AI

MDPI

Article

Governing AI Outcomes in Civil and Construction Engineering Education: Toward Trustworthy and Ethical Implementation

Armita Dabiri and Amir H. Behzadan *

Department of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, CO 80309, USA; armita.dabiri@colorado.edu
* Correspondence: amir.behzadan@colorado.edu

Abstract

Artificial intelligence (AI) is transforming civil and construction engineering (CCE) education. CCE students must develop both AI technical proficiency and ethical awareness, ensuring that AI tools reflect the varied experiences of project clients. Integrating ethical AI instruction supports the formation of students' professional identity by encouraging them to internalize values of responsibility, integrity, and equity in future practice. Adopting a narrative literature review, this paper examines the ethical concerns surrounding AI in CCE education, implications of existing regulations, and the need for institution-specific risk mitigation policies. The synthesis of the literature indicates that while integrating AI into CCE education may enhance learning experiences and personalized instruction, it also raises key ethical risks, such as algorithmic bias, privacy concerns, lack of transparency, threats to academic integrity, and digital inequity. As such, we also offer guidelines for responsible AI use in educational settings and propose an output governance framework and practical recommendations for educators and institutions, including performing regular ethical and algorithmic audits of AI tools, involving students in technology deployment decisions, and providing faculty development on digital ethics. By detailing actionable strategies, formalizing human-in-the-loop validation workflows, and preserving chains of provenance for educational metrics, these recommendations aim to align AI innovation with the core values of engineering education, i.e., integrity, equity, and public welfare.

Keywords: artificial intelligence; engineering education; civil and construction engineering; accountability; ethics; output governance; audit trail

Academic Editor: Yusaku Fujii

Received: 1 July 2026
Revised: 11 September 2026
Accepted: 11 September 2026
Published: 14 September 2026

1. Introduction

1.1. Integrating AI in Pedagogy: Opportunities and Risks

Artificial intelligence (AI) is transforming higher education [1,2] and creating new student learning experiences that promise greater efficiency, personalization (i.e., catering content to students' needs, learning styles, and abilities), and access to educational resources [3–6]. Recent studies suggest that the use of generative AI (GenAI) may increase academic productivity, as reflected in higher publication rates [7]. Cheng et al. [8] conducted a survey of 68 construction faculty members across 58 universities worldwide, examining the role of AI in civil and construction engineering (CCE) education, and reported that over 90% of respondents demonstrated a high level of AI awareness, and most viewed AI as a beneficial tool, while also recognizing its ethical implications.

These perceptions illustrate that AI integration in education is rapidly shifting from theory to practice, offering both impactful learning opportunities and unique risks [8].

Automated assessment systems can now grade quizzes, verify computer code, and provide feedback on design iterations [9–11]. Intelligent tutoring systems (ITS) extend this further by adapting to individual learners and offering personalized guidance [12]. At the same time, emerging large language models (LLMs), customized for specific design or engineering problems, can support self-directed learning by answering student inquiries, providing on-demand access to learning resources, and improving student peer feedback, freeing faculty time to focus on instruction [13–16]. Moreover, AI-assisted proctoring tools are increasingly used in remote learning environments to detect cheating behaviors through biometric data and behavioral cues [17,18].

In a profession where decisions directly affect public safety and well-being, the stakes could not be higher. The bedrock principles of CCE practice revolve around public trust, sustainability, integrity, and the minimization of risk or harm [19]. AI learning tools in CCE range from automated feedback systems that guide students through structural analysis problems to building information modeling (BIM) simulations that show how construction plans might work in the real world [20–22]. As the use of AI becomes increasingly prevalent in CCE applications, supporting tasks from optimizing concrete mix designs and predicting structural degradation to automating construction progress monitoring and enhancing site safety [23–25], it is essential that future engineers develop both AI technical skills and a deep understanding of their ethical obligations, recognizing that using these tools responsibly constitutes a core professional duty.

The White House's 2025 report, "Winning the Race: America's AI Action Plan", prioritizes multi-sector AI adoption, workforce reskilling, and global competitiveness, while also raising concerns about bias in AI systems, the need for trustworthy and objective information tools, the protection of the workforce, and the prevention of misuse and emerging risks associated with advanced technologies [26]. Specifically, the rapid and often untested integration of AI into the education sector can introduce a complex array of ethical questions. Concerns related to fairness in algorithmic decision-making [27,28], privacy and security of sensitive student data [29], transparency of AI-driven design or analysis [30], and the perennial challenge of academic integrity [31,32] are becoming central to discussions about AI in CCE education. These concerns are not merely theoretical: technologies and tools enhance learning outcomes only when they are designed ethically and responsibly and implemented through careful strategies. As such, educators must balance their instructional benefits with a thorough evaluation of their ethical implications. The distinct nature of CCE, characterized by applied problem-solving, hands-on experimentation, and high-stakes decision-making, both amplifies the value of AI and highlights its limitations. Many context-sensitive design and engineering tasks demand human oversight and ethical discernment. Thus, integrating AI into CCE education requires a deliberate, risk-aware approach that enhances learning while reinforcing the ethical competencies essential for responsible practice in the built environment.

Although numerous studies have extensively discussed model and data governance in artificial intelligence, output governance has received comparatively less attention. Much of the existing literature focuses primarily on input and model architectures, leaving the output side significantly under-researched. Consequently, while issues such as algorithmic bias and privacy are frequently highlighted, a critical gap remains in how to actively handle these challenges directly at the output stage. In this paper, we explore these ethical implications and challenges, making a strong case for appropriate human oversight and robust output governance. This is particularly crucial in civil and construction engineering (CCE) education and built environment applications, where AI-assisted decision-making requires a distinct governance approach to ensure transparency, equity, and safety. To address this, our work proposes a novel output governance framework derived from, and de-

signed to complement, established Verification and Validation (V&V) measures in systems engineering. Unlike existing educational technology or institutional AI guidelines, this framework comprehensively accounts for all parties involved in the decision-making setting. Specifically, we examine the mechanisms necessary to verify, audit, and systematically record AI-generated educational artifacts to prevent silent algorithmic failures and combat human automation bias. Furthermore, the framework explicitly illustrates the necessity of pre-adoption review, longitudinal benchmarking, and capacity building, extending the focus well beyond the governance of the output itself. Ultimately, we argue for a proactive and principled approach to AI adoption in CCE education that implements rigorous multi-tiered verification systems and immutable provenance logs, thus preparing ethically aware engineers capable of critically evaluating and responsibly applying AI while upholding the core values of integrity, equity, and public welfare in an increasingly digital world. While the ethical principles discussed throughout this paper draw on the well-established literature on AI, educational, and engineering ethics, the paper's primary contribution lies in translating these principles into a concrete, CCE-specific output governance framework and actionable institutional guidelines.

1.2. Review Approach

We adopt a narrative review approach, illustrated in Figure 1, given its aim of conceptually synthesizing and interpreting the literature across a heterogeneous body of scholarship rather than comprehensively covering it. Sources were primarily identified through Google Scholar and Web of Science, using an iterative search strategy, beginning with a set of key terms related to AI ethics, higher education, and CCE education and expanded primarily through backward snowballing, i.e., tracing the references cited within identified sources and supplemented to a lesser extent by forward snowballing to identify more recent works citing those sources. For each AI ethics concept under consideration (e.g., privacy, security), the concept was used as a keyword in combination with other generic keywords (e.g., higher education, construction, engineering), allowing the search to identify research at the intersection of specific ethical dimensions and the educational context. In order to capture both foundational and emerging perspectives on AI ethics and governance, no restriction was placed on publication date. Sources were selected based on their relevance to AI ethics and its implications for education broadly and CCE education in particular, their relevance to established guidelines and governance frameworks, and their conceptual contribution to the topic. Sources were excluded if they were not topically relevant or had been formally retracted or corrected. Claims drawn from industry reports and policy documents were treated with particular caution and are explicitly attributed in the manuscript as originating from an industry or institutional source, rather than presented as peer-reviewed findings. The identified literature was then organized thematically, and the paper's structure reflects the categorization of sources into the thematic clusters that emerged through this process. As with narrative reviews generally, this approach does not guarantee thorough coverage and rigorous synthesis of the relevant literature, a limitation that is characteristic of, and expected in, this review methodology [33,34].

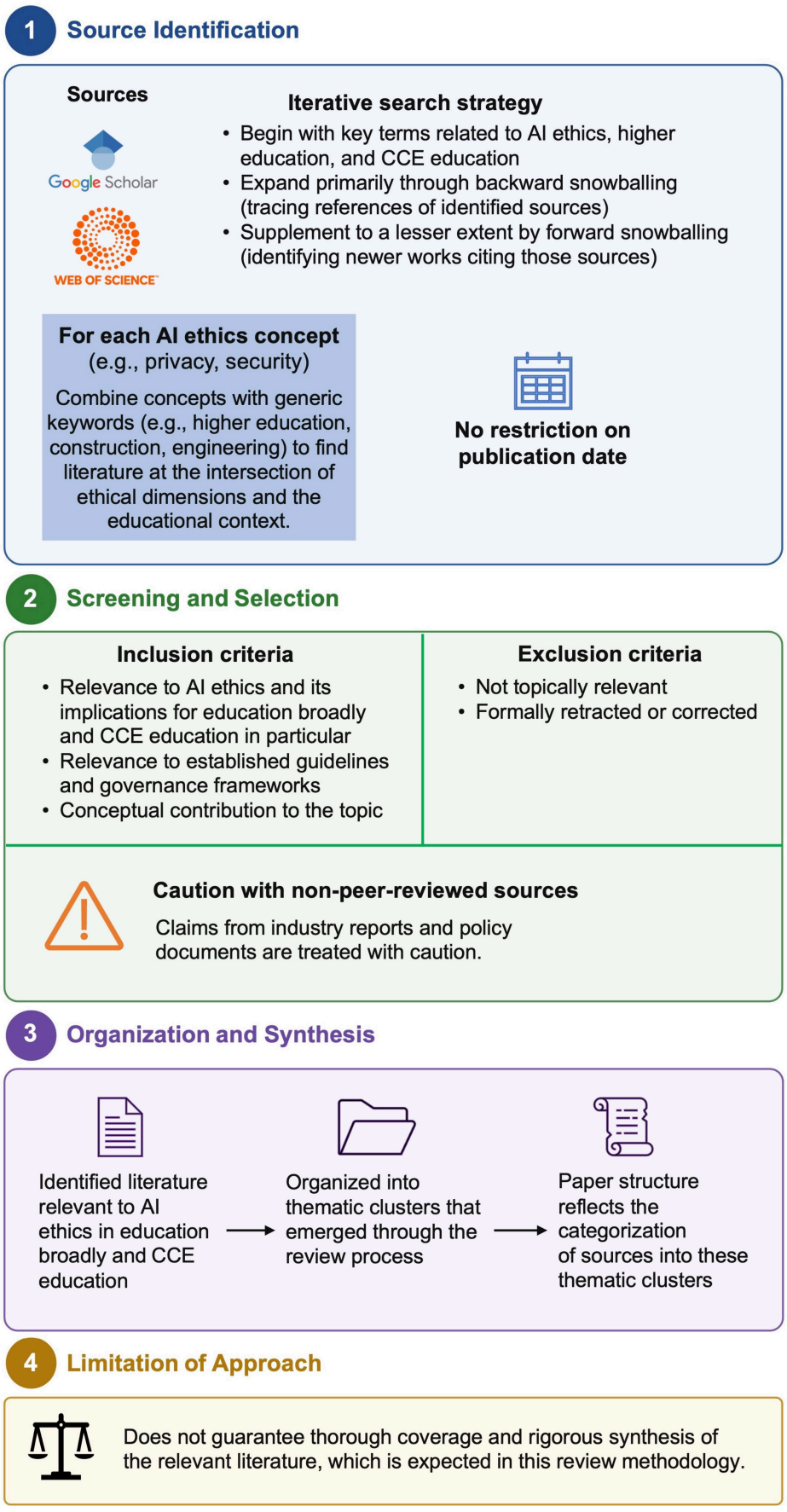

Figure 1. Overview of the narrative literature review methodology.

2. Ethical Implications of AI in Education

The literature on AI in the broader field of education highlights several critical ethical dimensions that directly impact the fairness of the learning environment, the integrity of academic processes, and the well-being of students.

