular attention for this review, as do their connections to the innovations and interventions that AIED researchers pursue. For example, an anthropocentric ethical perspective has been posited as a more precise lens for human-centered design [29], which scrutinizes subject-subject and object-object relationships; this analysis delineates three categories of AI agents: developers, operators, and users. Within AIED, developers correspond to engineering students and professionals in data science, analytics, testing, robotics, etc.; operators are those who configure AI systems, embedding ethical principles into their management, and users encompass anyone who acknowledges AI's significance in their daily life. The role of users also implies the education and training necessary to comprehend AI concepts and functionalities. Additionally, [28] clarified that the label "Ethical AI" is misleading because technology lacks four essential conditions for moral action: consciousness—the prerequisite for moral experience; intentionality—the need for genuine choice; accountability—a personal sense of responsibility; and judgmental capacity—the ability to balance values, principles, and interests. The same author notes that fairness alone is insufficient for ethical assessment, and that a novel form of disappointment arises when the algorithm fabricates an illusion, rendering the interface a trap [28]. Although the algorithmic exploration lies beyond the scope of this systematic review, it establishes a baseline for future analysis of AIED ethics interventions and the foundational frameworks employed in curriculum implementation.

[30, p. 258] proposes an AIED framework for teachers comprising seven components: (a) pertaining to the philosophy of technology—overcoming metaphysics—and continuous improvement toward further automation; (b) encompassing major contemporary approaches: systemic design, individuality, and accountability; and (c) three sustainable parts: the social system and the ecological system. To guarantee ethical and moral standards in education and enhance quality, [31] present the *AIED Transformer*, a framework for Large Language Models (LLM) comprising five steps: (a) rigorous evaluation and selection of educational base tools and their associated models, such as BERT and GPT; (b) tracking data sources and early training results with the selected model to preserve ethical behavior in real-world applications; (c) guiding models gradually from simple to complex tasks, from elementary to higher grades; (d) monitoring integration of complex activities through AI self-adjustment and overall optimization; and (e) reviewing AIED models to supervise education equity, quality, and personalized learning.

Some other frameworks are developed in AIED that do not incorporate the philosophical perspective yet incorporate relevant factors pertinent to AIED ethics exploration and the pursuit of improvement. [32] establishes AI agents to assume roles as tutors, practice partners, co-creators, and personal assistants in teaching and learning processes, although the authors deliberately avoid attributing any human-like skill to the AI agents and emphasize that many users may do so. However, the research remains in a nascent stage, far from full maturity. Additionally, the authors refer to the resistance generated in the learning experience through AI as *frictions* and classify them as (a) sensorimotor and temporal, (b) interpersonal, (c) epistemic, and (d) moral frictions. In essence, educational interventions align the roles of AI agents with each category of friction, and reducing these distractions—frictions—would facilitate a successful learning process [32]. Importantly, much of the literature on AIED—including the study of frictions—remains grounded in engineering paradigms, underscoring the necessity of engaging deeply with these technical domains while simultaneously bridging them with philosophical and pedagogical perspectives to advance a more comprehensive ethics of AI in education.

In the engineering domain, [33, p. 891] propose "The Model" to mitigate bias while preserving model performance, evaluating its efficacy with Accuracy and Bias Reduction Metrics. The approach comprises dataset selection, data preprocessing, hard debiasing techniques, and model training. Although this model is not proposed for the educational sector, it underscores an ethical perspective relevant to that domain. Another study in engineering [34] investigates ethical considerations in e-learning and delineates actions that institutional administrators may adopt: (a) raise awareness of incorporating AI into education; (b) design e-learning systems aligned with ethical principles; (c) address challenges associated with AI implementation in universities; and (d) foster partnerships with companies leading AI evolution. Furthermore, drawing on general ethical guidelines, [35] develop a data-ethics framework for AIED in engineering characterized by three components: (a) data ethics—encompassing principles such as transparency, privacy, accountability, inclusiveness, and security; (b) algorithms, data, and policies—as the central element; and (c) data processing—covering collection, consumption, maintenance, and disposal. In this framework, each data-ethics component interacts with the central component, and the data-processing component engages with both the central element and its own subcomponents.

This systematic review analyzes AIED interventions -specific programs, practices, or strategies employed by researchers to test impact studies aimed at improving educational outcomes- that incorporate ethical considerations, examining how ethics are defined and integrated through frameworks into AIED, rather than merely evaluating outcomes. It focuses on higher-education students and identifies the frameworks employed, as reported in articles and conference proceedings from 2021 to May 2025. While implementation outcomes matter [36], the primary concern is the conceptualization and integration of ethics in AIED.

This study addresses the following research questions:

1. What patterns can be identified in studies with AIED ethics interventions between 2021 and 2025?
2. What are the most used AI ethics frameworks in AI education interventions?
3. To what extent ethics theories used in AI education interventions are grounded in relevant philosophical movements?

Methodology

This systematic review focuses on curricula that incorporate ethics in AI for engineering education, published in articles and conference proceedings between 2021 and May 2025 due to the rise of generative AI in 2022, and the transformation it brought in many fields, remarkably in educational institutions. The review was conducted using nine major databases: IEEE Xplore, ACM Digital Library, ProQuest, Education Source, Academic Search Premier, Applied Science and Technology Source, Information Science and Technology, Scopus, and Web of Science. The search string included "Education*" AND "ethic*" AND ("AI" OR "artificial intelligence") AND ("Teach* OR Learn*"), in titles and abstracts of the literature.

This study followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) protocol to enhance consistency, dependability, credibility, and trustworthiness [37]. To improve rigor, two authors collaborated using the collaborative literature review software Covidence™ and negotiated interpretation strategies in three sessions resulting in an interrater agreement of 0.84 (Cohen's *kappa*).

Eligibility

To be included in this review, studies had to be (1) published between January 1, 2021 and May 30, 2025; (2) peer-reviewed articles, conference proceedings and dissertations; (3) reported in English or Spanish from anywhere in the world (i.e., languages spoken by the authors); (4) empirical studies; (5) conducted in formal and informal educational settings; (6) focus on higher education students (undergraduate or graduate); and (7) include "ethic*" in either research questions or hypotheses, objectives, or study purpose.

Screening

Initially, 4,031 publications were identified across the databases. After removing 1,729 duplicates—19 manually and 1,710 using Covidence™—a total of 2,302 records remained for title and abstract screening. Of these, 1,512 were excluded for not meeting the inclusion criteria. The remaining 790 publications were retrieved in full text, including both intervention and nonintervention studies, and imported into the Zotero™ reference manager for PDF collection. Two authors then conducted a second screening round to identify studies demonstrating an explicit ethics focus on their research questions, objectives, goals, or hypotheses. This Covidence™ supported process excluded 491 studies and retained 299. A subsequent indepth review refined the scope to studies reporting interventions designed to teach AI ethics concepts to specific stakeholder groups, thereby excluding systematic, literature, and narrative reviews. Exploratory case studies were also removed to prioritize research centered on the teaching–learning process. In addition, studies conducted in informal settings—such as organizations or bootcamps—were excluded, excluding 246 studies for a total of 53 publications in different education levels. Finally, 28 K–12 implementations were removed to align with the researchers' focus on higher education. Although the primary interest lies in engineering and related fields, studies from other disciplines were retained to enable crossfield comparison within higher education. This extended refinement process reduced the initial 4,031 studies to 25 intervention-based publications for a full text review, situated in formal higher education settings and focused on the teaching–learning process, as shown in Figure 1.

