**Nusantara Journal of Artificial Intelligence and Information Systems**
Vol. 2, No. 1, June 2026, pp 1-16 journal homepage: https://journal.unusia.ac.id/nuai
**Review Article** ISSN: 3090-8302 DOI: 10.47776/nuai.v2i1.1993

# Mapping Generative AI Ethics in Higher Education: A Systematic Review and Multidimensional Boundary Framework

**Nio Awandha Nehru[1*], Nik Haryanti[2]**
[1]Postgraduate Programs, Universitas Islam Negeri Sayyid Ali Rahmatullah, Tulungagung, Indonesia
[2]Communication Science Department, Universitas Islam Balitar, Blitar, Indonesia

[1]nioawandha25@gmail.com and [2]nikharyanti1983@gmail.com

**INFORMATION**

***Article history:***

Received February 18, 2026
Revised March 26, 2026
Accepted April 19, 2026

***Keywords:***

Artificial intelligence
Higher education
Human-AI Interaction
Cognitive augmentation
Systematic literature review

**ABSTRACT**

The rapid integration of generative artificial intelligence in higher education has generated a fundamental dilemma: does it function as a learning aid or as an academic shortcut? This study employs a qualitative systematic literature review guided by PRISMA procedures to map the ethical spectrum of generative AI use between cognitive augmentation and academic substitution. Thirty-eight core journal articles published between 2021 and 2026 were thematically synthesized. The findings indicate that generative AI can enhance learning when it supports metacognitive reflection, scaffolding, and self-regulated learning while preserving human evaluative control. Conversely, risks emerge when AI contribution replaces core cognitive labor, leading to authorship ambiguity, integrity violations, and superficial engagement. To reconcile these tensions, this study proposes a multidimensional ethical boundary framework structured along two dimensions: human cognitive engagement and level of AI contribution. This framework offers a conceptual basis for policy development, assessment redesign, and responsible pedagogical integration, positioning ethical AI use as a continuum grounded in sustained human intellectual accountability.

***To cite this article:***

Nehru, N. A., & Haryanti, N. (2026). Mapping Generative AI Ethics in Higher Education: A Systematic Review and Multidimensional Boundary Framework. *Nusantara Journal of Artificial Intelligence and Information Systems*, 2(1), 1-16.

***Corresponding Author:***

Nio Awandha Nehru
Universitas Islam Negeri Sayyid Ali Rahmatullah
Jl. Mayor Sujadi Timur No. 46, Kabupaten Tulungagung, Jawa Timur, Indonesia (66221)
Email: nioawandha25@gmail.com

## 1. INTRODUCTION

The rapid advancement of generative artificial intelligence (AI) has significantly reshaped the landscape of higher education. Tools such as ChatGPT, Claude, and Gemini enable students to generate academic texts, summarize scholarly articles, paraphrase sources, and construct essays within seconds. This technological development reflects not only digital innovation but also an epistemological shift in how knowledge is accessed, constructed, and evaluated within universities. Fundamentally, artificial intelligence is designed to simulate human cognitive processes through computer-based systems, allowing more adaptive and flexible learning environments. When integrated responsibly into pedagogy, large language models (LLMs) can enhance learning efficiency, provide adaptive feedback, and support self-regulated learning [1], [2]. However, systematic multidisciplinary studies highlight that generative AI also introduces challenges related to academic integrity, authorship, assessment practices, and instructional design [3], [4].

In the Indonesian context, policy responses to generative AI in higher education have developed rapidly following the publication of the *Buku Panduan Penggunaan Generative AI pada Pendidikan Tinggi* issued by

the *Direktorat Jenderal Pendidikan Tinggi*. The guideline emphasizes ethical use, transparency, and the strengthening of digital and AI literacy competencies in universities, positioning AI not merely as a disruptive threat but as a transformative educational instrument requiring adaptive governance. This national stance is consistent with broader international policy frameworks, including guidance from UNESCO, which highlights the importance of human-centered AI governance in education [5]. Despite this policy alignment, stakeholders demonstrate differing perceptions of AI's impact: lecturers tend to express caution regarding academic integrity, whereas administrative staff and students more often view it as a tool that enhances learning efficiency. These contrasting perspectives indicate the need for a balanced integration strategy, recognizing that generative AI may challenge traditional assessment and instructional practices while simultaneously offering opportunities for innovative teaching approaches, particularly within digital learning environments such as Open and Distance e-Learning (ODeL) systems [6].

Despite these promising developments, the integration of generative AI into academic practice raises a central dilemma, whether AI primarily functions as a learning aid or as an academic shortcut. On one hand, AI systems may operate as cognitive assistants, offering scaffolding in brainstorming, language refinement, drafting, and revision. Research suggests that when learners maintain cognitive control, AI can augment metacognitive awareness and support knowledge construction [1], [7]. On the other hand, concerns have emerged regarding overreliance on AI-generated outputs, potential erosion of critical thinking, and blurred authorship boundaries. Empirical and conceptual analyses warn that unregulated AI use may encourage superficial engagement and challenge established norms of academic integrity [8], [9]. This tension reflects the absence of clearly articulated ethical boundaries distinguishing legitimate academic assistance from problematic shortcut practices in higher education.

Ethical considerations further complicate this debate. Issues such as concealed plagiarism, AI-assisted ghostwriting, ambiguous attribution practices, and disparities in AI literacy among students and faculty have become increasingly visible in recent studies [8], [10]. International policy frameworks and global governance reports consistently call for clearer regulatory standards, explicit disclosure of AI-generated contributions, and strengthened institutional oversight mechanisms to ensure responsible implementation [11]. However, institutional responses vary significantly across universities, leading to inconsistent interpretations of acceptable AI practices. While some institutions emphasize prohibition, others advocate guided integration supported by AI literacy training and revised assessment design [12], [13]. AI systems in education depend on extensive collection and processing of sensitive student data, raising significant concerns about privacy, surveillance, and potential misuse in the absence of strong ethical and regulatory safeguards [14].

Existing studies above predominantly concentrate on either the pedagogical effectiveness of generative AI or the technical detection of AI-generated text. Comparatively limited attention has been devoted to synthesizing ethical, pedagogical, and human-AI interaction perspectives into a unified conceptual framework. This gap is particularly evident in the Indonesian higher education context, where regulatory developments have progressed more rapidly than theoretical consolidation. Addressing this limitation, the present study conducts a structured literature review to map the ethical spectrum of generative AI use, conceptualizing it along a continuum from learning augmentation to academic substitution. By integrating insights from educational technology research, academic integrity discourse, and human-AI interaction theory, this article proposes a conceptual model clarifying when generative AI operates as a legitimate learning aid and when it risks functioning as an academic shortcut.

The innovative contribution of this research lies in reframing the debate beyond binary acceptance or prohibition. Rather than treating generative AI as inherently beneficial or inherently problematic, this study introduces a multidimensional ethical boundary framework grounded in degrees of human cognitive engagement and AI contribution. This integrative approach offers a theoretically informed foundation for policy development, institutional regulation, and pedagogical design in the evolving ecosystem of higher education.

## 2. METHOD

This study adopts a qualitative Systematic Literature Review (SLR) to ensure methodological rigor, transparency, and replicability in synthesizing scholarship on generative AI in higher education. Unlike narrative reviews, SLR applies a structured and protocol-driven process to reduce selection bias and enhance analytical reliability [15]. The procedure follows PRISMA-based stages: identification, screening, eligibility, and inclusion to systematically map the ethical spectrum of generative AI use along the continuum between cognitive augmentation and academic substitution [16]. This design enables a comprehensive and theoretically grounded synthesis of emerging debates in educational technology and academic integrity research.