2.1. Fairness and Algorithmic Bias in Assessment

A central ethical concern is that AI-based assessment systems may perpetuate or introduce bias, while current methods for detecting and mitigating biased outcomes remain insufficiently developed [35,36]. Data used to train AI models may reflect or exacerbate existing human biases or systemic societal inequalities [37–39]. The presence of such biases in AI grading tools can unwittingly reinforce linguistic, cultural, or socioeconomic prejudices. A grading algorithm might unfairly penalize non-native English speakers for grammatical nuances or stylistic choices that deviate from its learned or pre-established linguistic patterns [40], even if the underlying technical concepts are sound. According to a 2024 report by the nonprofit organization Common Sense Media [41], course assignments submitted by African American students are disproportionately likely to be incorrectly flagged as AI-generated, raising concerns about bias embedded in AI detection tools. This finding is in line with another study where researchers, through the reflections and experiences of 46 Black high school students, found that nearly all the participants had been falsely flagged at least once, with many of the students reporting multiple false accusations [42]. Similarly, an AI evaluating a student's design submission might unknowingly favor approaches more commonly represented in its training dataset (e.g., choice of specific materials or design methodologies), potentially overlooking innovative or region-specific solutions (e.g., indigenous or culture-driven construction methods). Such embedded biases can unjustly impact grades, affect certain student groups, and undermine the key principles of equitable evaluation.

2.2. Privacy and Surveillance

The use of AI, particularly for remote learning, raises concerns regarding student privacy. AI-assisted proctoring systems, for instance, frequently collect personal data, including biometric information (e.g., facial scans, eye movements), ambient audio recordings of surroundings, and screen activity [43,44], which can lead to heightened anxiety among students and infringe upon their privacy [45,46]. ITS and personalized learning platforms also gather large amounts of student data, including academic performance, interaction patterns, learning behaviors, and even inferred emotional states. The secure storage, appropriate usage, and robust protection of this sensitive data can pose major ethical concerns. Risks such as data breaches, unauthorized access, or commercial exploitation represent major ethical issues, particularly given the power imbalance and vulnerable position of students in educational settings [47,48].

2.3. Transparency and Explainability of AI Outcomes

Many AI models operate as "black boxes", rendering their decision-making opaque and difficult to interpret [49–51]. In the absence of sufficient and context-informed explainability, it is not trivial to identify causal relationships between inputs and outputs, limiting the advancement of scientific theory [52]. In educational contexts, this limitation poses an ethical dilemma. Students need to understand the rationale behind their grades in order to learn from their mistakes and improve. When an AI-assisted grading tool assigns a score or provides feedback without a clear, human-interpretable explanation, it can impede student learning and erode trust in the assessment process [53,54]. To maximize students' learning experience, comprehending the reasoning behind AI-generated solutions is not merely

desirable but crucial. Research shows that overreliance on AI tools to recommend solutions without articulating the underlying logic, scientific principles, or inherent trade-offs (e.g., why a material or design choice was made) can limit students' creativity, critical thinking, problem-solving skills, and professional judgment [55].

2.4. Academic Integrity and Misuse of Generative AI (GenAI)

The emergence of GenAI tools and LLMs has sparked new debates and introduced unique challenges to academic integrity [56,57]. According to a 2024 survey of 1045 teen-parent pairs, over half (53%) of teenagers between 13 and 18 years old reported using GenAI tools for homework assistance [41]. When properly and responsibly used, GenAI can deliver pedagogical gains in tasks such as summarizing and synthesizing multi-sourced information, generating structured report outlines, streamlining computer coding, and assisting with citation formatting, allowing students and educators to engage in higher-level concepts earlier. However, it is critical to assess the long-term impact of GenAI use and conduct careful evaluation to determine whether observed improvements are sustained or primarily attributable to the novelty effect [58]. Research highlights that overuse or misuse of AI tools can undermine the core purpose of education, that is, to support critical thinking, develop problem-solving skills, and promote genuine intellectual understanding of complex engineering principles [59–61]. As more students rely on GenAI for academic writing, the resulting work may lack sufficient technical specificity and become increasingly homogeneous, reflecting patterned linguistic features and stylistic shifts that are consistent across groups [62,63]. Meanwhile, educational institutions are grappling with the complex task of distinguishing between legitimate AI assistance that enhances learning and outright plagiarism, a challenge compounded by traditional plagiarism detection tools often failing to identify AI-generated content [64,65]. This necessitates a thorough re-evaluation of established assessment methods and a thoughtful reconsideration of AI's appropriate role in facilitating, rather than hindering, intellectual development in students.

2.5. Digital Divide and Equity in Access

The increasing integration of AI into educational practices carries the risk of exacerbating the digital divide, i.e., the gap between people or communities that have access to modern information and communication tools (e.g., internet, computers, smartphones) and those that do not [66,67]. This divide can be caused by factors such as socioeconomic status, geography, education, infrastructure, social support, and digital literacy, and further reinforces existing inequalities [68,69]. Students from disadvantaged backgrounds or those enrolled in institutions with fewer resources may consequently lack the infrastructure or financial means to fully utilize AI-enhanced learning opportunities. This could potentially result in a two-tiered educational system, wherein some students significantly benefit from personalized learning and new AI tools while others are left behind [70]. In particular, in engineering education, where specialized software and substantial computational resources are often needed to carry out design and analysis, the presence of the digital divide can have a particularly pronounced and detrimental impact on students' ability to gain relevant professional competencies and access competitive employment opportunities.

2.6. Human Oversight and Accountability

As AI's influence over teaching and learning grows, the line between pedagogical support and control is blurred, raising concerns about how much autonomy should be ceded to AI, and who must be the responsible party for major pedagogical decisions, student well-being, and the upholding of academic integrity [71–73]. If an AI system makes an error in grading, provides misleading information, or inadvertently contributes to a student's distress during a proctored exam, the question arises: who is ultimately

accountable? Is it the AI developer, the educational institution, or the instructor who used the AI? The overlapping responsibilities between AI systems and human educators call for the establishment of clear guidelines for human intervention, continuous monitoring, and the unambiguous assignment of accountability. Overreliance on AI without adequate human review can lead to unintended negative consequences and diminution of the vital human element in teaching and learning, particularly in disciplines that demand rigorous ethical judgment and professional responsibility.

3. Ethical Risks in AI-Facilitated CCE Education and Practice

In this section, we extend our discussion on AI implementation risks and ethical challenges, considering the specific context of CCE education and practice. The ethical issues discussed in this section are grounded in the established literature on AI ethics, educational ethics, and engineering ethics, and serve as the normative foundation for AI-supported CCE education. However, they do not by themselves demonstrate that a specific AI system or output satisfies those requirements. Their practical implementation therefore requires governance mechanisms that translate each principle into defined acceptance criteria, verification procedures, assigned responsibilities, documentary evidence, human-review requirements, and corrective or appeal processes, which will be elaborated on in Section 5. In the construction domain, AI is commonly applied to tasks such as predicting schedule delays and automated scheduling [74–76], estimating cost overruns [77,78], supporting hazard detection [79,80], and offering guidance and preventive actions that function as safety recommendations [81,82]. These applications rely on large datasets collected through wearable sensors and visual data captured by ground-mounted cameras or unmanned aerial systems (UASs) [83–85], thus raising important ethical and legal questions related to privacy, surveillance, data ownership, and embedded algorithmic biases. In educational settings, which mirror many of these professional contexts, such datasets are frequently used for hands-on learning and instructional purposes. However, emphasis is typically placed on tangible AI outputs, such as cost estimates or hazard recognition results, rather than on the broader ethical, legal, and professional implications of how these tools are developed, tested, and used. As a result, critical issues related to privacy and labor data, bias and risk allocation, explainability and liability, safety normalization, and professional accountability are often underexamined in training environments.

In the CCE domain, AI-generated outputs directly affect public safety. Hence, discussing the output governance of AI systems is equally important for preparing students as future engineers. Classroom instruction should therefore address output control mechanisms, including output verification, auditing, historical evaluation, provenance, and traceability. Addressing these concerns early in CCE education is essential to prepare students to become responsible engineers who use AI as a decision support tool rather than a mere replacement for human judgment.

3.1. Privacy, Surveillance, and Labor Data

A central issue in privacy discussions is what personal data is collected and who retains control over that data. AI systems frequently acquire information through practices that lack an explicit consent process. Examples include large-scale data mining, web scraping, facial recognition, and online tracking, which could limit individuals' ability to oversee or regulate the use of their own data [86–88]. In CCE practice, privacy and surveillance concerns arise when, for example, field workers are captured in visual data without informed consent, or when sensitive or confidential project information is inadvertently collected [89]. However, these concerns are rarely translated into educational contexts where resulting datasets from drone footage, wearable sensors, and systems that track

workers' movements on construction sites are used by students. While the use of these data in assignments and course projects can support specific pedagogical objectives and provide students with an opportunity to integrate real-world contexts into their learning experience, questions surrounding how data were collected in the first place are rarely addressed. Classroom discussions frequently move directly to technical analysis without regard for workers' consent, data ownership, or the ethical conditions under which these datasets were generated. In a construction methods class, for instance, when students are asked to solve a problem on productivity analysis, the grading rubric is primarily based on finding creative ways to improve operational efficiency, treating workers' privacy as an afterthought. In this and similar cases, students are seldom asked to reflect on how and under what conditions collected field data should be anonymized, reused, shared, or monetized beyond their intended purpose. In the field, this learning experience may be translated into inadvertently normalizing intrusive equipment and worker monitoring by framing surveillance as a non-negotiable requirement for achieving better safety or efficiency, thereby limiting comprehensive ethical reflection on its broader consequences.

3.2. Algorithmic Bias and Contractual Risk Distribution

Although the final product of an engineering project (i.e., the built facility) may appear standardized and similar to previous examples (e.g., airport facilities or electric power plants in different locations), each project is fundamentally unique, involving its own distinct set of challenges, constraints, expectations, and requirements. Because no single comprehensive dataset can capture the complex processes and nuanced conditions involved in the design, construction, and operation of all built environment projects, AI systems trained on existing data are prone to inheriting and reproducing the biases embedded in their underlying data [89–91]. When applied to problems at the interface of the built environment and the human quality of life, these models may yield strong performance for some groups while simultaneously exacerbating social inequities for others [92,93]. This issue is particularly relevant to the CCE domain, where project resources (and risks) must be responsibly distributed among all stakeholders, and the inclusion of diverse perspectives is essential for designing resilient and equitable infrastructure that serves all segments of society. In the classroom, when historical datasets are frequently and uncritically used, they allow past inequities to creep into student-generated solutions, especially when equity-aware approaches that explicitly consider fair risk distribution are rarely solicited or rewarded. By grading students based on criteria such as minimizing project duration or cost rather than evaluating who ultimately absorbs uncertainty, AI-assisted scheduling or cost estimation assignments tend to prioritize optimization metrics without addressing how risks are allocated among different parties. This risk may be compounded when AI cost estimation tools are trained on historical bid data reflecting a particular level of design maturity. If that maturity level is not made explicit to students, they may be unable to judge whether an AI-generated estimate's implied accuracy (e.g., per the AACE International Class 1-5 classification system) matches the actual project phase.

3.3. Explainability, Liability, and Defensibility of AI-Assisted Decisions

When AI-assisted decision-making tools do not clearly explain how they arrive at specific outputs, their unchallenged adoption may result in significant liability concerns, particularly when the outcome results in loss or harm. Claims based on such decisions may be difficult to defend, as the underlying logic guiding the AI recommendations is not transparent or understandable [94,95]. Although an AI model may demonstrate strong technical performance, its lack of transparency could limit users' ability to interpret how inputs relate to specific outputs [96]. Research also argues that reduced explainability,

particularly in GenAI and deep learning systems, undermines human understanding and trust [97,98] and creates barriers to their broader adoption in architectural and structural design problems [99]. A similar issue arises in educational settings, where students may submit AI-generated schedules, cost estimates, or safety plans without articulating the assumptions or evaluating the credibility, limitations, and methodological rigor behind them. This is particularly consequential for scheduling, where AI-driven optimization can shift or compress the critical path in ways that are not traceable to the student or instructor. Because forensic delay analysis depends on being able to reconstruct why a given activity was critical at a given time, a black-box schedule undermines the reconstructability that both classroom evaluation and real-world dispute resolution require. Grading rubrics often emphasize the completeness and timeliness of the deliverable rather than its interpretability. Similarly, students are rarely asked to complete reflective learning or role-playing exercises to demonstrate in-depth critical thinking and problem-solving skills, and to justify or defend an AI-generated workflow or decision in the event of a dispute or counterargument [100].

3.4. Safety, AI Authority, and Risk Normalization

In safety-critical contexts, key decisions such as hazard recognition and safety planning must be made with care. AI has been increasingly used to support construction safety by detecting potential hazards on jobsites, including monitoring personal protective equipment (PPE) compliance [24,101], enabling real-time hazard identification [79,80], and preventing accidents such as collisions between workers and construction machinery [102–104]. On construction sites, AI-powered tools may therefore function as decision-support assistants for safety management. However, in educational settings, AI recommendations may not fully align with established safety standards (e.g., OSHA requirements) or site-specific conditions. For example, a recent study found that when GPT was used to generate welding safety guidelines, it omitted certain relevant safety requirements and did not cite the specific OSHA provisions on safety training and education that a human safety manager would typically be expected to include [105].