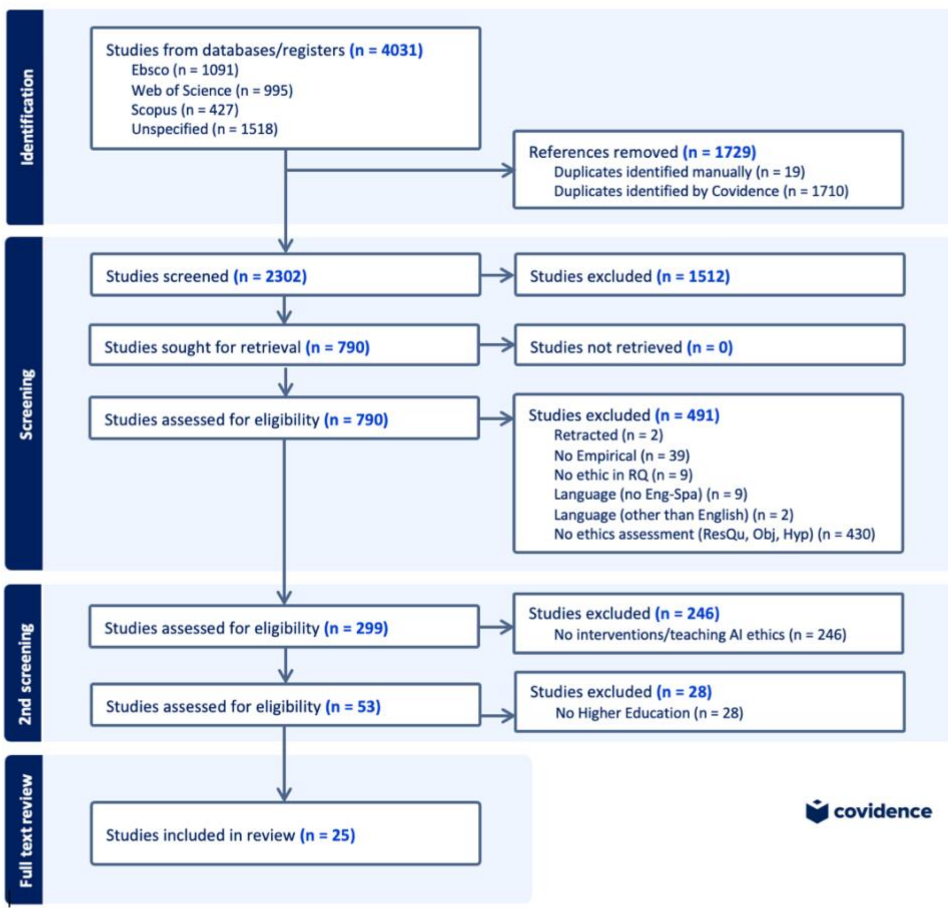

Figure 1. PRISMA protocol followed by the authors developed by Covidence™

To develop the codebook for analyzing the educational interventions, two authors conducted an in-depth analysis of 10 studies and documented the process through an audit trail. The final codebook (Appendix A) was adapted from the Institute of Education Sciences (IES) systematic framework for researchers, to integrate implementation research into high-quality impact evaluations [38], linking assessment, process and findings, in sum to the authors' defined inclusion criteria. Then, the authors focused on those studies where AIED ethics was explicitly taught during the educational interventions. Some studies include ethics as part of the findings – not as a research interest; others assessed ethics behavior in different scenarios –real or simulated, but did not actually teach any ethics aspect. This detailed review of each publication implied the constant dialogue between the authors as well as a constant evaluation of the inclusion criteria to determine which publications fit accurately for inclusion in the final selection. Finally, authors agreed that 25 publications (Appendix B) qualified for full-text review and data extraction.

Results

The selected publications (*n* = 25) were analyzed using the codebook (Appendix A) and a tracking spreadsheet. The majority of studies were published as Journal Articles (*n* = 21, 84%), while four were published in Conference Proceedings (*n* = 4, 16%). Most of the studies were published during 2024 (*n* = 13, 52%), five until May 2025 (*n* = 5, 20%) and 2023 (*n* = 5, 20%), and two during 2021 (*n* = 2, 8%).

In terms of the geographic distribution, most of the 25 studies were produced in the USA (*n* = 5, 19%). Australia, Taiwan and South Korea each produced three publications (12% each). The UK generated two publications (8%), and other publications were distributed with one per each of these countries: Tunisia, Saudi Arabia, Singapore, Indonesia, Spain, Ireland, France, Israel, Canada, and China. The eight engineering and akin studies originate from seven different countries: two generated in the United States representing 25%, and the remaining studies contribute 12.5% each, representing research precedence from Ireland, the United Kingdom, Spain, Australia, Tunisia and Israel.

All of the educational interventions were implemented in formal settings for undergraduate students (*n* = 21, 84%) and graduate students (*n* = 4, 16%). The most common modality was online (*n* = 11, 44%), then in-person (*n* = 8, 32%), and the rest were hybrid (*n* = 6, 24%). There were five different disciplines identified, most of the studies were by scholars in Engineering and Computer Sciences (*n* = 8, 32%), and Social Sciences and Humanities (*n* = 8, 32%). Six were implemented with students in a variety of disciplines categorized as Multidisciplinary (*n* = 6, 24%), Health Sciences (*n* = 5, 20%), and Business (*n* = 3, 12%).

Table 1 presents the specifications of AIED ethics interventions in the engineering field with respect to discipline, type of intervention, and ethics frameworks employed during the educational intervention. Additionally, Appendix C lists the engineering and related fields set of research questions, objectives, or purposes as documented in each publication.

Table 1. AIED ethics interventions in the Engineering field

Title	Sub-discipline	Type of intervention	Ethics frameworks used
Embedding Responsible AI in Technical Education Curriculum: A Case Study in an Asynchronous Online Advanced Data Analytics Course	Computing, data science, and related fields	Asynchronous online course integrating case studies, coding activities, and ethical reflections	Microsoft principles of responsible AI ACM FAccT (Fairness, Accountability and Transparency) Google principles of responsible AI IBM principles of responsible AI Meta principles of responsible AI UNESCO Recommendation on the Ethics of Artificial Intelligence European Commission Guidelines for Trustworthy AI
Ethics, AI, and Society: Shaping Responsible AI Engineers with CDIO Framework	Engineering	Course “Ethics, AI, and Society” designed under CDIO framework	CDIO Standards (1, 2, 3, 7, 8, 11)
The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements	STEM	Classroom module using AI Incident Database (AIID)	AIID classroom activity (incident exploration, reporting, reflection)
Harnessing generative AI in chemical engineering education: Implementation and evaluation of large language models ChatGPT v3.5	Chemical Engineering Education	Lab session integrating ChatGPT v3.5 into fluid dynamics experiment	IDEE (Identify desired outcomes, Determine level of automation, Ensure ethics, Evaluate effectiveness) FAST principles (Fairness, Accountability, Sustainability, Transparency)
Case Studies About Moral Dilemmas to Apply Ethical Theories in Engineering Education	Computer Engineering	Ethics courses (Fundamentals of Business Ethics and Innovation; Social & Environmental Aspects of Computing)	Ethics approach: students apply Kantianism, Utilitarianism, Social Contract Theory, and Virtue Ethics to real dilemmas
ETHICS4EU: Designing new curricula for Computer Science Ethics Education: Case studies for AI ethics	Computer Science	Case study-based ethics curriculum (delivered in programming module)	ProPublica COMPAS case study (Dressel & Hany, 2021)
Unpacking the role of AI ethics online education for science and engineering students	Electrical Engineering, Computer Engineering, and Information System Engineering	Online explicit-reflective learning module on AI ethics	OECD (2021) – Recommendation of the Council on Artificial Intelligence IEEE (2019) – Global Initiative on Ethics of Autonomous and Intelligent Systems

Applied Ethics via Encouraging Intuitive Reflection and Deliberate Discourse	Computer Engineering	SIMDE framework learning intervention (Social Intuitionist Model + Discourse Ethics)	Social Intuitionist Model (Haidt, 2001) Discourse Ethics (Habermas, 1990s) SIMDE framework applied to AI ethics dilemma PROPOSED Markkula Center for Applied Ethics framework (decision-making) UNESCO AI ethics video resources

Regarding the format of the educational interventions, the implementation of Courses was the most common identified (*n* = 14, 56%), Class Sessions (more than one) was the second type (*n* = 9, 36%), and Workshops and Class Activity (one session), were tied in the last place (each *n* = 1, 4%). The interventions had different durations, expressed in semesters, units, sessions, weeks, hours, and minutes. One study did not report the specific duration of their intervention. In the case of Courses, the longest lasted 14 weeks, and the shortest – 20 hours; Class Sessions ranged from a maximum of 8 hours to a minimum of 2 hours plus independent work. The Workshop was reported to last 10 hours, and there was one Class Activity of two hours long.

Only five studies used a research design that employed both a control and a treatment group (*n* = 5, 20%). The most popular research methodology was Mixed Methods (*n* = 12, 48%), Qualitative (*n* = 8, 32%), and Quantitative (*n* = 5, 20%). All studies explored Bloom's learning outcomes [39] in the Cognitive domain (*n* = 25, 100%). Additionally, almost half of the studies assessed both Cognitive and Behavioral domains (*n* = 19, 42%), and only one study focused on all three Bloom's domains, including the Affective domain (*n* = 1, 2%).

The quantity of data sources to explore the outcomes varied from one study to another. In total, 88 data sources were identified in the 25 publications under review. The most common data sources were pre- post-questionnaires (*n* = 12 and n = 17, 14% and 19% respectively), participants' reflections/discussions (*n* = 15, 17%). Table 2 shows the data sources used in the 25 publications.

Table 2. Data sources in the 25 studies under analysis

Data Collection Method (intervention groups)	Count	Percentage
Pre-questionnaire	12	14%
Post-questionnaire (feedback)	17	19%
Pre-work submissions (activities)	1	1%
Post-assessment activities	1	1%
Interview	3	3%
Focus group	3	3%
Participants' discussion forums (reflections/discussions)	15	17%
Reflective essays	1	1%
Teacher debriefings (experts / expert Delphi)	2	2%
Observations	4	5%
Video-recording	1	1%
Logs by Researcher	2	2%
Logs by Participants (self-reported data)	3	3%

Logs by Systems	3	3%
Follow-up survey	3	3%
Participants artifacts	2	2%
Participants projects	3	3%
Participants generated materials (drawing, essays, task completed)	4	5%
Lab session data analysis	1	1%
Peer assessment	4	5%
Evaluation rubrics	2	2%
Thematic analysis	1	1%
Total	88	100%

Even though all studies focused on AI ethics aspects to meet our inclusion criteria, a wide variety of ethics approaches were evident during the analysis. We approached this level of analysis from three different standpoints: (a) AI ethics frameworks mentioned either in the literature, theoretical background, introduction, or any other topic prior to the study design and implementation; (b) AI ethics frameworks explicitly used or adapted for the intervention design; and (c) explicit statement of the ethics philosophical foundation in their ethics approach employed. While these perspectives allowed us to determine the ethics foundations used in the studies, we found that some studies addressed AI ethics in their educational intervention design without explicitly mentioning any underlying ethics framework or perspective.

Regarding the (a) AI ethics frameworks in the studies' literature reviews, 50 frameworks were identified across 19 publications (76%,). Six publications did not explicitly mention any ethics framework (24%). The most common were the UNESCO AI Ethics Guidelines ($n = 4$, 8%), The European Commission's High-Level Expert Group (HLEG) on AI developed Ethics Guidelines for Trustworthy AI ($n = 3$, 6%), and three frameworks were each mentioned twice: ACM Code of Ethics & Professional Conduct (4%), UK QAA & Department for Education Guidance (4%), and the OECD Recommendation of the Council on AI (4%). Indeed, these frameworks correspond to general outlines, not specifically to the educational field. The rest of the frameworks are presented in Table 3.