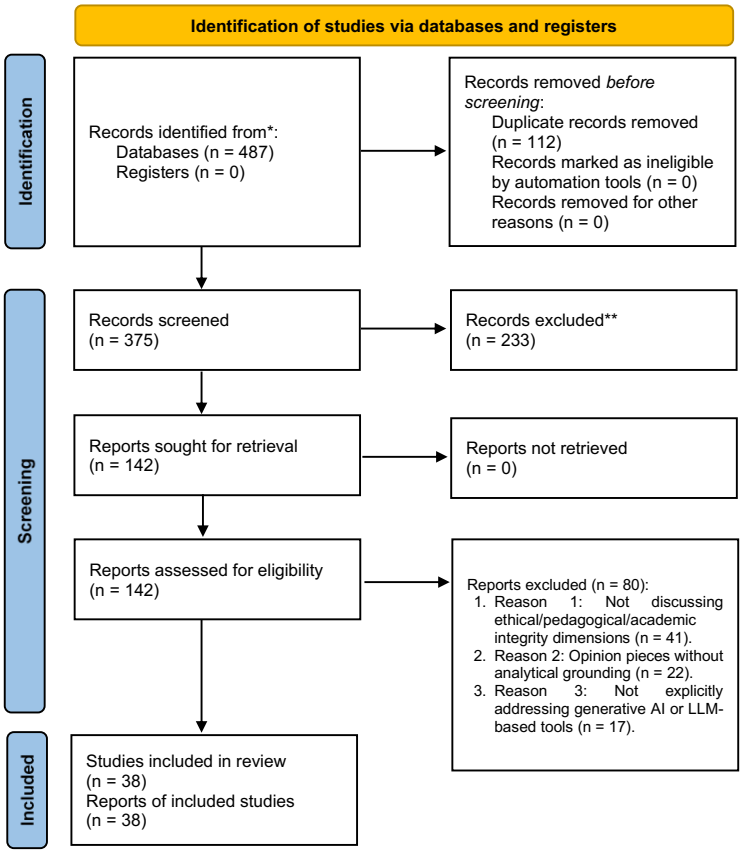


Figure 1. The PRISMA 2020 flow diagram of this research [17]

The systematic literature search was conducted across Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar to ensure multidisciplinary coverage. Keyword combinations included terms such as "generative AI in higher education," "AI and academic integrity," "ethical issues of AI in education," "AI-assisted writing," and "human–AI interaction in learning". The timeframe was limited to 2021–2026 to capture the rapid development of large language model discourse, and only peer-reviewed journal articles were included to maintain scholarly credibility. The PRISMA flow diagram summarizes the systematic selection process of this review. A total of 487 records were identified from five academic databases, and after removing 112 duplicates using Mendeley, 375 unique articles proceeded to title and abstract screening. At this stage, 233 records were excluded due to lack of relevance to generative AI in higher education or absence of ethical and pedagogical dimensions, leaving 142 articles for full-text assessment. Following eligibility evaluation, 80 articles were excluded for reasons including insufficient analytical grounding, lack of focus on generative AI or LLM-based tools, or absence of academic integrity discussion. Ultimately, 62 studies met the eligibility criteria, from which 38 core articles were selected for in-depth thematic synthesis, forming the analytical foundation of the Results and Discussion section.

Table 1. Inclusion and exclusion criteria

| Criteria Type | Description | Inclusion and Exclusion |
|---|---|---|
| Publication Type | Peer-reviewed journal and conference proceeding articles | Inclusion: Article published in indexed international journal.<br>Exclusion: Conference abstracts and opinion blogs/website. |
| Publication Year | Published between 2021–2026 | Inclusion: 2021 to 2026 empirical study on ChatGPT in higher education.<br>Exclusion: Under 2021 AI ethics paper. |
| Research Focus | Discusses generative AI or AI tools in academic/educational contexts | Inclusion: Study on AI-assisted academic writing.<br>Exclusion: AI in medical diagnosis without educational focus. |
| Ethical Dimension | Explicitly addresses ethics, academic integrity, cognitive engagement, governance, or boundary issues | Inclusion: Article on AI and plagiarism concerns.<br>Exclusion: Purely technical NLP performance study. |
| Educational Setting | Conducted in formal education settings (K–12, higher education, vocational, teacher education) | Inclusion: University-level AI usage study.<br>Exclusion: Corporate AI training context. |
| Language | Published in English | Inclusion: English-language journal article.<br>Exclusion: Non-English publication. |
| Accessibility | Full-text accessible for analysis | Inclusion: Open-access or institutionally accessible article.<br>Exclusion: Abstract-only record. |

To ensure consistency and relevance in the article selection process, predefined inclusion and exclusion criteria were applied. These criteria were aligned with the research objectives and guided the screening and eligibility stages of the review. The detailed criteria are presented in table 1. The selective coding process was conducted through iterative thematic refinement, in which axial categories were systematically reviewed and consolidated into higher-order conceptual themes. This process was performed manually by the authors to ensure interpretive depth and contextual sensitivity. Each article was repeatedly examined to identify recurring patterns, relationships, and conceptual overlaps, leading to the emergence of the core themes of "Cognitive Augmentation" and "Academic Substitution." Although qualitative data analysis software was not employed, the manual coding process followed rigorous qualitative analysis principles, including constant comparison and analytical memoing, to ensure reliability and conceptual coherence.

The proposed multidimensional ethical boundary framework advances existing AI ethics models by moving beyond binary classifications of AI use (e.g., acceptable vs. unacceptable) toward a continuum-based approach grounded in the interaction between human cognitive engagement and AI contribution. Unlike many prior frameworks that focus primarily on principles such as transparency, fairness, or accountability, this model introduces a pedagogically grounded lens that explicitly links ethical evaluation with learning processes and assessment design. Conceptually, the framework is both descriptive and normative. It is descriptive in mapping how AI is currently used across varying levels of cognitive engagement and contribution, and normative in proposing boundaries that distinguish ethically aligned augmentation from problematic substitution.

## 3. RESULTS AND DISCUSSION

### 3.1. Findings in the reviewed studies

A temporal analysis of the reviewed literature reveals a significant shift in research focus over time. Early studies in 2023 predominantly emphasized concerns related to academic dishonesty, plagiarism, and the potential misuse of AI as a shortcut mechanism. However, more recent contributions (2025–2026) demonstrate a transition toward exploring responsible pedagogical integration, ethical co-creation, and assessment redesign. This evolution indicates that the discourse is moving from reactive skepticism toward constructive frameworks that seek to align AI use with educational values and learning objectives. To systematically synthesize the findings of the included studies, a multi-stage coding process was employed, consisting of axial coding to identify recurrent analytical categories and selective coding to distill overarching themes. Table 2 presents the consolidated categories and corresponding main themes derived from the reviewed literature.

Table 2. Categories and main themes from the included studies

| Study ID | Axial Coding (Categories) | Selective Coding (Main Themes) |
|---|---|---|
| S1, S2, S3, S4, S5, S6, S7, S9, S10, S11, S12 | Pedagogical Integration and Human–AI Collaboration. | Conceptualizing Generative AI as Cognitive Augmentation in Higher Education. |
| S8, S13, S14 | AI Literacy and Cognitive Development. | |
| S15, S16, S17, S18, S19, S20, S21, S22, S23 | Academic Misconduct and AI Misuse. | Generative AI as Academic Substitution and the Ethical Risks and Integrity Challenges in Higher Education. |
| S24, S25, S26, S27 | Ethical AI Use and Responsible Writing. | |
| S28, S30, S31, S32 | Governance, Policy, and Institutional Regulation. | Toward a Multidimensional Ethical Boundary Framework for Human–AI Academic Interaction. |
| S34, S35, S38 | Human Capability, Skill Erosion, and AI Dependency. | |
| S29, S33, S36, S37 | Critical and Systemic Perspectives on AI in Higher Education. | |

As shown in Table 2, the included studies converge around shared analytical concerns, despite disciplinary and methodological differences. The axial categories capture recurring focal points such as cognitive support, human–AI collaboration, academic integrity risks, regulatory ambiguity, student dependency, and governance structures. These categories are subsequently integrated into broader selective themes that reflect the evolving positioning of generative AI within higher education—from augmentation to substitution, and ultimately toward the formulation of structured ethical boundaries. This thematic consolidation demonstrates that the discourse is progressively shifting from exploratory enthusiasm toward critical governance-oriented frameworks.

Beyond thematic categorization, the reviewed studies were also analyzed in terms of publication trends to understand the temporal development of the discourse. Table 2 summarizes the distribution of studies by publication year and journal outlet. This distributional overview provides insight into how scholarly attention to generative AI in higher education has intensified and diversified across publication venues.