Furthermore, GenAI tools may lack situational awareness and not be able to fully grasp the organizational, emotional, and ethical complexities that often surround safety incidents [106]. Even more concerning are hallucinations, in which the model generates false information, such as citing OSHA standards that do not actually exist [107]. When an AI safety tool fails to flag a hazard, the failure may originate from small, supervised training data [108], occlusion and overlapping of objects on the construction site [109], or operator over-reliance on the tool's output rather than independent hazard recognition. In educational settings, these failure modes might be rarely distinguished, leaving students unable to determine whether a given AI-assisted safety judgment failed due to data, model, or human factors. Moreover, the efficacy and practicality of AI-driven safety recommendations are rarely subjected to critical analysis in coursework. For example, a construction scheduling assignment may prioritize schedule compression or productivity gains without explicitly penalizing decisions that compromise safety. This risk extends beyond safety standards to design codes such as ASCE 7, ACI 318, and the IBC, which are revised on multi-year cycles. An AI model trained on historical data may inadvertently reference an outdated code edition without indicating that a newer version applies. This possibility highlights the importance of ensuring that CCE students can independently verify the applicability and currency of referenced codes and suggests that structured code-version checks could be incorporated into output governance processes.

3.5. Accountability and the Erosion of Engineer-of-Record (EoR) Responsibility

Another critical issue associated with AI-enabled decision-making in CCE concerns the engineer-of-record (EoR) responsibility. In professional practice, the EoR assumes responsible charge for the technical aspects of the project, including design, construction, and operation [110,111]. The National Council of Examiners for Engineering and Surveying model law defines responsible charge as the "direct control and personal supervision of engineering or surveying work", a definition that is also adopted by the National Society of Professional Engineers [112,113]. However, when decisions are informed or generated by AI systems, responsibility becomes ambiguous, particularly in cases where AI recommendations contribute to errors or lead to adverse outcomes. In safety-critical construction contexts, the lack of transparency resulting from the black-box nature of AI tools further introduces ethical and legal ambiguity about liability when decisions result in failure [114]. In a recent workshop sponsored by the U.S. National Science Foundation (NSF), it was reported that GenAI may not perform reliably under conditions of uncertainty or long-term horizon reasoning, making human oversight essential [52]. Professional bodies have begun to take positions on accountability for AI-assisted engineering work. For instance, the National Society of Professional Engineers [115] has taken the position that those responsible for designing, building, deploying, or supervising AI systems that affect public safety should face the same accountability expectations as licensed engineers. At the state level, the Florida Board of Professional Engineers, the Texas Board of Professional Engineers and Land Surveyors, and the North Carolina Board of Examiners for Engineers and Surveyors have announced that only a licensed Professional Engineer can take responsibility for the work, and the licensee is responsible for reviewing the output of the AI software to ensure it is accurate and consistent with engineering principles [116–118]. In educational contexts, students are increasingly using AI tools to draft construction plans and specifications, write code for structural analysis, or formulate capstone design projects. However, the concept of EoR responsibility is not discussed in tandem with planning and design exercises, as mistakes in AI-assisted outputs typically carry minimal academic consequences and receive limited ethical scrutiny. As a result, students are rarely required to articulate when reliance on AI would be inappropriate, unethical, or inconsistent with professional standards.

4. Oversight Framework and Regulatory Aspects of AI in CCE Education

Effectively addressing the ethical implications of AI in CCE education requires the establishment of a robust oversight framework, complemented by an understanding of existing regulations and the development of institutional policies. Such a framework serves to provide an indispensable guiding compass for the responsible use and integration of AI in the educational context, while also aligning with the principles of professional engineering practice.

4.1. Ethical Principles for Responsible AI Use and Integration

Drawing upon principles from general AI ethics and the specific tenets of educational ethics, a foundational framework for AI use and integration in CCE education should be grounded in the following core principles:

- Informed consent: Students and educators must be fully and transparently informed about how and why AI is being utilized, what specific data is collected, and how that data is stored, processed, and shared. Consent should be freely given, highly specific, and readily revocable. This demands genuinely transparent communication and a full understanding of the AI's functionalities and implications for all involved parties. Discussing informed consent in CCE classrooms is also essential for shaping students' future professional identities as engineers who will inevitably confront these

issues in the design and deployment of smart systems in built environments. For example, occupants of smart buildings often have limited awareness of the data being collected about them, and even when they are aware, they may compromise their privacy in exchange for the convenience these systems provide [119]. Under such conditions, engineers have a responsibility to design and deploy smart technologies in ways that preserve user convenience while also promoting data privacy, transparency, and meaningful informed consent for building occupants.

- Equity: AI systems must promote equitable learning for all students, irrespective of their background or learning differences. This principle entails actively mitigating algorithmic bias in AI outcomes, as well as ensuring equitable access to AI resources. It also involves preventing AI from inadvertently creating or exacerbating existing disparities that could hinder an inclusive and supportive learning environment. In a parallel context, i.e., built environments, where the use of AI systems is rapidly expanding, it is essential to ensure that these technologies do not exacerbate existing inequalities, but rather support equitable and responsible technological advancement in the construction industry [120].
- Accountability: Clear and unambiguous lines of responsibility must be established to ensure that educators and institutions remain accountable for pedagogical decisions, student welfare, and the fairness of assessment processes. To this end, robust mechanisms for appeal and prompt redress in instances of AI error or demonstrable bias in grading student work are essential. These provisions reflect the accountability structures inherent in professional engineering practice. As our presented framework assumes human-led education supported by AI rather than autonomous AI instruction, operationalizing this principle requires specifying the respective authority and responsibility of each party involved in AI-mediated educational decisions, summarized in Table 1.

Table 1. Authority and responsibility of instructors, institutions, AI providers, and students under human-led, AI-supported decision-making in educational settings.

Actor	Authority and Responsibility
Instructor	Reviews AI outputs before they affect a student (required for grades, feedback, and pathway/placement decisions), may reject or modify any AI output, holds final responsibility for pedagogical and assessment decisions.
Institution	Vets and procures AI tools prior to deployment, sets policy on when human review is mandatory, conducts periodic audits, holds responsibility for systemic fairness and compliance.
AI provider	Responsible for technical reliability and transparency
Student	Entitled to know when AI was used in a decision affecting them, may appeal through instructor/institutional channels, not responsible for AI system errors.

- Transparency: The decision-making processes of AI tools should be as transparent and explainable as is technically feasible. Nevertheless, this expectation contrasts with current practice. For instance, while LLMs have been widely adopted in construction for tasks such as scheduling and safety monitoring, explainable AI has not been sufficiently integrated into their design [121]. Educators and students should be able to comprehend how AI systems arrive at their assessments, generate feedback, or formulate recommendations. This leads to user trust, enables effective learning, and supports the timely identification and subsequent correction of errors or biases. When black-box models are unavoidable, their outputs must undergo rigorous human validation and critical review.

Together, these principles should guide the selection, implementation, and evaluation of AI in CCE education to support broader educational goals and student well-being, while allowing for continuous reflection and adaptation as AI evolves.

4.2. Implications of Existing Regulations

The integration of AI in educational settings does not occur in a legal vacuum. Existing data privacy regulations play a pivotal role in shaping best practices and imposing legal obligations. In the United States, the Family Educational Rights and Privacy Act (FERPA) governs the privacy of student educational records [122]. AI tools that collect, store, or process student data, e.g., performance metrics or behavioral data, must strictly comply with FERPA's provisions regarding parental and eligible student rights to inspect and review their records, and to control the disclosure of personally identifiable information (PII) derived from those records. To this end, educational institutions must ensure that all third-party AI vendors adhere unequivocally to these regulations. Similarly, educational institutions operating in the European Union are mandated to follow the General Data Protection Regulation (GDPR), which imposes requirements on the collection, processing, and storage of personal data [123]. These regulations include explicit consent requirements, the right to access and rectify personal data, the right to erasure (a.k.a. the "right to be forgotten"), and stringent data security protocols. AI systems that profile students or make automated decisions based on their personal data, such as recommending specific career paths based on student performance records, must be fully GDPR compliant and, if the processing is likely to result in high risks to the rights and freedoms of an individual, undergo frequent data protection impact assessments (DPIAs), as described in Article 35 of the GDPR [124]. It should be noted that this discussion of legal and regulatory obligations is grounded primarily in U.S. (FERPA) and EU (GDPR) frameworks, reflecting the regions most represented in the reviewed literature.

While these regulations provide a fundamental baseline for data privacy, they may not fully encompass the unique and evolving ethical challenges exclusively posed by AI use and integration in educational settings, particularly concerning issues of algorithmic bias or the nuanced implications of pervasive student surveillance.

4.3. Need for Institution-Specific Policies

Given the rapid advancement of AI technologies and the current limitations of broad regulatory frameworks, it is imperative that institution-specific policies for AI use and integration in education be developed [125]. These policies must go beyond general compliance and instead reflect the unique technical, ethical, and pedagogical demands to ensure that AI enhances, rather than undermines, both the educational mission and program-level priorities. A tailored approach should address several core areas, starting with data governance, i.e., ensuring that protocols for data collection, anonymization, secure storage, retention, and deletion are rigorously defined [126,127]. This is especially critical when managing sensitive data such as student performance metrics, where data misuse or mismanagement could have significant consequences.

Equally important is the establishment of transparent and ethically grounded procurement and vetting procedures for AI tools and software used in CCE instruction. Evaluations should assess not only technical performance but also potential risks, including algorithmic bias and lack of transparency. Tools used for design optimization, project scheduling, or automated feedback on student work must undergo bias and ethical audits. To support this, institutions should consider forming dedicated ethical review committees tasked with assessing new AI tools and software to ensure they are not only pedagogically effective but also socially responsible and aligned with institutional values. Faculty and student guide-

lines are also essential for responsible AI use. Clear policies should articulate acceptable uses of AI in coursework (and best practices for disclosing such use), establish expectations for academic integrity, and promote AI literacy [128,129]. For example, guidelines might require students to cite AI-generated code or disclose the use of AI tools in a capstone design project. To this end, research shows that courses that involve students in developing clear AI use policies and expectations tend to see higher levels of engagement and reduced concerns about academic integrity [130]. Additionally, institutions must proactively address issues of accessibility and equity, ensuring all students have access to necessary AI tools regardless of socioeconomic status. Finally, protocols must be in place to clarify the role of human oversight in AI-mediated learning, with instructions for educators on how to intervene when AI outputs are flawed or misleading.

5. Output Governance and the Verification of Educational AI Artifacts

The governance mechanisms detailed below are developed primarily around automated assessment and AI-generated educational outputs, given the acute stakes these applications carry for grading fairness, liability, and academic integrity. However, the underlying principles are intended to generalize across the broader range of AI applications discussed throughout this paper, including intelligent tutoring systems, personalized instruction, content generation, and AI-assisted proctoring, while requiring use-case-specific adaptation in their implementation. Elements such as pre-adoption ethical review, human-in-the-loop validation, provenance documentation, and longitudinal auditing constitute a common core applicable regardless of the specific AI application in use. By contrast, the operational form these mechanisms take must be tailored to the application at hand. The remainder of this section illustrates this framework primarily through the lens of automated assessment, as the application with the most immediate and consequential impact on students, while the principles articulated are presented as a template adaptable to the other AI applications noted above.

Data governance involves establishing authority and oversight over how data are managed, with the goal of maximizing their value while reducing associated costs and risks [126]. Within AI systems, governance extends beyond the data themselves to the models that process them. Model governance refers to the mechanisms and controls used to oversee AI model development and operation, including the algorithms and analytical procedures employed, and to enforce security, privacy, fairness, and other constraints on model inputs and prediction outputs [131]. While data and model governance primarily address the resources and processes through which AI outputs are produced, output governance shifts attention to the results generated by the system and their subsequent use. Output governance refers to a system of rules, procedures, and technical mechanisms intended to ensure that AI-generated outputs align with human values, objectives, and task-specific constraints. Developing a structured output governance framework ensures the validity, traceability, and defensibility of AI-generated artifacts in CCE education. In pedagogical environments where algorithmic systems generate summative assessment feedback or score student work, AI output cannot be assumed to be deterministically correct. In these scenarios, AI applications instead produce stochastic outputs that are prone to silent failures, i.e., when AI produces plausible but incorrect outputs due to drift, degradation, or flawed assumptions, without triggering traditional error alerts [132,133]. Consequently, higher education institutions must deploy systematic mechanisms for output verification, algorithmic auditing, and multi-layered record-keeping to integrate these systems without compromising educational standards.

Output verification protocol design must account for the acute threat of automation bias among human evaluators. Recent behavioral research demonstrates that human do-

main experts frequently miss critical algorithmic errors; for example, educators auditing automated grading outputs are statistically much less likely to correct unduly harsh marks when an evaluation is explicitly labeled as AI-generated than when it is attributed to a human peer [134]. This phenomenon underlines the necessity of deterministic policy bounding boxes to decouple evaluation logic from raw model generation. Within the CCE educational environment, this translates into executing a multi-tiered output governance framework, outlined in the following recommendations and strategies. Building on this framework, we recommend actionable strategies for educators and institutions to responsibly integrate AI into CCE education. These concrete recommendations are based on the three pillars of Prosper, Prepare, and Protect, and offer actionable items for those concerned with educational systems, including families, instructors, and technology companies [135].