Regarding the countries where these frameworks were employed, UNESCO was referenced in publications from Australia, Saudi Arabia, the UK, and Australia; HLEG in publications from Australia and South Korea; ACM in publications from Australia and Ireland-France; UK QAA in publications from the UK; and the OECD recommendations were cited in publications from Israel and Australia.

Table 3. AI ethics frameworks mentioned in the literature review section of the publications

Framework	Count	Percentage*
UNESCO AI Ethics Guidelines (UNESCO)	4	8%
The European Commission's High-Level Expert Group (HLEG) on AI developed Ethics Guidelines for Trustworthy AI	3	6%
ACM Code of Ethics & Professional Conduct (ACM)	2	4%
UK QAA & Department for Education Guidance (UK 2023)	2	4%
OECD Recommendation of the Council on AI (OECD)	2	4%
ABET guidelines (Accreditation Board for Engineering and Technology)	1	2%
Australia ACS Code of Ethics	1	2%

Microsoft Responsible AI Principles	1	2%
Accenture Responsible AI Principles	1	2%
Google Responsible AI Principles	1	2%
IBM Responsible AI Principles	1	2%
Meta Responsible AI Principles	1	2%
OpenAI Principles	1	2%
CDIO Framework (Conceive–Design–Implement–Operate)	1	2%
Academic Integrity Frameworks (Higher Education)	1	2%
Collaborative Learning Frameworks (Martinez-Maldonado et al.)	1	2%
AI-Enhanced Feedback Frameworks (Wang et al.)	1	2%
Elkhatat Framework (Elkhatat et al.)	1	2%
Chiu GenAI Competencies Framework (Chiu et al.)	1	2%
PAR Framework (Kemmis)	1	2%
Moral Reasoning (Kohlberg & McKeachie)	1	2%
Ng AI Literacy Framework	1	2%
Human-Centered AI – Moral Teaching	1	2%
Chiriatti Ethics of Authorship/Data (Floridi)	1	2%
Harvard Embedded EthiCS	1	2%
Stanford Computer Ethics	1	2%
FATE Frameworks (Fairness, Accountability, Transparency, Ethics)	1	2%
AI Incident Database (AIID)	1	2%
Hu 2024b – Thinking Aloud Pair Problem Solving (TAPPS)	1	2%
Alan Turing Institute Guidance	1	2%
IDEE (Identify desired outcomes, etc.)	1	2%
Raso Curriculum (bias awareness)	1	2%
Asilomar AI Principles	1	2%
Value-Sensitive Design (Flanagan et al.)	1	2%
Floridi's Information Ethics	1	2%
American Medical Association (AMA) Triage Ethics	1	2%
Crisis Standards of Care (CSC 2009)	1	2%
IEEE Global Initiative on Ethics of Autonomous Intelligent Systems	1	2%
PAPA Framework (Mason 1986; Parrish 2010)	1	2%
Social Intuitionist Model (Haidt)	1	2%
Discourse Ethics (Habermas)	1	2%
Deliberative Democracy (Habermas)	1	2%
Not explicitly mentioned	6	

* Percentage calculated on 19 publications out of 25 (50 frameworks).

Regarding (b), AI ethics frameworks explicitly used or adapted for the design of the educational intervention, 22 were identified in only 11 of 25 publications (*n* = 11, 44%); each of the 22 frameworks were either used or adapted only once per publication (5% each). The complete list of frameworks is provided in Table 4. Other publications relate either to specific ethics concepts, principles or aspects without specifying any formally established framework. One publication from Australia adapted seven frameworks in its intervention, one from the US used four, one from the UK used two, one from Israel also used two frameworks, and the remainder used only a single ethics framework as the primary resource for their intervention curriculum.

Table 4. AI ethics frameworks explicitly used or adapted for the curriculum intervention

Framework (name – source)	Count	Percentage
Microsoft principles of responsible AI	1	5%
ACM FAccT (Fairness, Accountability and Transparency)	1	5%
Google principles of responsible AI	1	5%
IBM principles of responsible AI	1	5%
Meta principles of responsible AI	1	5%
UNESCO Recommendation on the Ethics of Artificial Intelligence	1	5%
European Commission Guidelines for Trustworthy AI	1	5%
CDIO Standards (1, 2, 3, 7, 8, 11)	1	5%
PAR framework (Kemmis)	1	5%
AI literacy, understanding, and awareness of AI ethics.	1	5%
IDEE (Identify desired outcomes, Determine level of automation, Ensure ethics, Evaluate effectiveness)	1	5%
FAST principles (Fairness, Accountability, Sustainability, Transparency)	1	5%
CTS framework (Science, Technology, & Society)	1	5%
Crisis Standards of Care (CSC, 2009)	1	5%
OECD (2021) – Recommendation of the Council on Artificial Intelligence	1	5%
IEEE (2019) – Global Initiative on Ethics of Autonomous and Intelligent Systems	1	5%
PAPA framework (Mason, 1986; Parrish, 2010) – privacy, accuracy, property, accessibility	1	5%
Social Intuitionist Model (Haidt, 2001)	1	5%
Discourse Ethics (Habermas, 1990s)	1	5%
* SIMDE framework applied to AI ethics dilemma	1	5%
Markkula Center for Applied Ethics framework (decision-making)	1	5%
* Anti-Plagiarism Literacy Practices (APLP) programme	1	5%

* Frameworks proposed by the authors.

Finally, most of the publications lacked (c) an ethics philosophical/conceptual foundation for the AI ethics approach used in the research. Only seven (*n* = 7, 28%) included an explicit statement identifying such a philosophical foundation. These seven publications mentioned eight ethics approaches, with applied ethics being most popular (*n* = 3, 30%), as shown in Table 5.

Table 5. Ethics approaches explicitly stated in the publications

Group	Count	% of the ethics-approaches reported
Applied ethics	3	30%
Human-centric ethics learning	1	9%

Formative ethics learning	1	9%
Accountability focused ethics	1	9%
Value alignment	1	9%
Normative ethics theories	1	9%
Górniak hypothesis	1	9%
Reflexive “Principlism” [40]	1	9%
Not explicitly mentioned	18	--

We took as reference the framework created by the Institute of Education Sciences (IES) to evaluate educational interventions (IES) [38], which aim to assess the intervention fidelity – also referred as integrity by specifying components in three stages: planned, offered, and received components. We found that all studies presented directly the features or their interventions ignoring these stages, which made difficult to determine the intervention fidelity by stage, as stated by IES. Even though, all studies clearly structured the process implemented for effective interventions, without major changes between what the planned, and what they delivered to students.

Regarding the report of the results, 20 of the 25 studies identified positive outcomes in terms of improvements in ethical thinking—principles of fairness, bias recognition, and awareness; reflection on the use of AI tools to align with ethical aspects; increased engagement and reflection through structured activities; hands-on activities for AI understanding; and greater engagement with the AI-ethics topic. However, some studies revealed the need for specific tutorials on AI tool usage [41], a lack of depth beyond mere usage in AI knowledge [42], and the necessity for institutional policies [43] as well as limited change and moral tension in their initial intuitions about AI ethics [44]. Additionally, one study reported ambivalent perceptions: some participants were unaware that an avatar was AI until informed, and initial novelty diminished engagement with repeated exposure [45]. Three of the studies with modest or negative results implemented a qualitative approach [41], [42], [43], and the last two employed a mixed-methods design [44] and [45].

These study results can be grouped into four key areas: (a) the development of ethical awareness, (b) the improvement of learning readiness and feedback, (c) the identification of AI limitations through critical scrutiny, and (d) the impact on academic integrity and performance.

(a) Development of Ethical Knowledge and Awareness. Researchers consistently show that educational interventions strengthen students’ ability to recognize and reason through AI-related ethical dilemmas. Usher and Barak found that specialized online modules significantly increase awareness of issues such as data privacy and bias [6]. Shih et al. reported a positive correlation between general AI understanding and ethical awareness, suggesting that technical proficiency supports ethical foresight [46]. Feffer et al. showed that exposure to real-world AI failures via the AI Incident Database deepens students’ understanding of potential harms and the need for accountability [47]. Kwon noted that ethics education is particularly effective in cultivating responsibility and safety awareness [48]. Ajailia et al. demonstrated that embedding ethics within structured engineering frameworks improves students’ practical handling of ethical issues [49].

(b) Improvement of Learning Readiness and Feedback Efficiency. Several studies found AI tools effective as instructional companions and feedback systems. Hu showed that generative-AI debate tools enhance student readiness and confidence for in-person discussions without increasing cognitive load [50]. Akiba and Garte reported that students view LLM-generated feedback as clear, specific, and actionable, supporting iterative improvement [51]. Lee et al. concluded that automated essay scoring provides scalable formative feedback with substantial agreement to human raters [52]. Mansour and Wong found that GenAI tools increase students' confidence and knowledge regarding ethical considerations in clinical fieldwork [53].

(c) Identification of AI Limitations and Critical Scrutiny. Across studies, students appreciate AI's efficiency but quickly identify its limitations. Shin et al. reported that over 90% of nursing students perceived low reliability in AI-generated data due to scientific errors [42]. Aladsani found that postgraduate students noted accuracy problems and fabricated citations [54]. Keith et al. observed that technical-field students recognized mathematical and logical errors in AI outputs [41]. Vallis et al. showed that students frequently question the accuracy and authenticity of AI-generated content, treating AI as an assistant rather than a substitute for human judgment [45].