Table 3. Distribution of reviewed studies by publication year and journal name

| Publication Year | Journal Name | Study ID |
|---|---|---|
| 2021 | Hawaii International Conference on System Sciences | S1 |
| | Proceedings of the ACM on Human-Computer Interaction | S2 |
| 2022 | CRC Press Taylor & Francis Group | S3 |
| | IGI Global Scientific Publishing | S4 |
| | Computers and Education: Artificial Intelligence | S5 |
| 2023 | Journal of AI | S6 |
| | International Journal of Educational Technology in Higher Education | S7 |
| | Computers and Education: Artificial Intelligence | S8 |
| | Journal of Applied Learning & Teaching | S9 |
| | Journal of University Teaching & Learning Practice | S10 |
| | International Journal for Educational Integrity | S11 |
| | UNESCO Publishing | S12 |
| | Learning and Individual Differences | S13 |
| | Higher Education | S14 |
| | Australasian Journal of Educational Technology | S15 |
| | Library Hi Tech News | S16 |
| 2024 | Computers and Education: Artificial Intelligence | S17 |
| | Innovations in Education and Teaching International | S18 |
| | Intelligent Service Robotics | S19 |
| | Studies in Higher Education | S20 |
| | Smart Learning Environments | S21 |
| | Innovations in Education and Teaching International | S22 |
| | Journal of University Teaching and Learning Practice | S23 |
| 2025 | International Journal of Educational Technology in Higher Education | S24 |
| | Advancing Educational Innovation with Artificial Intelligence | S25 |
| | Jurnal Saintekom: Sains, Teknologi, Komputer dan Manajemen | S26 |
| | Pendas: Jurnal Ilmuah Pendidikan Dasar | S27 |
| | Computers and Education Open | S28 |
| | 6th International Congress of Education, Social and Cultural Studies with Futurist Education | S29 |
| | Jurnal Kajian Akuntansi | S30 |
| | Advances in Simulation | S31 |
| | Project Leadership and Society | S32 |
| | International Journal of Educational Research Open | S33 |
| | Assessment & Evaluation in Higher Education | S34 |
| | Pitfalls of AI Integration in Education | S35 |
| | Studies in Continuing Education | S36 |
| | Educational Research Review | S37 |
| 2026 | Riwayat: Educational Journal of History and Humanities | S38 |

The distribution presented in Table 3 indicates a clear surge in publications between 2021 and 2026, reflecting the rapid emergence of generative AI tools in academic contexts. The concentration of studies in educational technology, higher education, and interdisciplinary journals suggests that the discussion is not confined to a single field but has become a cross-disciplinary concern. This temporal clustering also signals that the field remains in a formative stage, characterized by rapid conceptual development and ongoing normative debate.

To further contextualize the scope of the reviewed literature, the studies were classified according to educational level. Figure 2 illustrates the distribution of research across primary, secondary, tertiary, and informal or cross-level contexts. This classification helps clarify the primary educational setting in which generative AI is being examined.

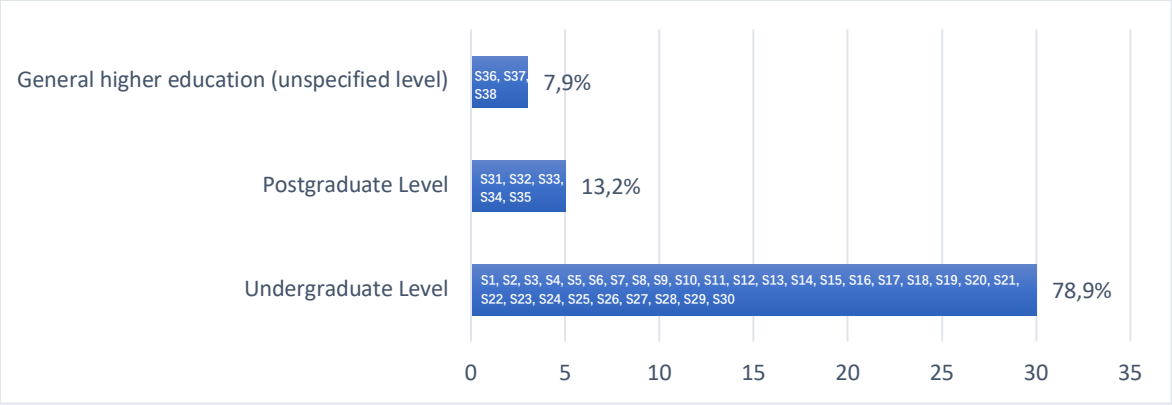


Figure 2. Educational level in the reviewed studies

As illustrated in Figure 2, the overwhelming majority of studies focus on higher education contexts. Only a limited number of works address school-level education or informal learning environments. This concentration confirms that universities function as the primary testing ground for generative AI integration, ethical negotiation, and governance experimentation. The dominance of tertiary-level studies also explains why issues such as academic integrity, research writing, and institutional regulation emerge as central concerns throughout the literature.

In addition to examining educational levels, the reviewed studies were also analyzed based on their geographical contribution. Figure 3 presents the distribution of contributing countries represented in the included studies. This mapping provides insight into the global spread of scholarly engagement with generative AI in higher education and highlights which national contexts are most actively shaping the discourse.

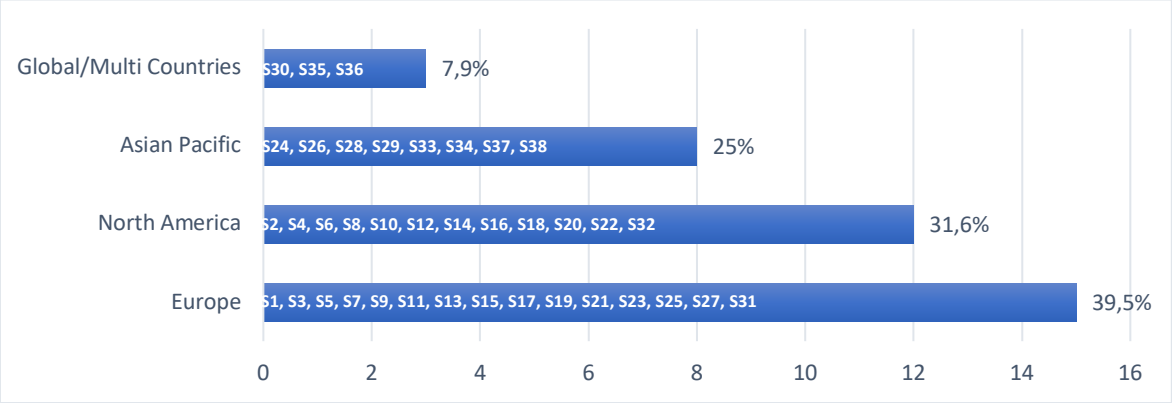


Figure 3. The contribution of each country in the studies

As shown in Figure 3, the contributions are concentrated in a limited number of countries, particularly those with strong research infrastructures and early adoption of generative AI technologies. Countries from North America, Europe, and parts of Asia Pacific dominate the scholarly output, while contributions from developing regions remain comparatively limited. This uneven distribution indicates that the current discourse is largely shaped by institutions with greater technological readiness and regulatory capacity. At the same time, the growing representation of diverse national contexts suggests that concerns related to governance, academic integrity, and human–AI interaction are becoming globally relevant rather than regionally confined. The geographical spread also underscores the need for context-sensitive policy development, as ethical boundaries and academic norms vary across educational systems.

In addition to educational level, the reviewed studies were analyzed based on their contextual focus. Figure 4 presents the distribution of research contexts, including classroom implementation, institutional policy, student perception, teacher preparedness, human–AI collaboration design, and governance frameworks. This contextual mapping allows for a more granular understanding of how generative AI is situated within academic practice.