Figure 2 presents our proposed output governance framework for AI artifacts in CCE education, illustrating how pre-adoption review, capacity building, verification, explainability, documentation, validation, and longitudinal evaluation can collectively support responsible AI integration in pedagogical settings. The proposed workflow also operationalizes the Prosper, Prepare, and Protect pillars articulated by Burns et al. [135]. Capacity Building corresponds to Prepare, building the AI literacy and professional development needed for students, educators, and institutions to engage with these systems effectively. Pre-Adoption Review, together with the Verify-Explain-Validate-Document sequence and the escalation and student-appeal processes, corresponds to Protect, embedding safeguards and responsible governance into both the selection and ongoing use of AI systems. Longitudinal Evaluation corresponds to Prosper, supporting the kind of sustained inquiry into when and how AI genuinely benefits student learning that the Prosper pillar calls for. This framework was derived from the principles of accountability, transparency, informed consent, and equity established in Section 4, operationalized in response to the automation-bias risks and regulatory obligations discussed earlier. The framework begins with a pre-adoption ethical review, which closely resembles the IRB process. At this stage, institutions evaluate AI tools through ethical review processes to assess their potential risks, impacts, and pedagogical value, with students also having opportunities to participate in the evaluation process. If an institution determines that the overall pedagogical benefits of a specific AI tool outweigh its potential risks and that appropriate mitigation strategies are in place to safely address identified risks, the tool's outputs must first undergo verification, explanation, and validation. Verification and validation are treated as distinct, sequential stages, consistent with the standard engineering distinction between the two: Verification is a set of activities that compare a system to a predefined set of characteristics and establish whether an output conforms to explicit, deterministic constraints, such as syntax, boundary conditions, or codified engineering rules, and can therefore be checked algorithmically without human judgment [136,137]. Explanation sits between verification and validation because meaningful validation depends on tracing an output back to the inputs, rules, or features that produced it. A human reviewer cannot judge fitness for purpose without first understanding how the output was derived. Validation establishes whether a verified output is actually fit for its educational purpose, a determination that necessarily requires human judgment because it must account for context, fairness, and defensibility that rule-based checks cannot capture. Verification ensures that the system is implemented correctly, while validation delves into a more fundamental question: Is the right system built [138]? Validation demonstrates whether a system satisfies its intended use and user needs, and requires external evidence and comparison to real-world behavior [137,139]. It also involves human-in-the-loop assessment to minimize automation bias, and regular performance and ethical audits to inspect data privacy, algorithmic bias, and transparency. During validation, if the output is rejected by the human reviewer, it is withheld from use

and undergoes correction, regeneration, and reassessment through updates to the prompt, input data, or methods. If the final human-reviewed decision substantially affects a student, opportunities for student correction and appeal should be provided when necessary. This stage, therefore, should involve informing the student of the decision and providing explanation and rationale.

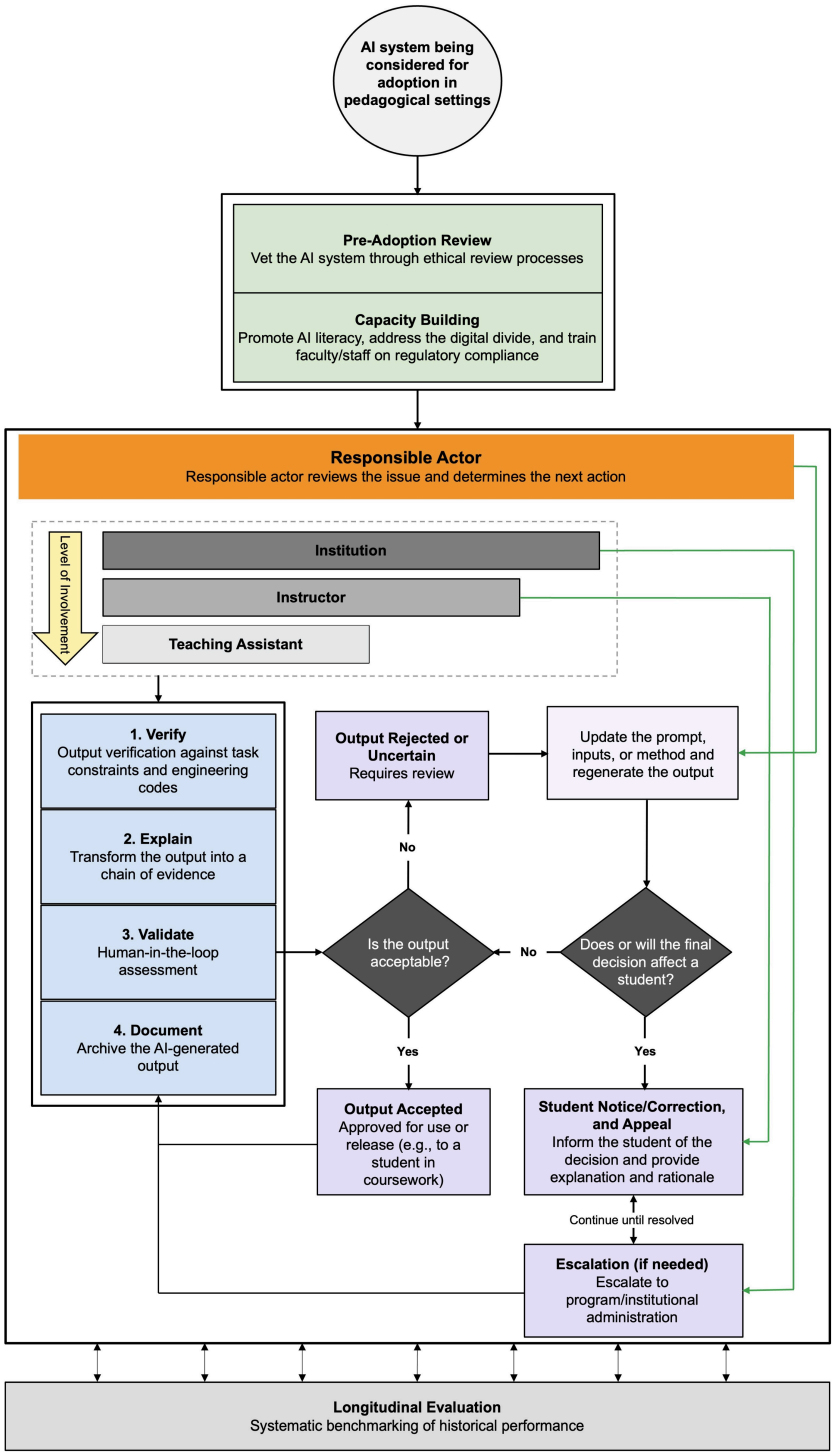

Figure 2. An institutional workflow for adopting AI systems, beginning with an Institutional Review Board (IRB)-style review to vet the tool before implementation, and continuing through verification, explanation, validation, and documentation. Capacity building supports the process, while human oversight, student appeal, escalation, and longitudinal evaluation promote accountability and continuous improvement. The roles and responsibilities of the actors depicted here (instructor, institution, student) are defined in detail in Table 1.

In cases involving systematic and repeated high-risk outputs with serious consequences for the students or unresolved issues at the responsible actor level, concerns should be escalated to program and institutional administration. If the final output is deemed acceptable, the output and relevant review processes are documented. Documentation involves archiving the input context, prompt configuration, model metadata, and human-review outcomes associated with each decision, thereby creating the audit trail that supports both individual student appeals and the longitudinal evaluation described at the end of the framework. An integral component of this framework is longitudinal evaluation, in which historical performance data are recorded and analyzed to determine whether the governance mechanisms improve output accuracy, enhance process transparency, and strengthen accountability among parties involved in AI-assisted decision-making. Capacity building is an integral and continuous component of this cycle. Faculty and staff must receive training in AI literacy, digital ethics, regulatory compliance, and digital equity considerations to effectively mitigate bias, maximize pedagogical benefits, and contribute to efforts toward equitable access to AI-enabled educational resources. Table 2 summarizes the purpose of each component.

Table 2. The purpose of each component of the output governance framework.

Component	Purpose
Pre-adoption Ethical Review	Evaluates AI tools (similar to an IRB process) for risks, impacts, and pedagogical value before adoption, incorporating student input. It compares benefits and risks and ensures that mitigation strategies are in place.
Verification	Algorithmically checks whether an output conforms to explicit, deterministic constraints (syntax, boundary conditions, codified rules)
Explanation	Bridges verification and validation by tracing an output back to the inputs, rules, or features that produced it, so a human reviewer can understand how the output was derived.
Validation	Human-in-the-loop judgment on whether a verified output is fit for its educational purpose, accounting for context, fairness, and defensibility. Includes audits for privacy, bias, and transparency.
Correction/Regeneration	If a human reviewer rejects an output during validation, it is withheld and revised through updated prompts, input data, or methods, then reassessed.
Student Notification, Correction, and Appeal	When a human-reviewed decision substantially affects a student, the student is informed with rationale and given the chance to correct or appeal.
Escalation	Routes systematic, high-risk, or unresolved issues to program/institutional administration when consequences for students are serious.
Documentation	Archives input context, prompt configuration, model metadata, and human-review outcomes for each decision, creating an audit trail.
Longitudinal Evaluation	Analyzes historical performance data over time to assess whether governance mechanisms improve accuracy, transparency, and accountability.
Capacity Building	Ongoing training for faculty/staff in AI literacy, digital ethics, regulatory compliance, and equity to support the entire cycle.

Although the above ethical principles and governance stages apply broadly across AI-supported CCE education, their implementation must be adapted to the nature and consequences of each application. In particular, the level of governance should be proportionate to the potential consequences of the AI use case. Low-risk applications with limited impact, e.g., generating practice questions, summarizing course material, or supporting early-stage idea development, may require relatively light safeguards. In contrast, high-risk uses that can affect students, professional judgment, or public safety, e.g., admission or assessment decisions, behavioral monitoring, or AI-assisted evaluation of engineering designs, call for stronger review, documentation, and human oversight. Table 3 operational-

izes this framework across several representative educational use cases by identifying the relevant AI output, evaluation criteria, responsible actors, verification and documentation requirements, and mechanisms for human review, escalation, and appeal.

Table 3. Application of the proposed AI governance framework across representative AI-supported CCE educational use cases.

Use Case	AI Output	Rule/Criterion	Responsible Actor	Verification and Documentation	Review, Escalation, and Appeal
Automated grading and feedback	Score, rubric assessment, written feedback	Alignment with rubric, factual and technical accuracy, and non-discrimination based on linguistic, cultural, or stylistic variation	Instructor, teaching assistant	Rule-based rubric checks; sampled double-grading; record model version, prompt, rubric, and output; final grade; responsible actor acceptance	Instructor approves consequential grades, anomalies should be reported to the program administration, student may request human regrading
AI-assisted structural design assessment	Evaluation of load combinations, member capacities, design assumptions, code compliance and flagging where a provision is violated	Assignment requirements, applicable code edition, physical constraints, distinction between compliance and engineering sufficiency	Instructor, teaching assistant	Deterministic checks where technically feasible, code-edition verification, documented input assumptions and assignment requirements, responsible actor acceptance	Outputs with conflicting constraints, incomplete load paths, redundancies, or ambiguous code interpretation require responsible actor's review, student may submit justification or correction and request human assessment
Scheduling and cost-estimation assignments	Critical path, activity durations, dependencies, sequence of activities	Traceable assumptions, valid dependencies and sequences, stated estimate accuracy consistent with project phase	Instructor, teaching assistant	Preserve schedule logic and input assumptions; record model version, prompt, rubric, and output; and compare with baseline or student-developed analysis, responsible actor acceptance	Black-box changes to critical path or unexplained cost assumptions, dependencies, and sequences trigger review, students may challenge the result with documented counter analysis
Personalized tutoring or ITS	Recommended content, feedback, learning pathway, learner profile, difficulty level	Pedagogical relevance, privacy limitations, non-discrimination, preservation of student autonomy	Instructor, teaching assistant, institution for data-governance controls	Review of recommendations, profile correction mechanism, record categories of data used, major interventions, and performance patterns, responsible actor acceptance	Instructor can override recommendations, persistent errors escalated to program or vendor review, students may inspect and correct relevant profile information
AI-assisted proctoring	Behavioral flag or suspected violation	Defined evidence threshold, privacy and proportionality, no automatic finding of misconduct	Instructor and designated academic integrity officer	Human review of evidence, record reason for flag, data used, reviewer decision, and retention period, responsible actor acceptance	No penalty based solely on AI flag, case escalated through existing academic integrity procedures, student receives notice and opportunity to contest

5.1. Practical Recommendations for Educators

Educators must conduct regular ethical audits of AI tools and software prior to adoption. These evaluations should extend beyond technical performance to include critical analyses of aspects such as transparency, algorithmic bias, data privacy, potential

impact on student autonomy, and the reliability of AI-generated outputs. For instance, when deploying AI-driven feedback systems to grade student work, instructors must assess whether the algorithms exhibit bias (e.g., favoring a particular terminology or writing style) or obscure their decision-making. Rigorous ethical vetting not only prevents unintended harm but also helps align AI applications with broader learning objectives and professional standards of the CCE practice. Audits can be operationalized through several complementary mechanisms, as described below. To guide both immediate pedagogical practice and future technical development, these recommendations are organized across two distinct tiers: (1) operational governance capabilities (both existing and near-term) that educators can adopt immediately using policy, workflow design, and standard logging practices; and (2) forward-looking research directions that outline design patterns requiring dedicated system-level V&V and integration.