(d) Enhancement of Writing Quality and Academic Integrity. Studies on AI-supported academic writing report positive outcomes when AI use is framed as a "pedagogical partnership." Akiba and Garte argued that shifting grading toward the quality of revisions reduces incentives for passive plagiarism [51]. Zhang et al. [55] found that targeted AI systems improve students' "anti-plagiarism literacy," boosting confidence in identifying plagiarism and improving writing quality. Yuriev et al. showed that requiring students to document and evaluate AI-generated reports strengthens their understanding of the interplay between human expertise and AI, promoting better verification habits [56]. Abadia emphasized that although ethics training remains challenging, integrating these topics directly into technical curricula is both necessary and effective [57].

Across the four categories, clear differences emerge in the depth and breadth of learning outcomes. **Course** interventions consistently produced the most substantial gains in ethical reasoning, AI literacy, and reflective judgment, as seen in studies by [57], [49], [54], [52], [48], [58], and [59]. These multi-week structures allowed students to apply frameworks, debate dilemmas, and integrate ethics into technical practice. **Workshops**, represented by [45] generated more modest shifts, often limited to awareness and critical reflection rather than deep skill development. **Class sessions** interventions such as [6], [23], [41], [44], [46], [47], [50], [51], [55], [56], produced meaningful but narrower improvements, typically in readiness, literacy, or short-term ethical awareness. Finally, **Single session activities**, including [53], offered targeted exposure that successfully prompted ethical reflection but lacked the sustained engagement needed for deeper transformation.

The implementation of interventions guided by global ethics frameworks in engineering and related technical fields has yielded positive results in student awareness and engagement, as observed by researchers such as Abadia [57], Ajailia et al. [49], and Usher and Barak [6]. For example, Guerrero and Forment [58] found that 84.6% of students reported that applying classical ethical theories like Kantianism and Utilitarianism through roleplay was effective for learning to manage moral dilemmas and improving critical thinking. In chemical engineering studies utilizing the IDEE framework, Keith et al. [41] reported that 85% of students (combining those who "Agree" and "Strongly Agree") agreed their understanding of how to use generative AI ethically had improved. Research by Feffer et al. [47] incorporating the AI Incident Database (AIID) found that interacting with real-world failure cases increased the motivation for 66.7% of participants to study responsible AI, while shifting their value priorities toward safety and accountability. Furthermore, online modules based on OECD and IEEE guidelines implemented by Usher and Barak [6] significantly enhanced ethical knowledge, raising students' mean scores from 6.37 to 10.26 out of 12. Despite these positive educational outcomes, institutional challenges persist; as noted by O'Sullivan et al. [59], a pan-European survey found that 36% of universities still do not teach any computer ethics, and 67% of those that do dedicate 10 hours or less per semester to the subject.

Main results were extracted (Appendix D) and the text was cleaned by deleting superfluous words –non-relevant– to refine the corpus. The optimized corpus was then employed to construct a word cloud (Figure 2) reflecting three key themes: *awareness*, *improved*, and *ethical*, thereby underscoring the central focus on learning and ethical development. Additional salient terms include *literacy*, *reflection*, *critical*, *reasoning*, *enhanced*, *design*, *skills*, and *engagement.*

Figure 2. A word cloud of the key results evidenced on publications

Discussion

This systematic review was driven by the need to identify the ethics frameworks referred to and used in AIED higher education, particularly in the design of engineering education innovations. Results demonstrate a growing interest in AI ethics among engineering education scholars and curriculum designers, with most studies published during 2024 (n = 12, 52 %), reflecting AIED’s growing concerns with AI ethics in curricula [3], [4], [5], [6], [7].

Alshanetsky promotes the use of the protocol called Interaction Layer in Education to cultivate trust in technology-mediated educational environments throughout honest and contributive debates under teachers' norms in the classroom [12]. Results have shown that 15 data sources (17%) have included participants' discussion forums to scaffold students' reflections and conversations regarding AI ethics in their curricula. Additionally, post-questionnaires (n = 17, 19%) assessed reflections on both the intervention and the inclusion of AI ethics. Other studies employed teacher debriefings, interviews, and focus groups that assessed the integrity of the AI ethics intervention implemented.

Most of the studies included AI ethics into the curriculum through Courses (n = 14, 56%) instead of specific isolated modules, in line with Usher's approach [6]. Moreover, all studies (n = 25, 100%) evaluated the cognitive domain when examining participants' understanding of AI ethics, and only 19 (42%) focused on what [8] calls the *human-AI training experience* implying the behavioral domain.

Even though several authors declare the relevance of incorporating institutional stakeholders' responsibility in promoting ethical use of AI [6], [8], [9], the studies reviewed by us only focused on the relationship between teachers and learners, not including additional educational stakeholders like administrators, community members, and society.

Tapalova et al. claim that engaging online content favor the personalized learning in AIED [9]. In this review, we found that most of the studies were implemented through online content (n = 11, 44 %). In line with Drozdek and the reported loss of the "philosophical flavor" in AI evolution [15, p. 277], our results show that only seven studies (28 %) explicitly ground their work in philosophical approaches in their literature. Specifically, these authors referred to eight philosophical underpinnings in their studies with the most prevalent being Applied ethics. Others mentioned only once were Human centric ethics learning, Formative ethics learning, Accountability focused ethics, Value alignment, and Normative ethics theories. The ethical challenges highlighted by [13], [14], [15], appear to be present.

This systematic review was conducted to identify ethics frameworks used in engineering education research on AIED and to avoid the "moral vacuum" in AIED [22]. We have evidenced that all studies integrate AI into the learning process to confront ethical consequences during AI research, development, and deployment [5]. We are beginning to see AI ethics frameworks incorporated into educational interventions, which is encouraging for the field; however, these frameworks are still emerging and will require time to fully propagate. The "AI ethics boom" suggested by [25] seems to bring an explosion of ethics guidelines, not only for AI but for AIED as well.

Conceptually, we have identified a number of general AI ethics frameworks such as XAI, FATE [27], and those from large corporations [28], as well as frameworks generated by AIED scholars [30], [31] [32], [33], [34], [35]. Out of 50 frameworks mentioned in the 25 studies we reviewed, only the UNESCO AI Ethics Guidelines were referred to four times (8 %), The European Commission's High-Level Expert Group (HLEG) on AI developed Ethics Guidelines for Trustworthy three times (6 %), and ACM, UK guidance, and OECD frameworks each used only twice (4 %). The remaining frameworks were mentioned only once, which highlights the need for more usage of well-defined frameworks, and we are paving the way to that bright future. Should we, as a scholarly community, work toward greater coherence by concentrating on a few core AIED ethics frameworks? We need to keep examining this issue, as an excess of AI ethics frameworks may push the educational field into a poorly mapped domain.

As scholars, we should be less lax when making assumptions related to ethics, even if it is a recent hot topic and an ongoing concern. In fact, we should be capable of assessing ethics throughout any intervention, with the responsibility of conducting a prior review of AIED frameworks, and being explicit about which specific ethical approach was applied (transparency, accountability, fairness, etc.), rather than drawing conclusions about the 'ethics' concept as a whole, which creates ambiguity and does not help advance research or curriculum design.

The findings of this study aim to identify the AI ethics frameworks used in university-level engineering education AI ethics interventions. Our data reveal that stakeholders are enthusiastic about the inclusion of AI in educational environments within a framework of ethics principles, yet there is at the same time a lack of depth and precision in ethics definitions from a philosophical and conceptual standpoint. These findings suggest that experiential, theory-grounded AI curricula can advance technical competence and ethical judgment but must also address the identified gaps in AIED ethics framework structure, institutional support, and integral evaluation of AIED ethics, rather than relying solely on cherry-picked ethics principles.

The ethics principles that shape the actions and decisions impacting large numbers of education stakeholders continue to be explored as we develop this systematic review. Meanwhile, several frameworks have been developed by different organizations and entities, prompting several essential questions for all of us in education: Are we truly prepared to engage in a deep, meaningful dialogue about what we seek from AI? What constitutes a practical path to follow when integrating AI into education within an ethics framework? How do we, as researchers, dive into our community's behavior from a more robust ethical perspective? The critical outcome of this systematic review lies in the limited synergy between the design and assessment of AIED curricula focusing on AI ethics and the three major ethical traditions –virtue ethics, deontology, and utilitarianism– which should be explored further in future research on ethics in AI and in the AIED. What we envision as a main contribution is that the critical path forward begins with our collective commitment to broadening our perspective and strengthening the foundational principles that will allow the field to grow a strong trunk with deeper, more resilient roots, ideally before another technological hype diverts our attention. The room is now open to compare these many AI ethics frameworks, identify more robust commonalities, and build more useful and usable AI ethics models that are more coherent, actionable, and adaptable to AIED contexts.