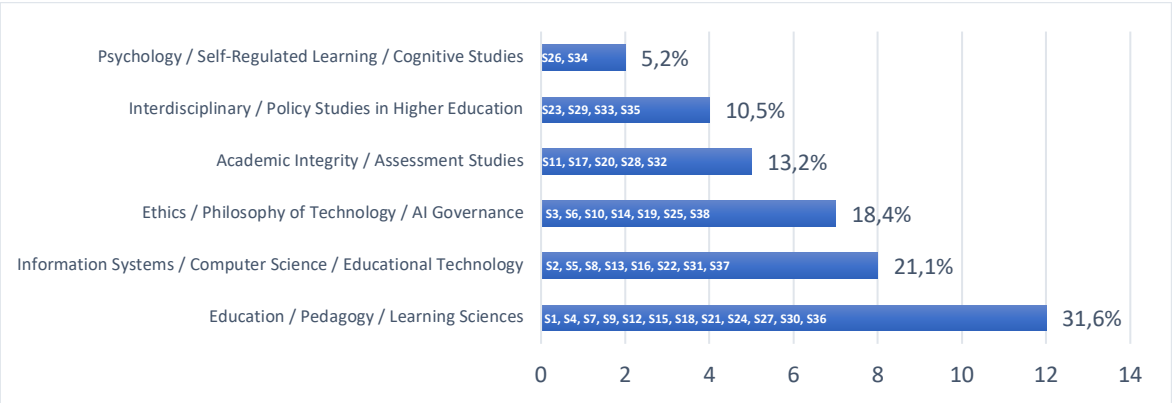


Figure 4. Context of the reviewed studies

Figure 4 demonstrates that the literature spans multiple layers of academic interaction, from micro-level classroom practices to macro-level governance and regulatory concerns. While early studies emphasized classroom experimentation and student perceptions, more recent contributions increasingly address institutional policy, ethical declarations, and governance mechanisms. This shift indicates a maturation of the

discourse, from tool-centered adoption toward systemic and multidimensional boundary-setting for responsible human–AI academic interaction.

### 3.2. Conceptualizing generative AI as cognitive augmentation in higher education

The integration of generative Artificial Intelligence (AI) in higher education is increasingly conceptualized as a form of cognitive augmentation rather than mere technological automation. This shift in perspective reflects a broader transformation in how digital technologies are understood within academic ecosystems, moving from efficiency-centered narratives toward cognition-centered interpretations. Emerging research argues that large language models (LLMs) function as epistemic mediators that reshape how learners organize information, construct arguments, and refine disciplinary discourse [18], [19]. Within this perspective, generative AI does not simply accelerate academic production; instead, it extends learners' analytical reach by offering dialogic interaction that stimulates iterative reasoning and conceptual reconsideration. Students are not merely consuming outputs but engaging in recursive exchanges that can provoke clarification, elaboration, and intellectual reorientation. Farrokhnia et al. suggest that generative AI tools can operate as intellectual partners when integrated within structured instructional contexts, enabling students to explore alternative interpretations, simulate counterarguments, and deepen conceptual clarity [20]. This framing situates AI within an augmentation paradigm, emphasizing its role as a cognitive amplifier embedded in pedagogical systems rather than as a shortcut detached from active engagement, thereby repositioning AI as a mediator of higher-order thinking rather than a substitute for it.

From a pedagogical standpoint, the augmentation model aligns with contemporary research on AI-enhanced instructional support systems that prioritize adaptive and learner-centered design. Luckin and Cukurova argue that AI can serve as an adaptive co-regulator of learning by providing real-time guidance tailored to learner progress, thereby strengthening instructional responsiveness beyond traditional static materials [21]. In generative AI environments, such scaffolding emerges through guided prompting, contextual elaboration, structured summarization, clarification of misconceptions, and iterative conversational refinement that evolves alongside student input. Rather than delivering fixed answers, the system participates in a dynamic exchange that encourages learners to reconsider assumptions and reorganize arguments. Empirical classroom-based investigations indicate that students who are guided to critically engage with AI-generated drafts demonstrate improved argumentative depth, stronger evidentiary support, and higher revision quality compared to those who use AI passively [22], [23]. These findings reinforce the view that AI-mediated scaffolding redistributes cognitive effort toward higher-order reasoning rather than diminishing intellectual rigor, particularly when pedagogical strategies explicitly require evaluation and justification of AI suggestions.

Recent studies also re-examine cognitive load dynamics within AI-supported learning environments, particularly in relation to how automation influences mental effort allocation. Kohnke et.al. report that generative AI can reduce extraneous load related to linguistic formulation, grammatical precision, and structural organization, thereby allowing learners to allocate more cognitive resources to synthesis, evaluation, and theoretical integration. This reduction of peripheral burdens can create cognitive space for deeper conceptual engagement, especially in complex academic writing tasks [24]. Similarly, Chiu emphasizes that when AI tools are used for preliminary drafting, brainstorming, or idea generation, they can enhance cognitive efficiency without compromising critical engagement, provided that learners retain evaluative control over final outputs [25]. In this sense, AI operates as a cognitive assistant that supports ideational exploration while preserving human authorship and interpretive authority. This redistribution mechanism exemplifies augmentation logic: routine processes are partially automated, while learners maintain responsibility for epistemic validation, coherence checking, and disciplinary alignment, ensuring that intellectual ownership remains human-centered.

Generative AI further intersects with contemporary models of self-regulated learning (SRL), which emphasize learners' proactive role in planning, monitoring, and evaluating their own progress. Lim et al. highlight that AI-driven feedback systems can support goal setting, progress monitoring, strategy adjustment, and performance reflection, core dimensions of SRL that are often difficult to sustain without timely feedback [26]. In higher education contexts, AI-generated formative feedback has been shown to encourage reflective iteration, particularly when students are instructed to document how AI suggestions were adopted, modified, or rejected [27]. This documentation requirement transforms AI use into a reflective exercise rather than a passive acceptance of generated content. By prompting learners to articulate reasons for accepting or revising suggestions, AI interaction fosters metacognitive awareness and strategic control. Rather than replacing reflective judgment, AI can act as a metacognitive trigger that stimulates deeper analysis, recursive revision cycles, and heightened awareness of one's cognitive processes, thereby reinforcing autonomous learning behaviors.

Metacognitive engagement represents a crucial dimension of augmentation because it determines whether AI enhances or undermines intellectual growth. Chan and Hu argue that generative AI tools can function as reflective mirrors, exposing inconsistencies, logical gaps, or underdeveloped claims in student writing that may

otherwise remain unnoticed [28]. When learners compare their original reasoning with AI-generated alternatives, they engage in evaluative comparison that strengthens epistemic awareness and critical discernment. This comparison process encourages students to question assumptions, reconsider evidence, and refine conceptual boundaries. Experimental studies in higher education settings reveal that students who are required to critique AI responses, identify inaccuracies, and justify revisions demonstrate greater analytical sophistication and argument coherence than those who merely incorporate AI-generated text [29]. This dialogic evaluation transforms AI into a cognitive interlocutor rather than an automated answer engine, positioning it as a stimulus for reflective reasoning rather than a replacement for intellectual effort.

Human-AI collaboration research further substantiates the augmentation paradigm by framing AI as a complementary partner in knowledge work. Dellermann et al. conceptualize collaborative intelligence as a process in which humans and AI systems jointly contribute to problem-solving, each leveraging complementary strengths such as pattern recognition, data processing speed, contextual understanding, and ethical reasoning [30]. In educational contexts, this means AI may assist with linguistic modeling, structural pattern detection, and rapid synthesis of information, while human learners retain contextual judgment, disciplinary interpretation, normative evaluation, and ethical accountability. Long and Magerko similarly emphasize that effective AI collaboration depends on maintaining human oversight, interpretive agency, and AI literacy to ensure responsible engagement [31]. These insights reinforce the notion that augmentation is contingent upon balanced interaction rather than unilateral automation, and that the quality of collaboration depends on learners' capacity to critically interpret and regulate AI contributions.

Nevertheless, augmentation remains conditional and context-dependent. Empirical analyses from 2024–2025 caution that without reflective integration and explicit instructional framing, generative AI can reduce productive cognitive struggle and promote surface-level engagement [32]. The determining factor lies not in the technology itself but in instructional design, assessment structure, and learner orientation. Assignments that incorporate reflective commentary, prompt disclosure, iterative revision logs, and comparative analysis between human and AI-generated reasoning preserve cognitive augmentation effects by ensuring active engagement. In contrast, unrestricted or unstructured use without metacognitive framing risks shifting from augmentation to substitution, where AI output replaces rather than stimulates intellectual effort. This conditionality highlights the importance of pedagogical intentionality in maintaining augmentation outcomes.