- Existing operational governance capability: Structured human-in-the-loop workflow to counteract automation bias. For example, human grading audits can be conducted using blind-labeling techniques where instructors review a randomized subset of both human-graded and AI-graded submissions without knowing the evaluator's identity. This approach is consistent with empirical work showing that blind, randomized evaluation designs help surface errors experts otherwise miss when they know an output was AI-generated [134]. In addition, a formal, transparent appeal process must be provided to students so that trust in AI remains an active sociotechnical construct strengthened with continuous human validation [140].
- Existing operational governance capability: For high-stakes applications (e.g., grading), AI-generated decisions must be archived along with their input context, prompt configuration, corresponding model metadata (e.g., model version, seed value), and the human auditor's intervention logs. The information that instructors and institutions can obtain from proprietary AI systems, e.g., model version, data governance information, privacy and security compliance, and adoption statistics, should be archived as well. Preserving this digital chain of custody guarantees that academic records remain transparent, verifiable, and fully compliant with evolving institutional and legal mandates over time.
- Near-term operational governance capability: Explainable AI (XAI) diagnostics that make each AI-generated result traceable through accessible evidence, such as the source inputs, user instructions, applied rules, verification steps, and reported uncertainty, rather than relying on access to model parameters or training datasets, which are typically unavailable in proprietary closed-weight systems. While these traceability measures can support manual explanation and interpretation of AI outputs, more comprehensive forms of automated, model-internal explanation remain an emerging research objective. Until such capabilities become technically mature and widely available, deterministic audit gates offer the most practical and reliable means of oversight.
- Forward-looking research direction: Deterministic, rule-based audit gates that validate the output of an AI system against task constraints before it reaches the student or instructor represent a promising design pattern rather than a demonstrated, off-the-shelf capability. For example, if an AI-based grading agent assesses a student's structural design optimization, the output must first pass through a syntax and boundary validation system to ensure the model's grading metrics do not violate fundamental physics or code-defined engineering constraints (e.g., ASCE 7 load combinations). Passing a discrete load-combination check, however, does not guarantee that a design satisfies the underlying intent of the code, such as redundancy, ductility, or continuous load path. Audit gates should therefore be designed to flag AI-generated structural outputs

for instructor review whenever they optimize narrowly to a stated constraint, since code compliance and code sufficiency are not equivalent, and this distinction is often where engineering judgment, rather than rule-following, is required. Developing and validating such deterministic checking systems for CCE-specific constraints remains an open research direction rather than a currently deployable practice.

It is also important to promote AI literacy and ethical reasoning among students to help them learn not only how to operate AI tools but also develop a sufficient understanding and appreciation of AI's capabilities, limitations, and societal implications. Instructors should incorporate discussions that explore real-world ethical challenges (e.g., algorithmic bias in predictive maintenance systems, data privacy concerns in infrastructure health monitoring) and accountability for AI-driven decision-making in CCE practice. To reframe AI as an "object of critique" rather than a shortcut, in certain assignments, students can be guided to compare outputs from various AI tools (e.g., ChatGPT, Gemini, Claude) with respect to reasoning quality, bias potential, and evidentiary support. By critically analyzing AI-generated outcomes, e.g., questioning the feasibility or equity implications of a spatial design, students are better equipped to engage with emerging technologies in a reflective and ethically responsible manner. Educators can further reinforce ethical engagement by designing assignments that require both the application and critical analysis of AI tools while emphasizing interpretation, judgment, and explanation of the outcome [141]. For instance, they might ask students to optimize a structural design using AI, then evaluate its long-term maintenance costs, environmental impact, or implications for underserved communities. Another example could involve using an LLM to draft a construction contract clause, followed by an assessment of its legal precision and susceptibility to misinterpretation.

As an illustrative teaching practice, inspired by the 2022 Blueprint for an AI Bill of Rights, in a graduate-level CCE course on AI for the built environment, we have incorporated scenario-based discussions on responsible design, development, and deployment of AI systems, considering safety and effectiveness, algorithmic discrimination, data privacy, notice and explanation, and human alternatives and fallback [142]. Student teams applied these principles to real-world CCE scenarios, spanning intelligent transportation systems, cost estimation and scheduling, construction robotics and automation, jobsite quality control and defect detection, water resource management, energy-efficient building design and optimization, and smart grids and demand-side energy management.

5.2. Institutional Strategies

Institutions must actively involve students in evaluating AI tools to promote transparency, build trust, and ensure that deployed technologies address real user needs. Students bring fresh insights on critical issues and perspectives that are often underrepresented in top-down administrative decisions. In the long term, institutional responsiveness to student concerns strengthens the legitimacy and effectiveness of AI integration in educational settings [143]. Universities should also utilize ethical review mechanisms similar to Institutional Review Boards (IRBs) for the vetting of instructional AI systems and assessing associated implementation risks. Building upon these review mechanisms, academic units must implement regular, programmatic ethical and performance audits of their deployed AI systems. Because educational models are prone to data drift (e.g., changing curriculum standards or student demographic shifts could alter the efficacy of the underlying algorithm over time), the system's historical evaluations must be continuously benchmarked. Among others, AI grading distributions must be cross-referenced against double-blinded evaluations executed by independent faculty panels to isolate and quantify systemic fairness gaps or output disparities.

Furthermore, building upon the recommendation for educators to document and verify AI-generated outputs, institutions should establish standardized procedures for recording AI-generated educational decisions. These records should include the original input context, prompt configuration, relevant model metadata, and any human auditor interventions. Additionally, institutions must address the issue of the digital divide by investing in equitable access to AI resources. Among other measures, this could include subsidizing software licenses, expanding access to high-performance computing, and offering financial support to ensure that all students have the means to fully engage with AI-integrated curricula. At the time this study was conducted, several U.S. universities had begun investing in institution-wide licenses for AI platforms, such as ChatGPT and Microsoft Copilot [144]. In parallel, some institutions were developing customized GPT-based tools to deliberately introduce flawed outputs, with the aim of discouraging academic misconduct and encouraging students to verify outputs, thus embedding fact-checking and critical evaluation into AI-supported learning processes [145,146]. In implementing these initiatives, it is critical that AI use is not mandated, faculty authority over pedagogy and assessment remains intact, safeguards are in place for privacy preservation and data supervision, and implementation is reviewed with input from faculty, staff, and students.

Sustaining ethical AI integration also requires structured training for faculty and staff on digital ethics, regulatory compliance, and bias mitigation. To support such efforts, institutions must provide clear guidelines on acceptable AI use in coursework [147] and invest in faculty development focused on digital ethics and equity [148,149]. Empowering instructors with the skills to critically integrate AI into their teaching ensures that students are both proficient in using these tools and prepared to confront their ethical complexities in practice. Training programs should emphasize both technical competencies and ethical discernment, equipping educators with the tools to guide students in using AI responsibly. Moreover, institutions must promote a culture of ethical inquiry and responsible innovation. This involves embedding AI ethics across the curriculum [150–152], supporting interdisciplinary research on AI (through campus-wide initiatives and seed grants), and facilitating public dialogue. Lastly, educational leaders and policymakers should prioritize ethical incorporation of AI in enhancing teaching practices, student support services, community engagement and communication, and policy formation and advocacy [153].

5.3. Governance of Personalized and Adaptive Instructional Systems

While the mechanisms above are illustrated primarily through automated assessment, personalized instruction and intelligent tutoring systems raise a parallel but distinct set of governance concerns, given their role in shaping, rather than merely scoring, a student's learning trajectory. Because these systems continuously profile learners, e.g., inferring skill level, engagement, or learning style from interaction data, erroneous profiling can silently misdirect a student toward inappropriate content and pacing. Institutions should therefore require periodic instructor review of the learning pathways assigned to a sample of students, similar to the blind-labeling audits used for grading, to check whether adaptive routing decisions correlate with demographic or linguistic characteristics rather than demonstrated need. Student autonomy is a related concern. When an adaptive system determines what a student sees next, students and instructors should have the ability to see the profile or classification driving that decision and to request a different pathway, preserving meaningful human agency over the learning process. This also requires clear procedures for reviewing and correcting AI-guided instructional decisions, mirroring the appeal processes established for grading, so that a student or instructor who disagrees with a system's placement or pacing decision has a defined route for escalation and correction, rather than having to accept an opaque recommendation. Finally, because personalized systems keep

detailed longitudinal records of student behavior and inferred characteristics, institutions must guard against secondary use of these records, i.e., their repurposing for administrative, disciplinary, or evaluative decisions beyond the original instructional purpose for which the data were collected, consistent with the data minimization principles embedded in FERPA and GDPR discussed earlier. These considerations carry particular weight in CCE education, where graduates will later be responsible for safety-critical professional judgments. Therefore, fostering early habits of scrutinizing, questioning, and correcting AI-guided recommendations, rather than deferring to them, is itself part of preparing students for that responsibility.

6. Summary and Conclusions

The integration of AI into CCE education represents a transformative shift, offering new opportunities for enhanced learning experiences, personalized instruction, and streamlined administrative processes. While AI holds immense promise for preparing future engineers for an increasingly AI-driven professional landscape, it also poses major concerns, such as hallucination, academic integrity, and impacts on critical thinking and problem-solving skills, the latter two being cornerstones of engineering education and practice. This paper highlighted key ethical concerns, including the potential for algorithmic bias, the significant privacy risks, the challenge of achieving transparency, the complex issues surrounding academic integrity, and the exacerbation of the digital divide. Each of these challenges necessitates a clear understanding of and strong commitment to upholding the core values of integrity, equity, and public welfare that fundamentally underpin the CCE profession, highlighting the parallel ethical responsibilities shared by AI-enabled construction practice and CCE education. To address this complexity, we proposed an ethical output governance framework and practical guidelines for responsible AI use and integration in CCE education, founded on the principles of informed consent, equity, accountability, and transparency. Furthermore, the discussions emphasized the importance of understanding existing regulations and the development of institutional policies to address the unique risks inherent in educational AI.

Ultimately, the future of the CCE domain depends on preparing professionals who combine technical proficiency with a strong ethical foundation. While we advocate for the responsible use of AI to enhance efficiency in certain educational and training contexts, such use cases must be designed with the mindset that AI tools do not replace fundamental educational goals, such as developing analytical and problem-solving skills and effective communication. By adopting practical recommendations, meticulously designing ethically engaging assignments, and promoting a culture of responsible innovation, CCE programs can ensure that AI serves as a powerful force for good. This proactive approach will enable future engineers to ethically use AI, making informed decisions that uphold public safety, promote social equity, and advance sustainable development. The ethical compass, therefore, is not merely an optional add-on but an integral and indispensable part of the AI-enabled CCE education by ensuring that innovation consistently aligns with the enduring values of the profession.

Limitations and Opportunities for Future Work

Several limitations of this work should be acknowledged. First, the paper is based on a narrative rather than a systematic literature review. Database selection, snowball sampling, the geographic representation of the selected sources, and the inclusion of English-language sources only may have introduced selection bias. Second, the proposed framework is a conceptual synthesis rather than an empirically validated model. While its component mechanisms are based on established governance frameworks and the existing literature

on AI in education and engineering ethics, the output governance framework proposed in this study has not been piloted or tested within a CCE program. The framework may also require different levels of resources depending on institutional capacity. Third, this synthesis draws predominantly on U.S. and European sources, reflecting the geographic concentration of published research on AI in education and AI ethics. CCE programs in other regions are therefore underrepresented in the sources informing this framework. The framework's applicability to these contexts should not be assumed without further investigation and local validation. On a relevant note, because the framework has not yet been piloted in these contexts, its applicability beyond the reviewed literature remains to be empirically established. Finally, the rapidly evolving nature of AI presents an inherent limitation to the continued relevance of the proposed framework. Accordingly, the framework must be reviewed and updated regularly to reflect technological advances, regulatory requirements and audit findings, institutional AI policies, and evolving data-governance practices.