References

[1] G. Cumming, “Artificial Intelligence in education: An exploration,” *Journal of Computer Assisted Learning*, vol. 14, no. 4, pp. 251–259, 1998. Accessed July, 15, 2025, doi: 10.1046/j.1365-2729. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2729.1998.1440251.x
[2] G. Cumming, and A. McDougal, “Mainstreaming AIED into Education?,” *International Journal of Artificial Intelligence in Education*, vol. 11, pp. 197-207, 2000. Accessed July, 15, 2025. [Online]. Available: https://telearn.hal.science/hal-00197331v1/document.
[3] A. Dahiya, K. Deepa. “Cultivating Ethical AI Practices: The Impact of AI Ethics Education on Responsible AI.” *Journal of Informatics Education and Research*, vol. 5, no. 2, 2025. Accessed July, 15, 2025. [Online]. Available: https://doi.org/10.52783/jier.v5i2.3073
[4] P. Palomino-Flores, R. Cristi-López and D. Paul, “Mastering Ethical Horizons: Exploring AI Integration in Advanced Studies of Engineering, Technology, and Informatics,” in *2024 IEEE World Engineering Education Conference (EDUNINE)*, Guatemala City, Guatemala, 2024, pp. 1-6. doi: 10.1109/EDUNINE60625.2024.10500654.
[5] W. Holmes, and K. Porayska-Pomsta, *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*, 1st ed. Routledge, 2022. [Online]. Available: https://doi.org/10.4324/9780429329067. Accessed: December, 22, 2025.
[6] M. Usher, M. Barak, and S. Erduran, “What role should higher education institutions play in fostering AI ethics? insights from science and engineering graduate students,” *International Journal of STEM Education*, vol. 12, no. 1, p. 51, 2025. Accessed December 22, 2025, doi: https://doi.org/10.1186/s40594-025-00567-x
[7] A. Nguyen, H. Ngo, Y. Hong, B. Dang, and B. Nguyen, “Ethical principles for artificial intelligence in education,” *Education and Information Technologies,* vol. 28, no. 4, pp. 4221-4241, 2023. doi: https://doi.org/10.1007/s10639-022-11316-w.
[8] J. B. Lyons, K. Hobbs, S. Rogers, S. H. Clouse, “Responsible (use of) AI,” *Frontiers in Neuroergonomics*, vol. 4, no. 1201777, 2023, Accessed July, 18, 2025. [Online]. Available: https://doi.org/10.3389/fnrgo.2023.1201777.
[9] Tapalova, Olga, and Nadezhda Zhiyenbayeva. “Artificial Intelligence in Education: AIEd for Personalised Learning Pathways,” *Electronic Journal of E-Learning*, vol. 20, no. 5, pp. 639-653, 2022. [Online]. Available: https://doi.org/10.34190/ejel.20.5.2597.
[10] C. Liu. “A comprehensive review of applications of AI technologies in higher engineering education,” *Discover Education*, vol. 4, no. 528, 2025. [Online]. Available: https://doi.org/10.1007/s44217-025-00954-0
[11] R. D. Roscoe, S. Salehi, N. Nixon, M. Worsley, C. Piech, and R. Luckin, “Inclusion and Equity as a Paradigm Shift for Artificial Intelligence in Education” in *Artificial Intelligence in STEM Education: The Paradigmatic Shifts in Research, Education, and Technology*, F. Ouyang, J. Pengcheng, B.M. McLaren, and A.H. Alavi, 1st. Ed, vol. 1. CRC Press, 2022, ch. 23, pp. 359-374. Available: https://doi.org/10.1201/9781003181187.
[12] E. Alshanetsky, “When AI Dissolves Trust: Education Can Pioneer New Infrastructure,” *Society,* vol. 62, pp. 982-987, October 2025, Accessed January, 3, 2025. [Online]. Available: https://doi.org/10.1007/s12115-025-01149-x.
[13] M.A. Boden, “Impacts of artificial intelligence,” *Futures*, vol. 16-1, pp. 60–70, 1984. Accessed July 16, 2025. [Online]. Available: https://doi.org/10.1016/0016-3287(84)90007-7.

[14] J. Beusmans, and K. Wieckert, “Computing, research, and war: If knowledge is power, where is responsibility?,” *Communications of the ACM*, vol. 32, no. 8, pp. 939–951, 1989. Accessed July 15, 2025. [Online]. Available: https://doi.org/10.1145/65971.65973.
[15] A. Drozdek, “Moral dimension of man and artificial intelligence,” *AI & Society,* vol. 6, pp. 271–280, 1992. Accessed July 16, 2025. [Online]. Available: https://doi.org/10.1007/BF02472799.
[16] R. Aiken, and R. Epstein, “Ethical guidelines for AI in education: Starting a conversation,” *International Journal of Artificial Intelligence in Education*, vol. 11, pp. 163-176, 2000. Accessed July 15, 2025. [Online]. Available: https://url-shortener.me/6XS6.
[17] C. Dowling, “Intelligent Agents in an e-literate society: Some ethical considerations,” *IFIP Advances in Information and Communication Technology*, pp. 85–92, 2003. Accessed July 15, 2025.[Online]. Available: https://doi.org/10.1007/978-0-387-35663-1_8.
[18] H. Drachsler, and W. Greller, Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics, in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK’16). Association for Computing Machinery*, New York, NY, USA, 89–98, 2016, https://doi.org/10.1145/2883851.2883893.
[19] E. Burton, J. Goldsmith, S. Koenig, B. Kuipers, N. Mattei, and T. Walsh, “Ethical considerations in Artificial Intelligence Courses,” *AI Magazine*, vol. 38, no. 2, pp. 22–34, 2017. Accessed July 15, 2025. [Online]. Available https://doi.org/10.1609/aimag.v38i2.2731.
[20] K. Kitto, and S. Knight, “Practical ethics for building learning analytics,” *British Journal of Educational Technology*, vol. 50, no. 6, pp. 2855–2870, 2019. Accessed July 22, 2025. [Online]. Available: https://doi.org/10.1111/bjet.12868.
[21] S. Hrastinski, A.D. Olofsson, C. Arkenback, S. Ekström, E. Ericsson, G. Fransson, J. Jaldemark, T. Ryberg, L.M. Öberg, A. Fuentes, U. Gustafsson, N. Humble, P. Mozelius, M. Sundgren, and M. Utterberg, “Critical imaginaries and reflections on Artificial Intelligence and robots in postdigital K-12 Education,” *Postdigital Science and Education*, vol. 1, no. 2, pp. 427–445, 2019. Accessed July 15, 2025. [Online]. Available: https://doi.org/10.1007/s42438-019-00046-x.
[22] W. Holmes, D. Bektik, B. Woolf, and R. Luckin, Ethics in AIED: Who cares?, in *20th International Conference on Artificial Intelligence in Education (AIED'19)*, pp. 25–29, June 2019, Chicago. Available: https://oro.open.ac.uk/60361/
[23] L. J. Wiese, I. Patil, D. S. Schiff, A. J. Magana, *“AI ethics education: A systematic literature review*,” Computers and Education: Artificial Intelligence, vol. 8, 2025. Available: https://doi.org/10.1016/j.caeai.2025.100405.
[24] A. Fox, “Educational Research and AIED: Identifying Ethical Challenges,” in *The Ethics of Artificial Intelligence in Education,* 1st ed., edited by W. Holmes and K. Porayska-Pomsta. Routledge, 2023, ch. 2, pp. 47-73. Available: https://doi.org/10.4324/9780429329067-4.
[25] N.K. Corrêa, C. Galvão, J.W. Santos, C. Del Pino, E.P. Pinto, C. Barbosa, D. Massmann, R. Mambrini, L. Galvão, E. Terem, and N. de Oliveira, “Worldwide AI ethics: A review of 200 guidelines and recommendations for AI Governance,” *Patterns*, vol. 4, no. 10, 2023. Accessed July 16, 2025. [Online]. Available: https://doi.org/10.1016/j.patter.2023.100857.
[26] I. Dekker, B. Bredeweg, W. Winkel, and I. van de Poel, “Ethical procedures for responsible experimental evaluation of AI-Based Education Interventions,” *AI and Ethics*, vol. 5, no. 3, pp. 2977–2986, 2024. Accessed July 15, 2025. [Online]. Available: https://doi.org/10.1007/s43681-024-00621-4.

[27] H. Khosravi, S. Buckingham Shum, G. Chen, C. Conati, Y. Tsai, J. Kay, S. Knight, R. Martinez-Maldonado, S. Sadiq, and D. Gašević, "Explainable Artificial Intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 3, 2022. Accessed July 15, 2025. [Online]. Available: https://doi.org/10.1016/j.caeai.2022.100074.
[28] G. Bronner-Martin. "Ethical AI is Impossible." *Philarchive*. 2025. Accessed December 3, 2025. [Online]. Available: https://philpapers.org/archive/BROEAI-5.pdf.
[29] E. Koval, N. Zhadunova and M. Martynova, "Artificial Intelligence Ethics: Anthropocentric Approach," in *IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, Yekaterinburg, Russian Federation, 2025, pp. 45-48, doi: 10.1109/USBEREIT65494.2025.11054203.
[30] S. T. H., Pham. "Get Ready for Future Education: A Phenomenological Study on Artificial Intelligence," (2021). Electronic Theses and Dissertations. 362. [Online]. Available: https://scholarworks.sfasu.edu/etds/362
[31] Y. Yan, H. Liu, "Ethical framework for AI education based on large language models," *Education and Information Technologies*, vol. 30, pp. 10891–10909, 2025. Accessed December 22, 2025. [Online]. Available: https://doi.org/10.1007/s10639-024-13241-6
[32] A. Peeters, M. Chichirău, T.N. Coggins, and S. Thill, "Disembodied Learning: A Critical Perspective on Flourishing with AI in Education," *Cristina Costescu, Shaping Children's Learning Through Technology. Cham: Springer*. 2025. Accessed November 22, 2025. [Online]. Available: https://philpapers.org/archive/PEEDLA.pdf.
[33] S. Ajith, M. Rithani and R. S. SyamDev, "Identifying and Mitigating Gender Bias in Language Models: A Fair Machine Learning Approach," in *Seventh International Conference on Image Information Processing (ICIIP)*, Solan, India, 2023, pp. 888-893, doi: 10.1109/ICIIP61524.2023.10537623.
[34] L. Homayouni, Y. Hejazi and N. Zarifsanaiey, "A Review of Ethical Considerations in Using Artificial Intelligence in E-Learning," in *11th International and the 17th National Conference on E-Learning and E-Teaching (ICeLeT),* Isfahan, Iran, Islamic Republic of, 2024, pp. 1-6, doi: 10.1109/ICeLeT62507.2024.10493100.
[35] Y. Hong, A. Nguyen, B. Dang and B. -P. T. Nguyen, "Data Ethics Framework for Artificial Intelligence in Education (AIED)," *International Conference on Advanced Learning Technologies (ICALT),* Bucharest, Romania, 2022, pp. 297-301, doi: 10.1109/ICALT55010.2022.00095.
[36] M.T. Schultes, "An introduction to implementation evaluation of school-based interventions," *European Journal of Developmental Psychology*, vol. 20, no. 1, pp. 189–201, 2021. Accessed July 17, 2025. [Online]. Available: https://doi.org/10.1080/17405629.2021.1976633.
[37] M. Page, J. Tetzlaff, and D. Moher, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *Value in Health*, vol. 23, no. 10, pp. S312–S313, May 2021, Accessed July 17, 2025. [Online]. Available: https://doi.org/10.1016/j.jval.2020.04.1154.
[38] C. J. Hill, L. Scher, J. Haimson, and K. Granito, *"Conducting implementation research in impact studies of education interventions: A guide for researchers,"* NCEE 2023-005, Washington, DC: U.S. Dept. of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2023. Accessed May, 2025 [Online]. Available: https://ies.ed.gov/ies/2025/01/conducting-implementation-research-impact-studies-education-interventions-guide-researchers