Across the reviewed studies categorized under cognitive augmentation (S1–S14), a consistent pattern emerges: generative AI contributes positively to higher education only when it operates within structures that preserve human evaluative authority and metacognitive engagement. Although the studies differ in methodological orientation, ranging from conceptual analyses to empirical classroom experiments, they converge in identifying scaffolding, reflective prompting, iterative drafting, and AI literacy as enabling conditions for augmentation. Importantly, none of the reviewed works frame augmentation as automatic; rather, its pedagogical value is conditional upon instructional design and learner agency. The synthesis therefore indicates that augmentation is not a technological property but a relational outcome shaped by how AI is embedded within assessment, feedback, and reflective practices. In conclusion, generative AI functions as cognitive augmentation when human learners retain epistemic control, actively interrogate AI outputs, and engage in iterative reasoning processes that deepen rather than displace intellectual labor.

For practical pedagogical application, augmentation-oriented use of generative AI can be operationalized through carefully designed assignments that explicitly position AI as a cognitive support tool rather than a substitute for student thinking. Drawing from the synthesis of the reviewed studies, particularly those emphasizing scaffolding, metacognitive engagement, and human–AI collaboration (S1–S14), AI can be integrated into early-stage learning processes to facilitate idea generation without diminishing intellectual responsibility. For example, students may be instructed to use AI tools to generate preliminary idea outlines, concept maps, guiding questions, or alternative argument structures. These outputs are not treated as final products but as initial cognitive stimuli that require further interrogation, validation, and refinement. Building on this approach, assignment designs can require students to critically evaluate AI-generated suggestions by identifying strengths, limitations, and potential inaccuracies before integrating them into their own work. This process may include comparative analysis between original student ideas and AI-generated alternatives, justification of selected arguments, and reflective commentary on how AI contributions influenced their thinking. Such structured interaction ensures that AI serves as a dialogic partner in the learning process rather than an autonomous producer of academic content. Empirical findings within the reviewed literature indicate that when students are guided to engage in iterative refinement and critical assessment of AI outputs, they demonstrate stronger argumentative coherence, deeper conceptual understanding, and enhanced metacognitive awareness.

Furthermore, augmentation-oriented practices can be strengthened by incorporating process-based assessment components, such as draft submissions, revision logs, or annotated outputs that explicitly document how AI-generated content was modified, accepted, or rejected. This aligns with research emphasizing that the educational value of AI lies not in the efficiency of output production, but in its capacity to stimulate reflective

reasoning and recursive knowledge construction. In this context, AI functions as a scaffolding mechanism that supports ideation, reduces cognitive barriers in the early stages of writing, and expands the range of perspectives considered by students, while ultimate epistemic control and authorship remain firmly grounded in human cognition.

### 3.3. Generative AI as academic substitution and the ethical risks and integrity challenges in higher education

Generative Artificial Intelligence (AI) has brought undeniable convenience to academic workflows, yet its rapid adoption in higher education has coincided with serious ethical concerns that risk undermining the very foundation of scholarly practice. Research indicates that the use of generative AI tools can blur the boundaries between legitimate support and academic misconduct, raising questions about plagiarism, authorship attribution, and the erosion of learner accountability [33]. A recent systematic review highlights how Generative AI (GenAI) technologies influence student behavior and academic honesty, emphasizing the complex interplay between the facilitation of task completion and potential violations of academic integrity standards across university contexts [34]. These concerns are particularly acute when students increasingly rely on AI-generated text in assignments without transparent disclosure or critical engagement with the underlying content, thus distancing the learning process from authentic intellectual effort. Over time, such dependency may gradually shift the locus of intellectual responsibility from the learner to the algorithm, weakening the formative dimension of academic writing as a process of reasoning, argument construction, and evidence-based justification. Empirical analyses published in 2023–2024 further suggest that the normalization of AI-generated drafts in coursework may reshape assessment cultures, prompting educators to reconsider how originality, process documentation, and authorship transparency are evaluated in digital learning environments [8], [27].

One of the most prominent ethical risks associated with AI-assisted academic work is the emergence of "AI-based plagiarism," sometimes termed "AI-giarism", where students incorporate AI-generated text into their submissions without proper attribution or reflective critical evaluation [35]. This phenomenon complicates existing definitions of plagiarism, as it differs from traditional copying of human-authored sources by involving outputs that are synthetically generated yet presented as original work. The literature reveals that existing academic misconduct frameworks have struggled to accommodate this new form of ethical breach, leaving institutions ill-equipped to distinguish between permissible assistance and academic dishonesty [36]. In this vein, questions of accountability arise: when a student's submitted work closely mirrors AI outputs, academic communities must decide whether such usage constitutes cheating, false authorship, or diminished scholarly involvement. Ref [37] argues that generative AI challenges long-standing assumptions about authorship and originality, requiring a reconceptualization of academic integrity that moves beyond binary distinctions between human and non-human production. The ambiguity surrounding authorship attribution also raises epistemological concerns, particularly regarding who can legitimately claim intellectual ownership when significant portions of reasoning, structure, or expression are algorithmically generated.

Student experiences with generative AI further illustrate the ethical ambiguity at play. Exploratory qualitative studies reveal that students often find it difficult to articulate clear moral or ethical boundaries regarding AI use in academic tasks, with many acknowledging that AI can save time but also expressing uncertainty about what constitutes acceptable or unacceptable use [38]. The absence of robust and unambiguous institutional guidelines exacerbates this moral hazard, as students often navigate a "shadow pedagogy" of informal practice norms that may conflict with formal integrity policies [39]. Such ambiguity can inadvertently promote superficial engagement with learning tasks, diminishing opportunities for deep cognitive processing and critical thinking. Moreover, when peer practices normalize undisclosed AI use, social pressures may further blur ethical judgment, reinforcing a culture in which efficiency is prioritized over intellectual authenticity and reflective scholarship. Watching in Indonesia university contexts reports that a significant proportion of students perceive AI-assisted writing as ethically "situational," depending on assessment type and perceived detection risk, which further complicates institutional enforcement and ethical clarity [40].

AI's potential to contribute inaccurate or fabricated information, commonly referred to as hallucinations, adds another layer of integrity risk to academic substitution. Although generative AI models can produce fluent and contextually plausible text, they occasionally generate false facts, incorrect citations, or misleading explanations [41]. When students accept such outputs uncritically, the risk of propagating misinformation within academic submissions increases, which in turn can undermine the credibility of educational programs and scholarly discourse. This challenge underscores the importance of ensuring that learners not only disclose AI involvement but also engage rigorously with the accuracy and validity of AI-assisted content. Without systematic verification and source triangulation, AI-generated inaccuracies may remain undetected, thereby compromising academic standards and weakening students' capacity to evaluate evidence independently. Zhai et al. further emphasize that overreliance on AI-generated explanations may reduce epistemic vigilance,

particularly among novice learners who lack sufficient domain knowledge to critically assess algorithmic outputs [29].

Institutional responses to these ethical challenges vary widely, often reflecting a lack of consensus on how to balance AI's benefits with integrity safeguards. Some universities have begun updating academic integrity policies to explicitly address AI use, advocating for transparency, documented draft histories, and clear acknowledgment of AI contributions [34]. However, other institutions have lagged in policy development, contributing to inconsistent enforcement and unclear expectations among students and faculty alike [42]. The gap between technological adoption and ethical regulation thus heightens the risk of academic substitution, where AI outputs effectively replace genuine intellectual effort rather than augment it. In the absence of coherent governance frameworks, disparities across departments or institutions may create confusion, leading to uneven accountability standards and potentially undermining trust within academic communities. International policy discussions, including UNESCO's 2023 guidance on generative AI in education, advocate for human-centered regulation, transparency requirements, and AI literacy initiatives to mitigate these risks while preserving academic values [5].