Future research should empirically evaluate the proposed framework across different educational and institutional contexts. First, the framework should be tested across representative AI-supported applications in CCE education, such as automated grading, cost estimation, and scheduling, to assess whether its governance mechanisms improve the accuracy, consistency, transparency, and accountability of AI-assisted decisions. Second, comparative studies should examine the applicability of the framework across geographic and institutional contexts beyond those represented in the current literature to determine whether differences in regulatory environments, institutional capacity, educational practices, and cultural attitudes toward AI affect its implementation. Third, survey- and interview-based studies involving faculty, students, administrators, and institutions could evaluate stakeholder perceptions of AI-supported education, including perceived benefits, barriers, risks, acceptable levels of automation, and expectations regarding human oversight.

Author Contributions: Conceptualization, A.D. and A.H.B.; methodology, A.D. and A.H.B.; formal analysis, A.D. and A.H.B.; investigation, A.D. and A.H.B.; resources, A.D. and A.H.B.; writing—original draft preparation, A.D. and A.H.B.; writing—review and editing, A.D. and A.H.B.; visualization, A.D. and A.H.B.; supervision, A.H.B.; project administration, A.H.B.; funding acquisition, A.H.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by an internal grant from the College of Engineering and Applied Science of the University of Colorado Boulder. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the University of Colorado Boulder.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. No empirical data, software code, or computational models were generated. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors used ChatGPT (GPT-5.2, OpenAI, San Francisco, CA, USA) solely to assist with language editing, including grammar and sentence flow. No AI was used to generate scientific content, analyses, interpretations, or conclusions. All content was reviewed and approved by the authors, who take full responsibility for the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chng, E.; Tan, A.L.; Tan, S.C. Examining the Use of Emerging Technologies in Schools: A Review of Artificial Intelligence and Immersive Technologies in STEM Education. *J. STEM Educ. Res.* **2023**, *6*, 385–407. [CrossRef]

2. Zhu, Y.; Jafari, A.; Behzadan, A.H.; Issa, R.R. A Roadmap to the Next-Generation Technology-Enabled Learning-Centered Environments in AEC Education. *J. Civ. Eng. Educ.* **2024**, *150*, 05024002. [CrossRef]
3. Ilbeigi, M.; Darestani Farahani, P. Future of Construction Engineering Education in the Automation Era. *J. Civ. Eng. Educ.* **2025**, *151*, 04025007. [CrossRef]
4. Kaswan, K.S.; Dhatterwal, J.S.; Ojha, R.P. AI in personalized learning. In *Advances in Technological Innovations in Higher Education: Theory and Practices*, 1st ed.; Garg, A., Babu, B.V., Balas, V.E., Eds.; CRC Press: Boca Raton, FL, USA, 2024; pp. 103–117. [CrossRef]
5. Maghsudi, S.; Lan, A.; Xu, J.; van der Schaar, M. Personalized Education in the Artificial Intelligence Era: What to Expect Next. *IEEE Signal Process. Mag.* **2021**, *38*, 37–50. [CrossRef]
6. Popenici, S.A.; Kerr, S. Exploring the impact of artificial intelligence on teaching and learning in higher education. *Res. Pract. Technol. Enhanc. Learn.* **2017**, *12*, 22. [CrossRef] [PubMed]
7. Filimonovic, D.; Rutzer, C.; Wunsch, C. Can GenAI Improve Academic Performance? Evidence from the Social and Behavioral Sciences. *arXiv* **2025**, arXiv:2510.02408. [CrossRef]
8. Cheng, J.Y.; Devkota, A.; Gheisari, M.; Jeelani, I.; Franz, B.W. AI Integration in Construction Education: A Survey on Awareness, Attitudes, and Implementation of AI among US and International Faculties. *J. Civ. Eng. Educ.* **2025**, *151*, 04025009. [CrossRef]
9. Dimari, A.; Tyagi, N.; Davanageri, M.; Kukreti, R.; Yadav, R.; Dimari, H. AI-Based Automated Grading Systems for open book examination system: Implications for Assessment in Higher Education. In Proceedings of the 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), Chikkaballapur, India, 18–19 April 2024. [CrossRef]
10. Gambo, I.; Abegunde, F.J.; Gambo, O.; Ogundokun, R.; Babatunde, A.; Lee, C.C. GRAD-AI: An automated grading tool for code assessment and feedback in programming course. *Educ. Inf. Technol.* **2025**, *30*, 9859–9899. [CrossRef]
11. Swiecki, Z.; Khosravi, H.; Chen, G.; Martinez-Maldonado, R.; Lodge, J.M.; Milligan, S.; Selwyn, N.; Gašević, D. Assessment in the age of artificial intelligence. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100075. [CrossRef]
12. Lin, C.C.; Huang, A.Y.Q.; Lu, O.H.T. Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learn. Environ.* **2023**, *10*, 41. [CrossRef]
13. Lin, C.J.; Wang, W.S.; Lee, H.Y.; Li, P.H.; Huang, Y.M.; Wu, T.T. Advancing self-directed learning in STEM education: Integrating GPT-based learning aid with multimodal learning analytics. *J. Res. Technol. Educ.* **2025**, 1–19. [CrossRef]
14. Pozdniakov, S.; Brazil, J.; Banihashem, S.K.; Noroozi, O.; Gašević, D.; Sadiq, S.; Khosravi, H. AI assistance in peer feedback provision: Pedagogically sound, but minimally adopted. *Comput. Educ.* **2026**, *248*, 105591. [CrossRef]
15. Shalong, W.; Yi, Z.; Bin, Z.; Ganglei, L.; Jinyu, Z.; Yanwen, Z.; Zequn, Z.; Lianwen, Y.; Feng, R. Enhancing self-directed learning with custom GPT AI facilitation among medical students: A randomized controlled trial. *Med. Teach.* **2025**, *47*, 1126–1133. [CrossRef] [PubMed]
16. Uddin, S.M.J.; Albert, A.; Ovid, A.; Alsharef, A. Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. *Sustainability* **2023**, *15*, 7121. [CrossRef]
17. Kulshrestha, A.; Gupta, A.; Singh, U.; Sharma, A.; Shukla, A.; Gautam, R.; Kumar, P.; Pandey, D. AI-based Exam Proctoring System. In Proceedings of the 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 11–12 May 2023. [CrossRef]
18. Tejaswi, P.; Venkatramaphanikumar, S.; Venkata Krishna Kishore, K. Proctor net: An AI framework for suspicious activity detection in online proctored examinations. *Measurement* **2023**, *206*, 112266. [CrossRef]
19. NSPE (National Society of Professional Engineers). Code of Ethics for Engineers. Available online: https://www.nspe.org/career-growth/ethics/nspe-code-ethics-engineers (accessed on 19 February 2026).
20. Fruchter, R.; Gluck, J.; Gold, Y.I. Application of AI programming techniques to the analysis of structures. *Comput. Struct.* **1988**, *30*, 747–753. [CrossRef]
21. Xiang, Y.; Liu, K.; Su, T.; Li, J.; Ouyang, S.; Mao, S.S.; Geimer, M. An Extension of BIM Using AI: A Multi Working-Machines Pathfinding Solution. *IEEE Access* **2021**, *9*, 124583–124599. [CrossRef]
22. Zhang, F.; Chan, A.P.; Darko, A.; Chen, Z.; Li, D. Integrated applications of building information modeling and artificial intelligence techniques in the AEC/FM industry. *Autom. Constr.* **2022**, *139*, 104289. [CrossRef]
23. Cha, Y.J.; Ali, R.; Lewis, J.; Büyüköztürk, O. Deep learning-based structural health monitoring. *Autom. Constr.* **2024**, *161*, 105328. [CrossRef]
24. Nath, N.D.; Behzadan, A.H.; Paal, S.G. Deep learning for site safety: Real-time detection of personal protective equipment. *Autom. Constr.* **2020**, *112*, 103085. [CrossRef]

25. Ziolkowski, P.; Niedostatkiewicz, M. Machine Learning Techniques in Concrete Mix Design. *Materials* **2019**, *12*, 1256. [CrossRef] [PubMed]
26. The White House. Winning the Race: America's AI Action Plan. Available online: https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf (accessed on 19 February 2026).
27. Baker, R.S.; Hawn, A. Algorithmic Bias in Education. *Int. J. Artif. Intell. Educ.* **2022**, *32*, 1052–1092. [CrossRef]
28. Kizilcec, R.F.; Lee, H. Algorithmic fairness in education. In *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*, 1st ed.; Holmes, W., Porayska-Pomsta, K., Eds.; Routledge: New York, NY, USA, 2022; pp. 174–202. [CrossRef]
29. Brown, M.; Klein, C. Whose Data? Which Rights? Whose Power? A Policy Discourse Analysis of Student Privacy Policy Documents. *J. High. Educ.* **2020**, *91*, 1149–1178. [CrossRef]
30. Hollanek, T. AI transparency: A matter of reconciling design with critique. *AI Soc.* **2023**, *38*, 2071–2079. [CrossRef]
31. Hysaj, A.; Dean, B.; Freeman, M. Exploring the purposes and uses of generative artificial intelligence tools in academic writing for multicultural students. *High. Educ. Res. Dev.* **2025**, *44*, 1686–1700. [CrossRef]
32. Lund, B.; Mannuru, N.R.; Teel, Z.A.; Lee, T.H.; Ortega, N.J.; Simmons, S.; Ward, E. Student Perceptions of AI-Assisted Writing and Academic Integrity: Ethical Concerns, Academic Misconduct, and Use of Generative AI in Higher Education. *AI Educ.* **2025**, *1*, 2. [CrossRef]
33. Sukhera, J. Narrative Reviews: Flexible, Rigorous, and Practical. *J. Grad. Med. Educ.* **2022**, *14*, 414–417. [CrossRef] [PubMed]
34. Byrne, J.A. Improving the peer review of narrative literature reviews. *Res. Integr. Peer Rev.* **2016**, *1*, 12. [CrossRef] [PubMed]
35. Chen, Z. Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanit. Soc. Sci. Commun.* **2023**, *10*, 567. [CrossRef]
36. Chen, F.; Wang, L.; Hong, J.; Jiang, J.; Zhou, L. Unmasking bias in artificial intelligence: A systematic review of bias detection and mitigation strategies in electronic health record-based models. *J. Am. Med. Inform. Assoc.* **2024**, *31*, 1172–1183. [CrossRef] [PubMed]
37. Ferrer, X.; van Nuenen, T.; Such, J.M.; Coté, M.; Criado, N. Bias and Discrimination in AI: A Cross-Disciplinary Perspective. *IEEE Technol. Soc. Mag.* **2021**, *40*, 72–80. [CrossRef]
38. Hall, P.; Ellis, D. A systematic review of socio-technical gender bias in AI algorithms. *Online Inf. Rev.* **2023**, *47*, 1264–1279. [CrossRef]
39. Sun, W.; Nasraoui, O.; Shafto, P. Evolution and impact of bias in human and machine learning algorithm interaction. *PLoS ONE* **2020**, *15*, e0235502. [CrossRef] [PubMed]
40. Goldshtein, M.; Alhashim, A.G.; Roscoe, R.D. Automating Bias in Writing Evaluation: Sources, Barriers, and Recommendations. In *The Routledge International Handbook of Automated Essay Evaluation*, 1st ed.; Shermis, M.D., Wilson, J., Eds.; Routledge: New York, NY, USA, 2024; pp. 421–444. [CrossRef]
41. Madden, M.; Calvin, A.; Hasse, A.; Lenhart, A. *The Dawn of the AI Era: Teens, Parents, and the Adoption of Generative AI at Home and School*; Common Sense: San Francisco, CA, USA, 2024.
42. Tanksley, T.; Cabral, B. "This Kind of Technology Can. . . Treat Students Like Threats": Black Youth Experiences, Reflections, and Articulations of Digital Discipline Under the New Jim Code. *Youth* **2026**, *6*, 12. [CrossRef]
43. Jia, J.; He, Y. The design, implementation and pilot application of an intelligent online proctoring system for online exams. *Interact. Technol. Smart Educ.* **2021**, *19*, 112–120. [CrossRef]
44. Verma, P.; Malhotra, N.; Suri, R.; Kumar, R. Automated smart artificial intelligence-based proctoring system using deep learning. *Soft Comput.* **2024**, *28*, 3479–3489. [CrossRef]
45. Conijn, R.; Kleingeld, A.; Matzat, U.; Snijders, C. The fear of big brother: The potential negative side-effects of proctored exams. *J. Comput. Assist. Learn.* **2022**, *38*, 1521–1534. [CrossRef]
46. Pokorny, A.; Ballen, C.J.; Drake, A.G.; Driessen, E.P.; Fagbodun, S.; Gibbens, B.; Henning, J.A.; McCoy, S.J.; Thompson, S.K.; Willis, C.G.; et al. "Out of my control": Science undergraduates report mental health concerns and inconsistent conditions when using remote proctoring software. *Int. J. Educ. Integr.* **2023**, *19*, 22. [CrossRef]
47. Prinsloo, P.; Slade, S. Student Vulnerability, Agency and Learning Analytics: An Exploration. *J. Learn. Anal.* **2016**, *3*, 159–182. [CrossRef]
48. UNESCO. *Minding the Data: Protecting Learners' Privacy and Security*; UNESCO: Paris, France, 2022.
49. Adadi, A.; Berrada, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [CrossRef]
50. Gillani, N.; Eynon, R.; Chiabaut, C.; Finkel, K. Unpacking the "Black Box" of AI in Education. *Educ. Technol. Soc.* **2023**, *26*, 99–111. [CrossRef]
51. von Eschenbach, W.J. Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philos. Technol.* **2021**, *34*, 1607–1622. [CrossRef]
52. Karpatne, A.; Deshwal, A.; Jia, X.; Ding, W.; Steinbach, M.; Zhang, A.; Kumar, V. AI-enabled scientific revolution in the age of generative AI: Second NSF workshop report. *npj Artif. Intell.* **2025**, *1*, 18. [CrossRef]