[39] Bloom, Benjamin S., ed. *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: The Cognitive Domain*. New York: David McKay Company, 1956.
[40] J. Beever and A. O. Brightman, "Reflexive Principlism as an Effective Approach for Developing Ethical Reasoning in Engineering," *Science and Engineering Ethics*, vol. 22, no. 1, pp. 275–291, Feb. 2016. Accessed January, 2026. [Online]. doi: 10.1007/s11948-015-9633-5.
[41] M. Keith, E. Keiller, C. Windows-Yule, I. Kings, P. Robbins, "Harnessing generative AI in chemical engineering education: Implementation and evaluation of the large language model ChatGPT v3.5", *Education for Chemical Engineers*, vol. 51, Pages 20-33, 2025. Accessed January, 2026. [Online]. Available: https://doi.org/10.1016/j.ece.2025.01.002.
[42] H. Shin, J. De Gagne, S. Kim, M. Hong, "The Impact of Artificial Intelligence-Assisted Learning on Nursing Students' Ethical Decision-making and Clinical Reasoning in Pediatric Care: A Quasi-Experimental Study". *Computers, informatics, nursing*. 2024. Accessed January, 2026. [Online]. doi: 10.1097/CIN.0000000000001177.
[43] M. Veras, J.O. Dyer, H. Shannon, B.J.M. Bogie, M. Ronney, H. Sekhon, D. Rutherford, P.G.B., Silva, D. Kairy, "A mixed methods crossover randomized controlled trial exploring the experiences, perceptions, and usability of artificial intelligence (ChatGPT) in health sciences education," *Digit Health*. December, 2024. Accessed June, 5, 2025. [Online]. doi: 10.1177/20552076241298485.
[44] L. Wiese, and A. Magana, "Applied Ethics via encouraging intuitive reflection and deliberate discourse", *2024 ASEE Annual Conference*. June, 23, 2024. Accessed June, 5, 2025. [Online]. doi https://doi.org/10.18260/1-2--46589
[45] C., Vallis, S., Wilson, D. Gozman, et al. "Student Perceptions of AI-Generated Avatars in Teaching Business Ethics: We Might not be Impressed," *Postdigital Science and Education*, vol. 6, pp. 537–555, 2024. Accessed June, 5, 2025. [Online]. doi: https://doi.org/10.1007/s42438-023-00407-7
[46] Shih, P.-K.; Lin, C.-H.; Wu, L.Y.; Yu, C.-C. Learning Ethics in AI—Teaching Non-Engineering Undergraduates through Situated Learning. Sustainability 2021, 13, 3718. https://doi.org/10.3390/su13073718
[47] M. Feffer, N. Martelaro, and H. Heidari, "The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements," in Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '23), Boston, MA, USA, Oct.–Nov. 2023, pp. 1–11.
[48] J. Kwon, "A Study on Ethical Awareness Changes and Education in Artificial Intelligence Society," Revue d'Intelligence Artificielle, vol. 37, no. 2, pp. 341–345, 2023.
[49] N. Ajailia, M. Zerai, N. Debbabi, A. Damergi, and M. Dchicha, "Ethics, AI, and Society: Shaping Responsible AI Engineers with CDIO Framework," in Proceedings of the 20th International CDIO Conference, Tunis, Tunisia, Jun. 2024, pp. 253–262.
[50] Y.-H. Hu, "Advancing Asynchronous Pre-Class Learning in Flipped Classrooms: Generative AI Companions in Business Ethics," Education and Information Technologies, Jan. 2025.
[51] D. Akiba and R. Garte, "Leveraging AI Tools in University Writing Instruction: Enhancing Student Success While Upholding Academic Integrity," Journal of Interactive Learning Research, vol. 35, no. 4, pp. 467–480, 2024.

[52] I. Lee, S. Ali, H. Zhang, D. DiPaola, and C. Breazeal, “Developing Middle School Students’ AI Literacy,” in Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE ’21), Virtual Event, USA, Mar. 2021, pp. 191–197.
[53] T. Mansour and J. Wong, “Enhancing Fieldwork Readiness in Occupational Therapy Students with Generative AI,” Frontiers in Medicine, vol. 11, Oct. 2024.
[54] H. K. Aladsani, “Developing Postgraduate Students’ Competencies in Generative Artificial Intelligence for Ethical Integration into Academic Practices: A Participatory Action Research,” Interactive Learning Environments, Apr. 2025. doi: 10.1080/10494820.2025.2487517.
[55] Y. Zhang, Y. Liu, and X. Wang, “Enhancing Anti-Plagiarism Literacy Practices Among Undergraduates with AI,” Interactive Learning Environments, Apr. 2025. doi: 10.1080/10494820.2025.2492783.
[56] E. Yuriev, M. G. Burton, D. Knott, A. E. Terrill, N. R. C. Jackson, and S. Chen, “Integrating Generative AI in Chemistry Education: Enhancing Career-Ready Writing Skills in Pharmaceutical Science,” Journal of Chemical Education, vol. 102, pp. 1991–2001, Apr. 2025.
[57] R. R. Abadia, “Embedding Responsible AI in Technical Education Curriculum: A Case Study in an Asynchronous Online Advanced Data Analytics Course,” CEUR Workshop Proceedings, vol. 3901, pp. 67–80, 2024.
[58] M. J. C. Guerrero and M. Alier Forment, “Case studies about moral dilemmas to apply ethical theories in engineering education,” *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, vol. 19, pp. 1–14, 2024, doi: 10.1109/RITA.2024.3368609.
[59] D. O’Sullivan *et al.*, “ETHICS4EU: Designing new curricula for computer science ethics education: Case studies for AI ethics,” in *Proc. EAEEIE/ETHICS4EU Conf.*, 2023, pp. 1–11. doi: 10.1145/3617694.3623223.

Appendix A. Codebook structured by the authors, adapted from the IES framework for implementation research [38] and own authors criteria.

Primary information	Classification	Description
Title		
Authors		
Publication type	Article; Conference proceeding	
Journal/Publication name		
Publication year		
Setting	Formal; Informal	
Modality	Virtual; In-person; Hybrid	
Discipline	Computing science; Social and Behavioral Sciences; Education; Multidisciplinary; Business; Humanities; Engineering; Health Sciences	
Ethics frameworks referred		Frameworks explicitly mentioned either in the theoretical approach
Ethics frameworks used		Frameworks used or adapted for the intervention curriculum
Ethics approach		Explicit approach of an ethics perspective
Pedagogical approach	Instructor Lecture; Role-play debates; Scenario based assessments (SBA) – simulations, dilemma cases; Co-creation activities; Hands-on activities (AI algorithm exploration/development)	
Type of intervention	1 = Class activity (one session) 2 = Class sessions (more than one) 3 = Workshop (not developed during a class) 4 = Course	
Study Design	Qualitative; Quantitative; Mixed-methods.	
Scale		Specific validated scales (if any) used to measure the intervention
Domain of dependent measurement	Cognitive; Behavioral; Affective	
Institute of Education Sciences (IES) criteria [X -2]		
Context	Country	
	Level of Education	Educational level: 1 = Higher (undergraduate) 2 = Higher (graduate)
	Unit (participants)	1 = Teachers; 2 = Students; 3 = Administrative & Staff; 4 = Employees; 5 = Community; 6 = Other
	Component	Short description of the core intervention
	Direct (D) and Support (S) components	(D): involve direct interaction with the intended population (S): Strategies, resources, or preparation that activate or support the direct components (professional development, training, coaching, or peer learning; supervision, monitoring, or feedback; setting up and using record-keeping or data systems [21]

Primary information	Classification	Description
	Duration	Defined in quantity of semesters, units, sessions, weeks, hours, minutes
Planned	Content; Quantity; Mode; Quality	As proposed by the publications' authors
Offered	Content; Quantity; Mode; Quality	As described by the publication authors
Received	Content; Quantity; Mode; Quality	As described in the results
Intervention fidelity	High; Medium; Low	Refers to the correspondence between the intervention components that were planned and the intervention components that were offered [21]
Data sources to measure intervention contrast	Pre-questionnaire; Post-questionnaire; Lag Sequential Analysis (LSA); Pre-focus group; Pre-drawing assessment; Pre-work submissions (activities); Post-assessment activities; Post-focus group; Long-term follow-up interviews; Interview; Focus group; Participants reflections & discussions; Reflective essays; Experts analysis; Observations; Audio-recordings; Video-recording; Logs by Researchers, Participants or Systems; Follow-up survey; Participants artifacts, projects & generated materials (drawing, essays, task completed); Peer assessment; Evaluation rubrics; Thematic analysis;	Data collection used to assess and evaluate the intervention