Beyond plagiarism and policy ambiguity, generative AI also raises concerns regarding equity and access. While some students benefit from advanced AI literacy and premium tools, others may lack the resources or skills to leverage these technologies effectively, potentially widening achievement gaps. Studies in digital ethics argue that unequal AI proficiency can distort assessment fairness, as students with greater technological competence may produce more polished outputs regardless of underlying conceptual mastery [1]. Consequently, academic substitution does not merely threaten integrity; it may also reshape competitive dynamics within higher education, requiring institutions to address issues of fairness, inclusivity, and digital competence alongside ethical compliance.

The cluster of studies grouped under academic substitution (S15–S27) reveals a shared concern regarding the erosion of authorship clarity and integrity norms when AI contribution exceeds human cognitive engagement. While individual studies focus on plagiarism, hallucination risks, regulatory ambiguity, or student misuse behavior, collectively they depict substitution as a structural risk emerging from disproportionate AI reliance combined with unclear institutional governance. The synthesis shows that ethical tension intensifies not merely due to AI usage but due to opacity, lack of disclosure, weak policy articulation, and diminished reflective accountability. Furthermore, cross-national evidence suggests that regulatory responses lag behind technological adoption, creating interpretive gaps that normalize shortcut practices. In summary, generative AI shifts from augmentation to substitution when algorithmic contribution replaces demonstrable human reasoning, resulting in compromised authorship transparency, weakened academic integrity, and heightened governance challenges.

In contrast, substitution-oriented practices emerge when students rely on generative AI to produce substantial portions or even the entirety of academic work with minimal modification, verification, or critical engagement. Based on the synthesis of the reviewed studies, particularly those grouped under academic substitution and integrity risks (S15–S27), this pattern reflects a shift in which AI no longer functions as a cognitive aid but instead assumes the role of primary knowledge producer. In such cases, the locus of intellectual responsibility is transferred from the student to the algorithm, resulting in diminished opportunities for reasoning, argument construction, and evidence-based justification. Empirical findings across the literature consistently indicate that when AI-generated outputs are adopted without scrutiny, students exhibit weaker conceptual understanding, reduced analytical depth, and limited capacity to defend their own arguments.

A common manifestation of substitution-oriented use can be observed in assignment practices where students prompt AI systems to generate complete essay drafts and submit them with only superficial edits, such as minor paraphrasing or formatting adjustments. This practice is often accompanied by the absence of reflective commentary, source verification, or acknowledgment of AI involvement, thereby obscuring authorship and undermining academic transparency. The reviewed studies highlight that such patterns not only challenge conventional definitions of plagiarism but also complicate the assessment of learning outcomes, as instructors are unable to distinguish between genuine student understanding and algorithmically generated reasoning. Furthermore, the risk is exacerbated by the increasing fluency and plausibility of AI-generated text, which can create an illusion of competence without corresponding cognitive engagement.

From a pedagogical perspective, this distinction underscores the necessity of designing assessment strategies that make human intellectual contribution both visible and measurable. Drawing on the synthesized evidence, effective countermeasures include requiring annotated drafts that demonstrate the evolution of ideas, revision logs that document iterative improvements, and structured reflections that explain how AI outputs were interpreted, validated, or rejected. These mechanisms re-anchor the learning process in human cognition by shifting evaluation criteria from final product to intellectual process. In addition, assignments can incorporate elements such as oral defense, in-class writing components, or personalized contextualization to further ensure the authenticity of student work.

### 3.4. Toward a multidimensional ethical boundary framework for human–AI academic interaction

The expanding integration of generative artificial intelligence in higher education has revealed the limitations of binary ethical framings that position AI either as a pedagogical enhancement or as a threat to academic integrity. Such dichotomous perspectives fail to capture the nuanced ways in which AI tools mediate learning processes, assessment practices, and epistemic responsibility within academic environments. Recent scholarship increasingly argues that ethical evaluation of human–AI academic interaction requires a multidimensional approach capable of capturing variations in learner agency, cognitive engagement, and degrees of algorithmic contribution. Rather than treating AI use as inherently acceptable or unethical, contemporary studies emphasize the need for boundary frameworks that situate AI practices along continua of human involvement and technological influence, thereby enabling more precise normative judgments and contextual sensitivity across disciplines [43], [44]. In this respect, ethical analysis must move beyond detection-oriented paradigms and instead interrogate how AI reshapes authorship, accountability, and the distribution of cognitive labor in higher education settings.

Central to this framework is the dimension of human cognitive engagement, defined as the extent to which learners actively exercise epistemic control, critical reasoning, and reflective judgment during AI-assisted academic work. This dimension underscores that meaningful learning occurs not merely through task completion but through sustained engagement in interpretation, evaluation, and knowledge construction. Research in higher education pedagogy demonstrates that learning quality is closely tied to students' engagement in sense-making processes, including evaluating sources, constructing arguments, and articulating original interpretations, regardless of technological mediation [45], [46]. In AI-supported contexts, ethical alignment is more likely when learners remain cognitively responsible for interpreting, validating, and contextualizing AI-generated outputs, rather than passively adopting them as ready-made solutions. Empirical study indicate that diminished cognitive engagement correlates with surface-level learning strategies and increased reliance on automated outputs, raising concerns about the erosion of academic authorship, intellectual accountability, and long-term epistemic development [47]. Thus, cognitive engagement functions not only as a pedagogical variable but also as a normative benchmark for ethical AI integration.

Complementing this dimension is the level of AI contribution, which refers to the scope and depth of AI involvement in producing academic artifacts. The literature distinguishes between low-level contributions, such as linguistic refinement, formatting support, or brainstorming prompts, and high-level contributions, including content generation, argument formulation, theoretical synthesis, or full draft production [8]. Ethical tension intensifies as AI contribution approaches substitution of core cognitive labor traditionally expected of students, particularly in summative assessment contexts where demonstration of independent reasoning is central. Several studies argue that ethical risk does not stem solely from AI usage per se, but from disproportionate AI contribution that eclipses demonstrable human intellectual effort, thereby challenging conventional notions of authorship, originality, and merit-based evaluation [48], [49]. In this sense, the level of AI contribution becomes a measurable indicator of whether AI operates as an augmentation or a substitution within academic practice.

Synthesizing these two dimensions enables the formulation of a conceptual ethical boundary model structured along intersecting axes of human cognitive engagement and AI contribution. Within this model, zones of ethical augmentation are characterized by high learner engagement coupled with limited or moderate AI support, where AI functions as a cognitive amplifier rather than a surrogate thinker. Conversely, zones of ethical concern emerge when high AI contribution coincides with low human engagement, signaling academic substitution, diminished agency, and potential integrity violations. This multidimensional mapping advances existing discourse by providing a theoretically grounded mechanism to differentiate acceptable AI-supported learning from ethically problematic dependency and by offering a scaffold for empirical operationalization in future research [27], [44]. Moreover, the model accommodates disciplinary variation by recognizing that acceptable thresholds of AI contribution may differ across fields, while still preserving the centrality of human intellectual accountability.

Beyond conceptual clarification, the framework carries important policy implications for higher education governance. Rather than relying on prohibitive or overly permissive AI regulations, institutions can operationalize ethical boundaries by articulating expectations around demonstrable human engagement and transparent AI contribution. Policy instruments such as reflective disclosures, process-based assessment, staged submissions, and documentation of human revision cycles align with research suggesting that integrity is better preserved through accountability mechanisms and formative design than through technological surveillance alone [50]. Furthermore, emerging scholarship emphasizes that AI governance should be educational rather than punitive, integrating AI literacy initiatives that cultivate students' understanding of responsible use, limitations of generative models, and implications for academic credibility [51]. Such policy alignment strengthens institutional coherence while reducing ambiguity around acceptable AI practices.

In the Indonesian context, the findings suggest several areas where policy clarification and strengthening are urgently needed to ensure responsible integration of generative AI in higher education. First, explicit

standards regarding acceptable levels of AI contribution in academic work remain underdeveloped, leading to inconsistent interpretations across institutions and disciplines. Second, existing institutional guidelines often lack clear and operational definitions of authorship in AI-assisted writing, particularly in distinguishing between legitimate support and unacceptable substitution. Third, there is still limited integration of AI literacy into formal curricula, which contributes to disparities in students' capacity to use AI tools critically, transparently, and ethically. In addition, the current regulatory landscape tends to emphasize general ethical principles without providing concrete implementation mechanisms at the classroom level, such as assessment design standards or documentation requirements for AI use. This gap creates ambiguity not only for students but also for lecturers in enforcing academic integrity consistently. Therefore, future policy development should focus on establishing clear disclosure requirements, defining measurable thresholds of acceptable AI involvement, and embedding AI ethics and literacy education systematically within higher education systems. Strengthening alignment between national policy, institutional regulations, and pedagogical practice will be essential to ensure that AI integration supports learning enhancement while preserving academic accountability.