53. Hall, E.; Seyam, M.; Dunlap, D. Exploring Explainability and Transparency in Automated Essay Scoring Systems: A User-Centered Evaluation. In Proceedings of the International Conference on Human-Computer Interaction (HCII 2024), Washington, DC, USA, 29 June–4 July 2024. [CrossRef]
54. Jackson, S.; Panteli, N. Trust or mistrust in algorithmic grading? An embedded agency perspective. *Int. J. Inf. Manag.* **2023**, *69*, 102555. [CrossRef]
55. Zhai, C.; Wibowo, S.; Li, L.D. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learn. Environ.* **2024**, *11*, 28. [CrossRef]
56. Meça, A.; Shkëlzeni, N. Academic Integrity in the Face of Generative Language Models. In Proceedings of the 6th EAI International Conference on Emerging Technologies in Computing (iCETiC 2023), Southend-on-Sea, UK, 17–18 August 2023; pp. 58–70. [CrossRef]
57. Sellnow, D.D. Reflection-AI: Exploring the challenges and opportunities of artificial intelligence in higher education. *Front. Commun.* **2025**, *10*, 1615040. [CrossRef]
58. Deng, R.; Jiang, M.; Yu, X.; Lu, Y.; Liu, S. Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Comput. Educ.* **2025**, *227*, 105224. [CrossRef]
59. Melisa, R.; Ashadi, A.; Triastuti, A.; Hidayati, S.; Salido, A.; Ero, P.E.L.; Marlini, C.; Zefrin, Z.; Al Fuad, Z. Critical Thinking in the Age of AI: A Systematic Review of AI's Effects on Higher Education. *Educ. Process Int. J.* **2025**, *14*, e2025031. [CrossRef]
60. OECD (Organization for Economic Co-operation and Development). *OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education*; OECD Publishing: Paris, France, 2026. [CrossRef]
61. Watts, F.M.; Dood, A.J.; Shultz, G.V.; Rodriguez, J.M.G. Comparing Student and Generative Artificial Intelligence Chatbot Responses to Organic Chemistry Writing-to-Learn Assignments. *J. Chem. Educ.* **2023**, *100*, 3806–3817. [CrossRef]
62. Eslami, M.; Collins, P.; Queen, B. How Generative AI Is Reshaping Student Writing: A Data-Driven Perspective for Writing Instructors. *Educ. Sci.* **2026**, *16*, 1. [CrossRef]
63. Sanz-Tejeda, A.; Domínguez-Oller, J.C.; Baldaquí-Escandell, J.M.; Gómez-Díaz, R.; García-Rodríguez, A. The impact of generative AI on academic reading and writing: A synthesis of recent evidence (2023–2025). *Front. Educ.* **2025**, *10*, 1711718. [CrossRef]
64. Giray, L. The Problem with False Positives: AI Detection Unfairly Accuses Scholars of AI Plagiarism. *Ser. Libr.* **2024**, *85*, 181–189. [CrossRef]
65. Khalil, M.; Er, E. Will ChatGPT Get You Caught? Rethinking of Plagiarism Detection. In Proceedings of the International Conference on Human-Computer Interaction (HCII 2023), Copenhagen, Denmark, 23–28 July 2023. [CrossRef]
66. Kitsara, I. Artificial Intelligence and the Digital Divide: From an Innovation Perspective. In *Platforms and Artificial Intelligence*, 1st ed.; Bounfour, A., Ed.; Springer: Cham, Switzerland, 2022; pp. 245–265. [CrossRef]
67. Srinuan, C.; Bohlin, E. Understanding the digital divide: A literature survey and ways forward. In Proceedings of the 22nd European Regional Conference of the International Telecommunications Society (ITS), Budapest, Hungary, 18–21 September 2011.
68. Bentley, S.V.; Naughtin, C.K.; McGrath, M.J.; Irons, J.L.; Cooper, P.S. The digital divide in action: How experiences of digital technology shape future relationships with artificial intelligence. *AI Ethics* **2024**, *4*, 901–915. [CrossRef]
69. Lythreatis, S.; Singh, S.K.; El-Kassar, A.N. The digital divide: A review and future research agenda. *Technol. Forecast. Soc. Change* **2022**, *175*, 121359. [CrossRef]
70. Bulathwela, S.; Pérez-Ortiz, M.; Holloway, C.; Cukurova, M.; Shawe-Taylor, J. Artificial Intelligence Alone Will Not Democratise Education: On Educational Inequality, Techno-Solutionism and Inclusive Tools. *Sustainability* **2024**, *16*, 781. [CrossRef]
71. Kachakova, V. Crisis of Responsibility: What Does ChatGPT Leave Unsaid About the Future of education? *Sociol. Probl. J.* **2024**, *56*, 251–265. [CrossRef]
72. Louis, M.; ElAzab, M. Will AI replace Teacher? *Int. J. Internet Educ.* **2023**, *22*, 9–21. [CrossRef]
73. Taufikin, M.S.I.; Supa'At; Azifah, N.; Nikmah, F.; Kuanr, J.; Parminder. The Impact of AI on Teacher Roles and Pedagogy in the 21st Century Classroom. In Proceedings of the 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), Chikkaballapur, India, 18–19 April 2024. [CrossRef]
74. Cho, M.; Park, M.; Park, J.; Park, S.K.; Cha, G.; Park, S. From prediction to prioritization: Automated framework for delay risk management in highway earthwork projects. *Autom. Constr.* **2026**, *181*, 106588. [CrossRef]
75. Gondia, A.; Siam, A.; El-Dakhakhni, W.; Nassar, A.H. Machine Learning Algorithms for Construction Projects Delay Risk Prediction. *J. Constr. Eng. Manag.* **2020**, *146*, 04019085. [CrossRef]
76. Yao, Y.; Tam, V.W.Y.; Wang, J.; Le, K.N.; Butera, A. Automated construction scheduling using deep reinforcement learning with valid action sampling. *Autom. Constr.* **2024**, *166*, 105622. [CrossRef]
77. Attalla, M.; Hegazy, T. Predicting Cost Deviation in Reconstruction Projects: Artificial Neural Networks versus Regression. *J. Constr. Eng. Manag.* **2003**, *129*, 405–411. [CrossRef]
78. Williams, T.P.; Gong, J. Predicting construction cost overruns using text mining, numerical data and ensemble classifiers. *Autom. Constr.* **2014**, *43*, 23–29. [CrossRef]

79. Fang, W.; Ding, L.; Luo, H.; Love, P.E.D. Falls from heights: A computer vision-based approach for safety harness detection. *Autom. Constr.* **2018**, *91*, 53–61. [CrossRef]
80. Hou, X.; Li, C.; Fang, Q. Computer vision-based safety risk computing and visualization on construction sites. *Autom. Constr.* **2023**, *156*, 105129. [CrossRef]
81. Ayhan, B.U.; Tokdemir, O.B. Safety assessment in megaprojects using artificial intelligence. *Saf. Sci.* **2019**, *118*, 273–287. [CrossRef]
82. Sudharsan, R.; Vinayagam, K. The Role of Artificial Intelligence in Behavioural Safety and Strategy in Construction Industry. In Proceedings of the 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 15–16 March 2024. [CrossRef]
83. Awolusi, I.; Marks, E.; Hallowell, M. Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Autom. Constr.* **2018**, *85*, 96–106. [CrossRef]
84. Kim, K.; Kim, S.; Shchur, D. A UAS-based work zone safety monitoring system by integrating internal traffic control plan (ITCP) and automated object detection in game engine environment. *Autom. Constr.* **2021**, *128*, 103736. [CrossRef]
85. Li, L.; Huang, Z.; Wang, J.; Du, B.; Dai, L. Automated construction monitoring based on computer vision: A comprehensive review. *Dev. Built Environ.* **2026**, *25*, 100832. [CrossRef]
86. Debnath, R.; Veeraraghavan P, V.; Hapse, N. AI and Privacy: Ethical Concerns in Data Collection and Surveillance. *Int. J. Multidiscip. Res.* **2024**, *6*, 1–8. [CrossRef]
87. Elliott, D.; Soifer, E. AI Technologies, Privacy, and Security. *Front. Artif. Intell.* **2022**, *5*, 826737. [CrossRef] [PubMed]
88. Fontes, C.; Hohma, E.; Corrigan, C.C.; Lütge, C. AI-powered public surveillance systems: Why we (might) need them and how we want them. *Technol. Soc.* **2022**, *71*, 102137. [CrossRef]
89. Liang, C.J.; Le, T.H.; Ham, Y.; Mantha, B.R.K.; Cheng, M.H.; Lin, J.J. Ethics of artificial intelligence and robotics in the architecture, engineering, and construction industry. *Autom. Constr.* **2024**, *162*, 105369. [CrossRef]
90. Speiser, K.; Seiß, S.; Boukamp, F.; Melzner, J.; Teizer, J. From fragmented data to unified construction safety knowledge: A process-based ontology framework for safer work. *Autom. Constr.* **2025**, *176*, 106293. [CrossRef]
91. Van Tam, N. How generative AI reshapes construction and built environment: The good, the bad, and the ugly. *Build. Environ.* **2025**, *284*, 113526. [CrossRef]
92. Behzadan, A.H. Formalizing Trust in Artificial Intelligence for Built Environment Decision-Making. In Proceedings of the AHFE 2024 International Conference on Human Factors in Design, Engineering, and Computing (AHFE 2024 Hawaii Edition), Honolulu, HI, USA, 8–10 December 2024. [CrossRef]
93. Helbing, D.; Fanitabasi, F.; Giannotti, F.; Hänggli, R.; Hausladen, C.I.; van den Hoven, J.; Mahajan, S.; Pedreschi, D.; Pournaras, E. Ethics of Smart Cities: Towards Value-Sensitive Design and Co-Evolving City Life. *Sustainability* **2021**, *13*, 11162. [CrossRef]
94. Gerke, S.; Minssen, T.; Cohen, G. Ethical and legal challenges of artificial intelligence-driven healthcare. In *Artificial Intelligence in Healthcare*, 1st ed.; Bohr, A., Memarzadeh, K., Eds.; Academic Press: London, UK, 2020; pp. 295–336. [CrossRef]
95. Roig, A. Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR). *Eur. J. Law Technol.* **2017**, *8*, 3. Available online: https://www.google.com/goto?url=CAESdQHrOzAV7rfAdtQO9rQ1w_SDCJ7PEYaEIJ8yytTccjp37kjn7BZmFTo9Xfs-GcyMsrJ_2mpNbKcbEwkk88AmNWqCZDwmSdK4WUvzx02TVbPEyHCVS7yTEofIJb1sF4RF1MSug0yzMAYZr6a6NBRXB7f1dsEAZA (accessed on 10 September 2026).
96. Zhan, J.; Fang, W.; Love, P.E.D.; Luo, H. Explainable Artificial Intelligence: Counterfactual Explanations for Risk-Based Decision-Making in Construction. *IEEE Trans. Eng. Manag.* **2024**, *71*, 10667–10685. [CrossRef]
97. Behzadan, A.; Dabiri, A. Factors influencing human trust in intelligent built environment systems. *AI Ethics* **2025**, *5*, 5841–5855. [CrossRef]
98. Liu, B. In AI We Trust? Effects of Agency Locus and Transparency on Uncertainty Reduction in Human–AI Interaction. *J. Comput.-Mediat. Commun.* **2021**, *26*, 384–402. [CrossRef]
99. Zarghami, S.; Kouchaki, H.; Yang, L.; Rodriguez, P.M. Explainable Artificial Intelligence in Generative Design for Construction. In Proceedings of the 2024 European Conference on Computing in Construction (EC3 2024), Chania, Greece, 14–17 July 2024. [CrossRef]
100. Kovari, A. Ethical use of ChatGPT in education—Best practices to combat AI-induced plagiarism. *Front. Educ.* **2025**, *9*, 1465703. [CrossRef]
101. Kothapalli, S. Computer Vision-Enabled Safety for Construction: A Deep Learning and Predictive Analytics Approach Across Project. *Int. J. Innov. Res. Eng. Multidiscip. Phys. Sci.* **2024**, *12*, 232598. [CrossRef]
102. Elelu, K.; Le, T.; Le, C. Collision Hazard Detection for Construction Worker Safety Using Audio Surveillance. *J. Constr. Eng. Manag.* **2023**, *149*, 04022159. [CrossRef]
103. Hoang, P.D.; Lee, J.; Choi, J.; Ahn, Y. Review the Application of ML in Construction Equipment Safety Management. In Proceedings of the Third International Conference on Sustainable Civil Engineering and Architecture (ICSCEA 2023), Da Nang City, Vietnam, 19–21 July 2023. [CrossRef]