Appendix B. List of publications, authors, publication type, country and year

Id	Title	Author(s)	Pub. Type	Country	Year
1	Embedding Responsible AI in Technical Education Curriculum: A Case Study in an Asynchronous Online Advanced Data Analytics Course	Abadia, R.	Article	Australia	2024
2	Ethics, AI, and Society: Shaping Responsible AI Engineers with CDIO Framework	Ajailia, N., Zerai, M., Debbabi, N., Damergi, A., Dchicha, M.	Conference proceeding	Tunisia	2024
3	Leveraging AI Tools in University Writing Instruction: Enhancing Student Success While Upholding Academic Integrity	Akiba, D., & Garte, R.	Article	USA	2024
4	Developing postgraduate students' competencies in generative artificial intelligence for ethical integration into academic practices: a participatory action research	Aladsani, H. K.	Article	Saudi Arabia	2025
5	A Human-Centric Automated Essay Scoring and Feedback System for the Development of Ethical Reasoning	Lee, A. V. Y., Luco, A. C., & Tan, S. C.	Article	Singapore	2023
6	Developing AI Literacy of Management Students using Problem and Project based Learning	Chen, Y. Y., Qin, Z.	Conference proceeding	UK	2023
7	STEM-based Artificial Intelligence Learning in General Education for Non-Engineering Undergraduate Students	Lin, C.-H., Yu, C.-C., Shih, P.-K. & Wu, L.-Y.	Article	Taiwan	2021
8	Exploring the Use of Generative AI in Student-Produced EFL Podcasts: A Qualitative Study	Baskara, FX. R., Dewangga Puri, A., & Laos Mbato, C.	Article	Indonesia	2024
9	The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements	Feffer, M., Martelaro, N., & Heidari, H.	Conference proceeding	USA	2023
10	Advancing asynchronous pre-class learning in flipped classrooms: Generative AI companions in business ethics	Hu, Y. H.	Article	Taiwan	2025
11	Harnessing generative AI in chemical engineering education: Implementation and evaluation of large language models ChatGPT v3.5	Keith, M., Keiller, E., Windows-Yule, C., Kings, I., & Robbins, P.	Article	United Kingdom	2025
12	A Study on Ethical Awareness Changes and Education in Artificial Intelligence Society	Kwon, J.	Article	South Korea	2023
13	Development of Convergence Course of Artificial Intelligence and Psychology Applying Team Teaching Method	Lee, K., & Park, H.-Y.	Article	South Korea	2024
14	Case Studies About Moral Dilemmas to Apply Ethical Theories in Engineering Education	Guerrero, M. J. C., & Forment, M. A.	Article	Spain	2024
15	Enhancing fieldwork readiness in occupational therapy students with generative AI	Mansour, T., & Wong, J.	Article	USA	2024
16	ETHICS4EU: DESIGNING NEW CURRICULA FOR COMPUTER SCIENCE ETHICS EDUCATION: CASE STUDIES FOR AI ETHICS	O'Sullivan, D., Gibson, J. P., Curley, A., Becevel, A., Murphy, E., & Gordon, D.	Article	Ireland, France	2023

Id	Title	Author(s)	Pub. Type	Country	Year
17	Learning Ethics in AI—Teaching Non-Engineering Undergraduates through Situated Learning	Shih, P.-K., Lin, C.-H., Wu, L. Y., & Yu, C.-C.	Article	Taiwan	2021
18	The Impact of Artificial Intelligence-Assisted Learning on Nursing Students' Ethical Decision-making and Clinical Reasoning in Pediatric Care: A Quasi-Experimental Study	Shin, H., De Gagne, J. C., Kim, S. S., & Hong, M.	Article	South Korea, USA	2024
19	Ethics Crisis Standards of Care Simulation	Switzer, D. F., & Knowles, S. G.	Article	USA	2024
20	Unpacking the role of AI ethics online education for science and engineering students	Usher, M., & Barak, M.	Article	Israel	2024
21	Student Perceptions of AI-Generated Avatars in Teaching Business Ethics: We Might not be Impressed	Vallis, C., Wilson, S., Gozman, D., & Buchanan, J.	Article	Australia	2024
22	A mixed methods crossover randomized controlled trial exploring the experiences, perceptions, and usability of artificial intelligence (ChatGPT) in health sciences education	Veras, M., Dyer, J. O., Shannon, H., Bogie, B. J. M., Ronney, M., Sekhon, H., Rutherford, D., Silva, P. G. B., & Kairy, D.	Article	Canada	2024
23	Applied Ethics via Encouraging Intuitive Reflection and Deliberate Discourse	Wiese, L., & Magana, A. J.	Conference proceeding	USA	2024
24	Integrating Generative AI in Chemistry Education: Enhancing Career-Ready Writing Skills in Pharmaceutical Science	Yuriev, E., Burton, M. G., Knott, D., Terrill, A. E., Jackson, N. R. C., & Chen, S.	Article	Australia	2025
25	Enhancing anti-plagiarism literacy practices among undergraduates with AI	Zhang, Y., Liu, Y., & Wang, X.	Article	China	2025

Appendix C. Engineering and related fields publications with type of intervention and ethics frameworks.

Title	Sub-discipline	Research questions (RQ) / Objective (O) / Purpose (P)	Type of intervention	Ethics frameworks used
Embedding Responsible AI in Technical Education Curriculum: A Case Study in an Asynchronous Online Advanced Data Analytics Course	Computing, data science, and related fields	O: 1. Describe how the course was designed to embed the principles of responsible AI in an asynchronous online course. 2. Discuss challenges faced in the course design. 3. Present best practices and lessons learned that can be used to replicate the success of the case.	Asynchronous online course integrating case studies, coding activities, and ethical reflections	Microsoft principles of responsible AI ACM FAccT (Fairness, Accountability and Transparency) Google principles of responsible AI IBM principles of responsible AI Meta principles of responsible AI UNESCO Recommendation on the Ethics of Artificial Intelligence European Commission Guidelines for Trustworthy AI
Ethics, AI, and Society: Shaping Responsible AI Engineers with CDIO Framework	Engineering	P: The paper underscores how this course, rooted in the CDIO framework, aims to equip learners with a holistic and in-depth understanding of the ethical challenges and considerations in the AI and societal interface.	Course "Ethics, AI, and Society" designed under CDIO framework	CDIO Standards (1, 2, 3, 7, 8, 11)
The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements	STEM	O: In this study, we look to see how engaging students using the AI Incident Database can raise their awareness of possible AI harms and enrich their overall thinking about AI ethics.	Classroom module using AI Incident Database (AIID)	AIID classroom activity (incident exploration, reporting, reflection) Motivation, applications, values, role of ML experts
Harnessing generative AI in chemical engineering education: Implementation and evaluation of large language models ChatGPT v3.5	Chemical Engineering Education	O: The research presented herein describes the development, implementation, and evaluation of an undergraduate chemical engineering lab session with a specific focus on using LLMs.	Lab session integrating ChatGPT v3.5 into fluid dynamics experiment	IDEE (Identify desired outcomes, Determine level of automation, Ensure ethics, Evaluate effectiveness) FAST principles (Fairness, Accountability, Sustainability, Transparency)
Case Studies About Moral Dilemmas to Apply Ethical	Computer Engineering	P: The key aspects underlying the course are:	Ethics courses (Fundamentals of Business	Ethics approach: students apply Kantianism,

Theories in Engineering Education		1. The practical lectures and case studies included in the course to help students apply ethical theories in real-world scenarios. 2. The importance of incorporating ethical reflection into engineering education and recognizing the societal impact of technology. 3. The use of philosophy, sociology, and history as tools to help engineers develop a critical and ethical perspective on technology and its impact on society	Ethics and Innovation; Social & Environmental Aspects of Computing)	Utilitarianism, Social Contract Theory, and Virtue Ethics to real dilemmas
ETHICS4EU: Designing new curricula for Computer Science Ethics Education: Case studies for AI ethics	Computer Science	P: In this paper we present and evaluate educational content that was developed as part of the project, specifically lessons that focus on ethics related to computer programming errors.	Case study-based ethics curriculum (delivered in programming module)	ProPublica COMPAS case study (Dressel & Hany, 2021)
Unpacking the role of AI ethics online education for science and engineering students	Electrical Engineering, Computer Engineering, and Information System Engineering	RQ: 1. Whether and to what extent does participation in the online explicit-refective learning module enhance knowledge of AI ethics? 2. Whether and to what extent does participation in the online explicit-refective learning module enhance both perceived and actual awareness of AI ethics? 3. Whether and to what extent does participation in the online explicit-refective learning module improve problem-solving skills in AI ethics?	Online explicit-reflective learning module on AI ethics	OECD (2021) – Recommendation of the Council on Artificial Intelligence IEEE (2019) – Global Initiative on Ethics of Autonomous and Intelligent Systems
Applied Ethics via Encouraging Intuitive Reflection and Deliberate Discourse	Computer Engineering	P: The purpose of this paper is to introduce and detail a learning intervention intended to enhance the ethical behavior of future AI developers and engineers	SIMDE framework learning intervention (Social Intuitionist Model + Discourse Ethics)	Social Intuitionist Model (Haidt, 2001) Discourse Ethics (Habermas, 1990s) SIMDE framework applied to AI ethics dilemma PROPOSED Markkula Center for Applied Ethics framework (decision-making) UNESCO AI ethics video resources