From a pedagogical design perspective, the framework underscores the necessity of aligning AI integration with learning outcomes that prioritize reasoning, evaluation, synthesis, and metacognitive reflection. Instructional designs that require students to critique AI outputs, compare human and AI-generated reasoning, justify final decisions, or explicitly document iterative revision processes have been shown to strengthen metacognitive awareness and ethical sensitivity [52]. By embedding AI within tasks that demand intellectual agency and reflective articulation, educators can transform potential ethical risk into an opportunity for cultivating responsible and critically aware human–AI collaboration. In doing so, pedagogy shifts from reactive restriction to proactive design, ensuring that AI serves as a scaffold for higher-order thinking rather than a replacement for it.

The final thematic grouping (S28–S38) reflects a maturation of the discourse, moving from problem identification toward conceptual integration and governance modeling. Across these studies, there is growing consensus that binary prohibition-versus-acceptance debates are analytically insufficient. Instead, researchers increasingly advocate multidimensional approaches that incorporate learner agency, AI literacy, contribution thresholds, and institutional accountability mechanisms. The synthesis demonstrates that ethical boundary-setting becomes most coherent when framed along intersecting dimensions of human cognitive engagement and AI contribution, rather than isolated misconduct detection. This integrated pattern across studies supports the construction of a continuum-based framework capable of differentiating responsible collaboration from problematic dependency. In conclusion, a multidimensional ethical boundary model offers the most theoretically consistent and practically adaptable pathway for aligning generative AI integration with academic values, ensuring that technological advancement remains anchored in sustained human intellectual responsibility.

### 3.5. Multidimensional ethical boundary framework

To enhance conceptual clarity, the proposed multidimensional ethical boundary framework is visually represented as a two-dimensional matrix integrating human cognitive engagement (vertical axis) and level of AI contribution (horizontal axis). This visualization enables a more intuitive interpretation of ethical positioning, illustrating how varying degrees of human–AI interaction produce distinct pedagogical and ethical outcomes. The matrix representation further allows educators and policymakers to identify threshold zones between acceptable augmentation and problematic substitution, thereby operationalizing the framework beyond abstract conceptualization.

As illustrated in Figure 5 ethical positioning is determined by the interaction between two principal dimensions: Human Cognitive Engagement and Level of AI Contribution. The upper-left quadrant (Ethical Augmentation Zone) reflects conditions where AI functions as a cognitive amplifier while human epistemic control remains dominant. The upper-right quadrant (Guided Human–AI Collaboration) represents a transitional space in which substantial AI support is ethically acceptable provided that active human evaluation, revision, and accountability are maintained. The lower-left quadrant (Minimal Engagement Zone) indicates limited intellectual depth despite low AI reliance, whereas the lower-right quadrant (Academic Substitution Risk Zone) signals heightened integrity concerns due to diminished human reasoning and excessive AI-generated content. The diagonal Ethical Alignment Gradient further emphasizes that ethical robustness increases as demonstrable human intellectual engagement remains proportionally stronger than AI substitution.

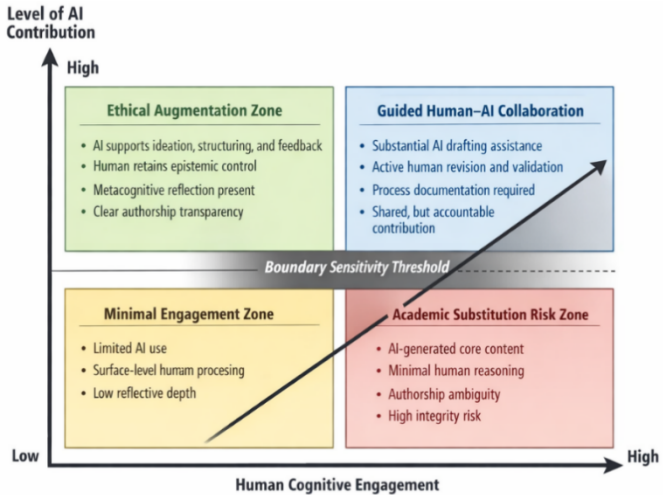


Figure 5. Multidimensional ethical boundary framework

To operationalize the proposed framework in instructional practice, Table 4 provides practical guidelines for lecturers to design AI-resilient assessments that balance ethical AI use with the preservation of academic integrity.

Table 4. Practical guidelines for designing ai-resilient assessments in higher education

| Principle | Implementation Strategy | Expected Outcome |
|---|---|---|
| Process Transparency | Require draft submission and revision logs. | Ensures visibility of student thinking. |
| AI Disclosure | Mandate explicit acknowledgement of AI use. | Promotes ethical accountability. |
| Critical Engagement | Include tasks requiring critique of AI outputs. | Enhances higher-order thinking. |
| Iterative Asessment | Use staged assignments (outline-draft-final). | Reduces reliance on AI-generated final products. |
| Personalization | Contextualize tasks with local or experiential data. | Limits generic AI responses. |
| Oral Validation | Conduct follow-up presentations. | Verifies authorship authenticity. |

These guidelines are derived from the synthesis of the reviewed literature, particularly studies emphasizing academic integrity, assessment redesign, and human–AI collaboration. The table outlines key principles, corresponding implementation strategies, and their expected pedagogical outcomes, highlighting how assessment design can simultaneously accommodate the ethical use of AI while safeguarding meaningful student engagement. By shifting the focus from product-oriented evaluation to process-oriented learning, these strategies aim to ensure that AI functions as a supportive tool rather than a substitute for human cognitive effort, thereby reinforcing transparency, accountability, and higher-order thinking within AI-integrated learning environments.

### 3.6. Limitations of the study

This study acknowledges several limitations that should be considered when interpreting the findings. First, the review was limited to English-language publications, which may exclude relevant studies published in other languages and introduce potential publication bias. As a result, important perspectives from non-English-speaking regions, including locally grounded policy discussions and pedagogical practices, may not be fully represented in the synthesis. Second, the dataset is heavily concentrated in higher education contexts, with limited representation from primary, secondary, or informal learning environments. This concentration reflects the current research trend but also restricts the generalizability of the findings across different educational levels. Ethical dynamics, patterns of AI use, and governance challenges may differ significantly in school-based or non-formal learning settings, which remain underexplored in the current literature. Third, the rapid evolution of generative AI technologies presents a temporal limitation. Given the review timeframe (2021–2026), newly emerging tools, features, and usage patterns may not yet be captured in the analyzed studies. This dynamic landscape means that conceptualizations of augmentation, substitution, and ethical boundaries may continue to evolve alongside technological advancements and shifting institutional responses. Finally, the existing body of literature is still largely dominated by conceptual analyses, perception-based studies, and cross-sectional designs, with relatively limited longitudinal empirical evidence examining the sustained impact of AI on learning outcomes, cognitive development, and academic integrity over time. This limits the ability to draw conclusions about long-term effects and the stability of observed patterns.

Taken together, these limitations highlight the need for future research that incorporates more diverse educational contexts, including primary, secondary, vocational, and informal learning environments, to ensure

broader applicability of the findings. In addition, expanding the scope to include multilingual and regionally grounded scholarship would provide a more inclusive and globally representative understanding of how generative AI is integrated and governed across different cultural and institutional settings. Future studies are also encouraged to adopt longitudinal and mixed-method research designs in order to capture the dynamic and evolving nature of human-AI interaction in education. Such approaches would enable a deeper examination of long-term impacts on learning outcomes, cognitive development, and academic integrity practices. Furthermore, empirical validation of the proposed framework across different disciplines and institutional contexts would be valuable in assessing its robustness and adaptability. Collectively, these efforts would contribute to refining the multidimensional ethical boundary framework and strengthening its relevance as both a conceptual and practical tool in real-world educational environments.