104. Lee, Y.S.; Kim, D.K.; Kim, J.H. Deep-Learning-Based Anti-Collision System for Construction Equipment Operators. *Sustainability* **2023**, *15*, 16163. [CrossRef]
105. Katooziani, A.; Jeelani, I.; Gheisari, M. GPT Applications for Construction Safety: A Use Case Analysis. *Buildings* **2025**, *15*, 2410. [CrossRef]
106. NSC (National Safety Council). *Exploring the Role of Generative AI in Occupational Environment, Health and Safety*; Campbell Institute, National Safety Council: Itasca, IL, USA, 2025.
107. Sammour, F.; Xu, J.; Wang, X.; Hu, M.; Zhang, Z. Responsible AI in Construction Safety: Systematic Evaluation of Large Language Models and Prompt Engineering. *J. Constr. Eng. Manag.* **2026**, *152*, 04025217. [CrossRef]
108. Rabbi, A.B.K.; Jeelani, I. AI integration in construction safety: Current state, challenges, and future opportunities in text, vision, and audio based applications. *Autom. Constr.* **2024**, *164*, 105443. [CrossRef]
109. Zaidi, S.F.A.; Yang, J.; Abbas, M.S.; Hussain, R.; Lee, D.; Park, C. Vision-Based Construction Safety Monitoring Utilizing Temporal Analysis to Reduce False Alarms. *Buildings* **2024**, *14*, 1878. [CrossRef]
110. Copeland, A.M. What it means to be an engineer of record. In Proceedings of the 6th International Mining and Industrial Waste Management Conference, Limpopo, South Africa, 29–31 October 2018.
111. Gillum, J.D. The Engineer of Record and Design Responsibility. *J. Perform. Constr. Facil.* **2000**, *14*, 67–70. [CrossRef]
112. NCEES (National Council of Examiners for Engineering and Surveying). *Model Law*; NCEES: Clemson, SC, USA, 2021.
113. NSPE (National Society of Professional Engineers). *NSPE Position Statement No. 10-1778: Responsible Charge*; NSPE: Alexandria, VA, USA, 2024. Available online: https://www.nspe.org/nspe-advocacy/explore-issues/professional-policies-and-position-statements/responsible-charge (accessed on 30 July 2026).
114. Savaş, S. Artificial intelligence in construction project management: Trends, challenges and future directions. *J. Des. Resil. Archit. Plan.* **2025**, *6*, 221–238. [CrossRef]
115. NSPE (National Society of Professional Engineers). Artificial Intelligence: NSPE Position Statement No. 03-1774. Available online: https://www.nspe.org/nspe-advocacy/explore-issues/professional-policies-and-position-statements/artificial-intelligence (accessed on 30 July 2026).
116. Raybon, Z. *From the Executive Director: Regulation in the AI Era*; Florida Board of Professional Engineers: Tallahassee, FL, USA, 2025. Available online: https://fbpe.org/regulation-in-the-ai-era/ (accessed on 10 August 2026).
117. Texas Board of Professional Engineers and Land Surveyors. *Policy Advisory Opinion Regarding the Use of Artificial Intelligence Software by Licensees Under the Board's Jurisdiction (Policy Advisory Request No. 71)*; Texas Board of Professional Engineers and Land Surveyors: Austin, TX, USA, 2024. Available online: https://pels.texas.gov/nm/2024/pao-71-response.pdf (accessed on 10 August 2026).
118. North Carolina Board of Examiners for Engineers and Surveyors. *Guidelines on the Use of Artificial Intelligence in Engineering and Surveying Services*; North Carolina Board of Examiners for Engineers and Surveyors: Raleigh, NC, USA, 2025. Available online: https://www.ncbels.org/wp-content/uploads/2025/07/Guidelines-on-the-Use-of-Artificial-Intelligence-in-Engineering-and-Surveying-Services-FINAL.pdf (accessed on 10 August 2026).
119. Pathmabandu, C.; Grundy, J.; Chhetri, M.B.; Baig, Z. Privacy for IoT: Informed consent management in Smart Buildings. *Future Gener. Comput. Syst.* **2023**, *145*, 367–383. [CrossRef]
120. Ogunade, T.; Aigbavboa, C.; Tunji-Olayeni, P. Strategies for Fair Machine Learning Applications in the South African Construction. In Proceedings of the Future Technologies Conference (FTC 2025), Munich, Germany, 6–7 November 2025. [CrossRef]
121. Love, P.E.D.; Fang, W.; Matthews, J.; Luo, H. Open the 'Black Box' of Large Language Models and Make Them Explainable in Construction. *IEEE Eng. Manag. Rev.* **2026**, 1–12. [CrossRef]
122. U.S. Congress. *Family Educational Rights and Privacy Act (FERPA)*; 20 U.S.C. § 1232g; U.S. Government Publishing Office: Washington, DC, USA, 1974.
123. EU (European Union). General Data Protection Regulation (GDPR). Available online: https://gdpr.eu/tag/gdpr/ (accessed on 20 February 2026).
124. Wolford, B. Data Protection Impact Assessment (DPIA). Available online: https://gdpr.eu/data-protection-impact-assessment-template/ (accessed on 27 January 2026).
125. Holmes, W.; Miao, F. *Guidance for Generative AI in Education and Research*; UNESCO Publishing: Paris, France, 2023.
126. Abraham, R.; Schneider, J.; vom Brocke, J. Data governance: A conceptual framework, structured review, and research agenda. *Int. J. Inf. Manag.* **2019**, *49*, 424–438. [CrossRef]
127. Janssen, M.; Brous, P.; Estevez, E.; Barbosa, L.S.; Janowski, T. Data governance: Organizing data for trustworthy Artificial Intelligence. *Gov. Inf. Q.* **2020**, *37*, 101493. [CrossRef]
128. Cotton, D.R.E.; Cotton, P.A.; Shipway, J.R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innov. Educ. Teach. Int.* **2024**, *61*, 228–239. [CrossRef]
129. McGrath, C.; Cerratto Pargman, T.; Juth, N.; Palmgren, P.J. University teachers' perceptions of responsibility and artificial intelligence in higher education—An experimental philosophical study. *Comput. Educ. Artif. Intell.* **2023**, *4*, 100139. [CrossRef]

130. An, Y.; Yu, J.H.; James, S. Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *Int. J. Educ. Technol. High. Educ.* **2025**, *22*, 10. [CrossRef]
131. Wei, W.; Liu, L. Trustworthy Distributed AI Systems: Robustness, Privacy, and Governance. *ACM Comput. Surv.* **2025**, *57*, 143. [CrossRef]
132. Buijsman, S.; Veluwenkamp, H. Spotting When Algorithms Are Wrong. *Minds Mach.* **2023**, *33*, 541–562. [CrossRef]
133. Tan, C.; Xu, X.; Ma, L. Can humans timely and accurately perceive silent failures in automated decision aids? Effects of failure severity, initial reliability, and error type. *Int. J. Ind. Ergon.* **2026**, *113*, 103924. [CrossRef]
134. Goulas, S.; Megalokonomou, R.; Sotirakopoulos, P. Why do experts miss AI's errors? Evidence from a randomized labeling experiment. *PNAS Nexus* **2026**, *5*, 146. [CrossRef] [PubMed]
135. Burns, M.; Winthrop, R.; Luther, N.; Venetis, E.; Karim, R. *A New Direction for Students in an AI World: Prosper, Prepare, Protect*; The Brookings Institution: Washington, DC, USA, 2026.
136. *BS ISO/IEC/IEEE 15288:2023*; Systems and Software Engineering—System Life Cycle Processes. British Standards Institution: London, UK, 2023.
137. *IEEE Std 1012-2016*; IEEE Standard for System, Software, and Hardware Verification and Validation. Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2017.
138. Debbabi, M.; Hassaïne, F.; Jarraya, Y.; Soeanu, A.; Alawneh, L. *Verification and Validation in Systems Engineering: Assessing UML/SysML Design Models*; Springer: Berlin/Heidelberg, Germany, 2010.
139. Jiao, R. Demystifying verification and validation in design research methodology—From underdeveloped practice to methodological rigour. *J. Eng. Des.* **2026**, *37*, 1143–1162. [CrossRef]
140. Garcia, M.B. Transparency, Ethical Framing, and User Agency as Determinants of Trust in AI-Mediated Assessment: Informing the Design of Trustworthy Systems. *Eval. Rev.* **2026**, 0193841X261451124. [CrossRef] [PubMed]
141. Chisale, P.B. Protecting creativity in the age of generative AI: Productive uncertainty, and visible thinking in scholarship and assessment. *Front. Educ.* **2025**, *10*, 1694819. [CrossRef]
142. The White House. The Blueprint for an AI Bill of Rights. Available online: https://www.privacysecurityacademy.com/wp-content/uploads/2022/09/EXCERPT-Biden-Blueprint-for-AI-Bill-of-Rights.pdf (accessed on 1 January 2025).
143. Brossi, L.; Castillo, A.M.; Cortesi, S. Student-centred requirements for the ethics of AI in education. In *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*, 1st ed.; Holmes, W., Porayska-Pomsta, K., Eds.; Routledge: New York, NY, USA, 2022; pp. 91–112.
144. Xiao, P.; Chen, Y.; Bao, W. Waiting, Banning, and Embracing: An Empirical Analysis of Adapting Policies for Generative AI in Higher Education. *arXiv* **2023**, arXiv:2305.18617. [CrossRef]
145. Eaton, S.E. Teaching Fact-Checking Through Deliberate Errors: An Essential AI Literacy Skill. Available online: https://drsaraheaton.com/2025/04/23/teaching-fact-checking-through-deliberate-errors-an-essential-ai-literacy-skill/ (accessed on 10 March 2026).
146. Pitts, G.; Rani, N.; Mildort, W.; Cook, E.M. Students' Reliance on AI in Higher Education: Identifying Contributing Factors. In Proceedings of the HCI International 2025—Late Breaking Posters: 27th International Conference on Human-Computer Interaction (HCII 2025), Gothenburg, Sweden, 22–27 June 2025. [CrossRef]
147. Jin, Y.; Yan, L.; Echeverria, V.; Gašević, D.; Martinez-Maldonado, R. Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Comput. Educ. Artif. Intell.* **2025**, *8*, 100348. [CrossRef]
148. Soomro, K.A.; Kale, U.; Curtis, R.; Akcaoglu, M.; Bernstein, M. Digital divide among higher education faculty. *Int. J. Educ. Technol. High. Educ.* **2020**, *17*, 21. [CrossRef] [PubMed]
149. Walsh, B.; Dalton, B.; Forsyth, S.; Yeh, T. Literacy and STEM Teachers Adapt AI Ethics Curriculum. In Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23), Washington, DC, USA, 7–14 February 2023. [CrossRef]
150. Alam, A. Developing a Curriculum for Ethical and Responsible AI: A University Course on Safety, Fairness, Privacy, and Ethics to Prepare Next Generation of AI Professionals. In Proceedings of the Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2023), Tirunelveli, India, 16–17 February 2023. [CrossRef]
151. Deb, D.; Taylor, G.; Betz, S.; Maddux, B.A.T.; Ebert, C.E.; Richardson, F.W.; Couto, J.L.S.; Jarrett, M.S.; Madjd-Sadjadi, Z. Enhancing University Curricula with Integrated AI Ethics Education: A Comprehensive Approach. In Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE TS 2025), Pittsburgh, PA, USA, 26 February–1 March 2025. [CrossRef]
152. Howley, I.; Mir, D.; Peck, E. Integrating AI ethics across the computing curriculum. In *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*, 1st ed.; Holmes, W., Porayska-Pomsta, K., Eds.; Routledge: New York, NY, USA, 2022; pp. 255–270. [CrossRef]
153. Sposato, M. Artificial intelligence in educational leadership: A comprehensive taxonomy and future directions. *Int. J. Educ. Technol. High. Educ.* **2025**, *22*, 20. [CrossRef]