Appendix D. List of publications with research design methodology and key main results

Id	Title	Authors (Name of all authors in APA)	Study Design	Key results
1	Embedding Responsible AI in Technical Education Curriculum: A Case Study in an Asynchronous Online Advanced Data Analytics Course	Abadia, R.	Qualitative	Students applied Responsible AI principles in coding tasks; shared reflections in forums; peer review and journaling
2	Ethics, AI, and Society: Shaping Responsible AI Engineers with CDIO Framework	Ajailia, N., Zerai, M., Debbabi, N., Damergi, A., Dchicha, M.	Mixed-methods	Students reported improved ability to tackle ethical dilemmas; broadened perspectives on AI's societal impact; valued guest speakers and case studies
3	Leveraging AI Tools in University Writing Instruction: Enhancing Student Success While Upholding Academic Integrity	Akiba, D., & Garte, R.	Mixed-methods	Students reported AI feedback was clear, specific, actionable; promoted deeper revisions; increased independence and motivation; reduced temptation to cheat
4	Developing postgraduate students' competencies in generative artificial intelligence for ethical integration into academic practices: a participatory action research	Aladsani, H. K.	Qualitative	Students developed competencies in four themes: technical/AI literacy, ethical/legal, critical/analytical, lifelong learning/interpersonal; demonstrated ability to evaluate GenAI outputs and present findings
5	A Human-Centric Automated Essay Scoring and Feedback System for the Development of Ethical Reasoning	Lee, A. V. Y., Luco, A. C., & Tan, S. C.	Mixed-methods	Students received formative feedback from AES system + peers; instructors provided feedback loop; AES slightly less accurate than human raters but improved over iterations
6	Developing AI Literacy of Management Students using Problem and Project based Learning	Chen, Y. Y., Qin, Z.	Quantitative	Students reported improved AI literacy (69% perceived increase); Business/Science students benefited more than Tech/Engineering peers; gained confidence in prescribing AI solutions and presenting recommendations
7	STEM-based Artificial Intelligence Learning in General Education for Non-Engineering Undergraduate Students	Lin, C.-H., Yu, C.-C., Shih, P.-K. & Wu, L.-Y.	Quantitative	Students reported improved AI literacy and understanding; significant gains in ethics awareness among low-literacy learners; paired t-tests confirmed improvements
8	Exploring the Use of Generative AI in Student-Produced EFL Podcasts: A Qualitative Study	Baskara, FX. R., Dewangga Puri, A., & Laos Mbato, C.	Qualitative	Students perceived AI as creative collaborator; improved linguistic quality and efficiency; noted limitations in creativity and cultural relevance; raised ethical concerns (authorship, data privacy); reported accessibility challenges
9	The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements	Feffer, M., Martelaro, N., & Heidari, H.	Qualitative	Students reported increased awareness of magnitude/severity of AI harms; shift in priorities (urgency around safety, governance, accountability); found AIID valuable for updating knowledge; feedback on limitations (coverage, timeliness)
10	Advancing asynchronous pre-class learning in flipped classrooms: Generative AI companions in business ethics	Hu, Y. H.	Quantitative	Experimental group showed significantly higher pre-class readiness and ethical learning confidence; no increase in perceived task load; students reported improved preparation for in-class debates
11	Harnessing generative AI in chemical engineering education: Implementation and evaluation of large language models ChatGPT v3.5	Keith, M., Keiller, E., Windows-Yule, C., Kings, I., & Robbins, P.	Mixed-methods	Students reported limited prior GAI experience but recognized usefulness; post-survey showed improved understanding of ethical use, critique of outputs, awareness of limitations; thematic analysis identified 8 themes including strong awareness of pitfalls and hallucinations
12	A Study on Ethical Awareness Changes and Education in Artificial Intelligence Society	Kwon, J.	Quantitative	Group B showed significant gains in responsibility/safety and tolerance/limits; moderate gains in employment/discrimination; strongest differences in copyright and AI rights recognition; Group A showed lower awareness across domains
13	Development of Convergence Course of Artificial Intelligence and Psychology Applying Team Teaching Method	Lee, K., & Park, H.-Y.	Mixed-methods	Students engaged in group discussions and presentations on AI & human ethics; demonstrated improved humanities literacy and convergence thinking; reported positive perceptions of interdisciplinary learning
14	Case Studies About Moral Dilemmas to Apply Ethical	Guerrero, M. J. C., & Forment, M. A.	Qualitative	Students applied ethical theories to real dilemmas (e.g., cryptocurrency crowdfunding for Ukraine war); developed arguments and counterarguments; voted on most persuasive

	Theories in Engineering Education			arguments; survey results showed improved ability to apply theories and enhanced critical thinking, though some discomfort with debate format
15	Enhancing fieldwork readiness in occupational therapy students with generative AI	Mansour, T., & Wong, J.	Mixed-methods	Lecture introducing case study and ChatGPT-generated interventions (including one contraindicated option for ethical evaluation); assignment requiring students to use ChatGPT to generate interventions, evaluate them with evidence-based resources, and reflect on ethical/safety considerations
16	ETHICS4EU: DESIGNING NEW CURRICULA FOR COMPUTER SCIENCE ETHICS EDUCATION: CASE STUDIES FOR AI ETHICS	O'Sullivan, D., Gibson, J. P., Curley, A., Becevel, A., Murphy, E., & Gordon, D.	Mixed-methods	Students reported increased appreciation of ethics in programming; valued interactive discussions and peer perspectives; many changed views on technology's role; highlighted importance of including diverse perspectives (minorities, disabilities). Lecturers noted strongest engagement with AI-focused cases (judicial sentencing, autonomous vehicles).
17	Learning Ethics in AI—Teaching Non-Engineering Undergraduates through Situated Learning	Shih, P.-K., Lin, C.-H., Wu, L. Y., & Yu, C.-C.	Quantitative	Students improved AI understanding, teamwork, attitudes, and ethics awareness; discovered importance of sufficient training data and negative samples; applied models to car kits; survey results showed positive correlations among AI understanding, attitudes, and ethics awareness
18	The Impact of Artificial Intelligence-Assisted Learning on Nursing Students' Ethical Decision-making and Clinical Reasoning in Pediatric Care: A Quasi-Experimental Study	Shin, H., De Gagne, J. C., Kim, S. S., & Hong, M.	Mixed-methods	Control group outperformed experimental group in ethical standards and nursing processes ($p < .001$); control group showed stronger critical thinking, evidence integration, and reflection; experimental group reported lower reliability but higher time efficiency and broader perspectives; reflective essays revealed differences in resource use (AI reliance vs. diverse sources)
19	Ethics Crisis Standards of Care Simulation	Switzer, D. F., & Knowles, S. G.	Qualitative	Students deviated from triage AI algorithm; implicit bias entered decision-making; experienced moral distress; used clinical judgment to care for all three patients; recognized ethical dilemmas in CSC; reported learning about fairness, bias, and moral distress in triage
20	Unpacking the role of AI ethics online education for science and engineering students	Usher, M., & Barak, M.	Mixed-methods	Students' knowledge of AI ethics improved significantly (mean score ↑ from 6.37 to 10.26/12, $p<0.001$, $d=1.19$); perceived awareness increased moderately ($M=3.69 \rightarrow 4.00$, $p<0.001$, $d=0.70$); actual awareness and problem-solving skills enhanced (students articulated privacy, bias, flawed datasets, social representation issues, and proposed solutions)
21	Student Perceptions of AI-Generated Avatars in Teaching Business Ethics: We Might not be Impressed	Vallis, C., Wilson, S., Gozman, D., & Buchanan, J.	Qualitative	Students reported ambivalent perceptions: some unaware avatar was AI until told; initial novelty but diminished engagement with repeated exposure; some valued "shock" as teaching strategy; tutors noted clarity for international students but limited impact overall; outcomes highlighted need for personalization, awareness, and integration into pedagogy
22	A mixed methods crossover randomized controlled trial exploring the experiences, perceptions, and usability of artificial intelligence (ChatGPT) in health sciences education	Veras, M., Dyer, J. O., Shannon, H., Bogie, B. J. M., Ronney, M., Sekhon, H., Rutherford, D., Silva, P. G. B., & Kairy, D.	Mixed-methods	ChatGPT rated higher for ease of use and quick learnability (SUS Q7, Q10 significant); students valued productivity, brainstorming, creativity, and research utility; concerns about reliability, accuracy, plagiarism, and unclear boundaries; strong call for clear policies, accountability, and ethical guidelines
23	Applied Ethics via Encouraging Intuitive Reflection and Deliberate Discourse	Wiese, L., & Magana, A. J.	Qualitative	Students engaged in dilemma resolution, peer discourse, perspective-taking, expert-informed reflection; reported moral tension, defended gut reactions, some broadened perspectives
24	Integrating Generative AI in Chemistry Education: Enhancing Career-Ready Writing Skills in Pharmaceutical Science	Yuriev, E., Burton, M. G., Knott, D., Terrill, A. E., Jackson, N. R. C., & Chen, S.	Mixed-methods	Students engage in co-creation with GenAI for report writing.
25	Enhancing anti-plagiarism literacy practices among undergraduates with AI	Zhang, Y., Liu, Y., & Wang, X.	Mixed-methods	Students in experimental group showed significant improvement: perceptions shifted from "important but vague" to "understanding and confident"; plagiarism level decreased (moderate → minor); writing quality improved (mean score 58 → 79/100). Control group showed no comparable gains.