## 4. CONCLUSION

This study set out to examine how generative AI is conceptualized, contested, and governed within higher education, and the findings demonstrate clear alignment between this objective and the synthesized results. The analysis confirms that the literature consistently frames generative AI across three interconnected trajectories: as cognitive augmentation that enhances learning and human-AI collaboration, as academic substitution that raises integrity and dependency risks, and as a catalyst for developing multidimensional ethical boundary frameworks at institutional and policy levels. The distributional findings further show that the discourse is concentrated in higher education contexts and led by technologically advanced countries, indicating both rapid adoption and uneven global participation. Collectively, these outcomes substantiate the expectation introduced at the outset that generative AI in academia is evolving from an experimental pedagogical tool toward a structurally embedded governance issue. Future research should move beyond conceptual and perception-based studies toward longitudinal, cross-cultural, and policy-implementation analyses to test sustainable boundary frameworks in practice. The findings also offer practical prospects for universities to design context-sensitive governance models, assessment reforms, and AI literacy initiatives that balance cognitive augmentation with ethical accountability in human-AI academic interaction.

## CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this paper.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

[1] E. Kasneci et al., “ChatGPT for good? On opportunities and challenges of large language models for education,” Learn. Individ. Differ., vol. 103, p. 102274, 2023.
[2] I. Hamsar, H. Dewantara, H. Andini, S. Ardiansyah, N. A. Afrizah, and M. G. P. Hasta, “Analisis literasi artificial intelligence mahasiswa pada perguruan tinggi,” J. Vocat. Informatics Comput. Educ., vol. 2, no. 1, pp. 72–81, 2024.
[3] N. Kshetri et al., “‘So what if ChatGPT wrote it?’ Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy,” Int. J. Inf. Manage., vol. 71, p. 102642, 2023.
[4] A. Tlili et al., “What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education,” Smart Learn. Environ., vol. 10, no. 1, p. 15, 2023.
[5] W. Holmes and F. Miao, Guidance for generative AI in education and research. Unesco Publishing, 2023.
[6] K. Sevnarayan and M.-A. Potter, “Generative Artificial Intelligence in distance education: Transformations, challenges, and impact on academic integrity and student voice,” J. Appl. Learn. Teach., vol. 7, no. 1, pp. 104–114, 2024.
[7] D. Boud and E. Molloy, “Feedback in higher and professional education,” Underst. it doing it well, p. 2013, 2013.
[8] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, “Chatting and cheating: Ensuring academic integrity in the era of ChatGPT,” Innov. Educ. Teach. Int., vol. 61, no. 2, pp. 228–239, 2024.
[9] T. Susnjak and T. R. McIntosh, “ChatGPT: The end of online exam integrity?,” Educ. Sci., vol. 14, no. 6, p. 656, 2024.
[10] T. Bretag, “Challenges in addressing plagiarism in education,” PLoS Med., vol. 10, no. 12, p. e1001574, 2013.
[11] E. Papagiannidis, P. Mikalef, and K. Conboy, “Responsible artificial intelligence governance: A review and research framework,” J. Strateg. Inf. Syst., vol. 34, no. 2, p. 101885, 2025.
[12] J. Tsao, “Trajectories of AI policy in higher education: Interpretations, discourses, and enactments of students and teachers,” Comput. Educ. Artif. Intell., p. 100496, 2025.
[13] L. J. Wiese and A. J. Magana, “A department’s syllabi review for LLM considerations prior to university-standard guidance,” in 2024 ASEE Annual Conference & Exposition, 2024.
[14] D. S. T. Prayogi, “Artificial Intelligence in Education: Opportunities and Challenges,” Kartika J. Stud. Keislam., vol. 5, no. 1, pp. 605–621, 2025.
[15] H. Snyder, “Literature review as a research methodology: An overview and guidelines,” J. Bus. Res., vol. 104, pp. 333–339, 2019.
[16] M. J. Page et al., “Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement,” J. Clin. Epidemiol., vol. 134, pp. 103–112, 2021.
[17] N. R. Haddaway, M. J. Page, C. C. Pritchard, and L. A. McGuinness, “PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis,” Campbell Syst.

Rev., vol. 18, no. 2, p. e1230, Jun. 2022, doi: https://doi.org/10.1002/cl2.1230.
[18] D. Baidoo-Anu and L. O. Ansah, "Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning," J. AI, vol. 7, no. 1, pp. 52–62, 2023.
[19] J. Rudolph, S. Tan, and S. Tan, "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?," J. Appl. Learn. Teach., vol. 6, no. 1, pp. 342–363, 2023.
[20] M. Farrokhnia, S. K. Banihashem, O. Noroozi, and A. Wals, "A SWOT analysis of ChatGPT: Implications for educational practice and research," Innov. Educ. Teach. Int., vol. 61, no. 3, pp. 460–474, 2024.
[21] R. Luckin, K. George, and M. Cukurova, AI for school teachers. CRC Press, 2022.
[22] A. Mills, M. Bali, and L. Eaton, "How do we respond to generative AI in education?: Open educational practices give us a framework for an ongoing process," J. Appl. Learn. Teach., vol. 6, no. 1, pp. 16–30, 2023.
[23] Y. Kim, D. Kim, J. Choi, J. Park, N. Oh, and D. Park, "A survey on integration of large language models with intelligent robots," Intell. Serv. Robot., vol. 17, no. 5, pp. 1091–1107, 2024.
[24] L. Kohnke, B. L. Moorhouse, and D. Zou, "Exploring generative artificial intelligence preparedness among university language instructors: A case study," Comput. Educ. Artif. Intell., vol. 5, p. 100156, 2023.
[25] T. K. F. Chiu, "Future research recommendations for transforming higher education with generative AI," Comput. Educ. Artif. Intell., vol. 6, p. 100197, 2024.
[26] T. Lim, S. Gottipati, and M. Cheong, "What students really think: unpacking AI ethics in educational assessments through a triadic framework," Int. J. Educ. Technol. High. Educ., vol. 22, no. 1, p. 56, 2025.
[27] M. Perkins, "Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond," J. Univ. Teach. Learn. Pract., vol. 20, no. 2, pp. 1–24, 2023.
[28] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," Int. J. Educ. Technol. High. Educ., vol. 20, no. 1, p. 43, 2023.
[29] C. Zhai, S. Wibowo, and L. Li, "The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review," Smart Learn. Environ., vol. 11, Jun. 2024, doi: 10.1186/s40561-024-00316-7.
[30] D. Dellermann, A. Calma, N. Lipusch, T. Weber, S. Weigel, and P. Ebel, "The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems," arXiv Prepr. arXiv2105.03354, 2021.
[31] D. Long, T. Blunt, and B. Magerko, "Co-designing AI literacy exhibits for informal learning spaces," Proc. ACM Human-Computer Interact., vol. 5, no. CSCW2, pp. 1–35, 2021.
[32] A. Nguyen, Y. Hong, B. Dang, and X. Huang, "Human-AI collaboration patterns in AI-assisted academic writing," Stud. High. Educ., vol. 49, no. 5, pp. 847–864, 2024.
[33] K. Bittle and O. El-Gayar, "Generative AI and academic integrity in higher education: A systematic review and research agenda," Information, vol. 16, no. 4, p. 296, 2025.
[34] R. Khalida, A. Rahmandri, S. A. M. Magren, and E. Nurmiati, "Etika teknologi informasi dalam dunia pendidikan: Tinjauan literatur atas penggunaan AI dan isu plagiarisme akademik," J. Saintekom Sains, Teknol. Komput. dan Manaj., vol. 15, no. 2, pp. 222–234, 2025.
[35] R. Salsabillah and G. Erlianti, "Fenomena AI-Giarism Mahasiswa dalam Aktivitas Akademik: Kajian Tantangan dan Strategi Literasi Digital," Riwayat Educ. J. Hist. Humanit., vol. 9, no. 1, pp. 128–135, 2026.
[36] I. N. Fitri, A. Thahira, M. Mahdum, and J. Copriady, "Etika Penggunaan AI dalam Menjaga Integritas Akademik pada Pembelajaran Bahasa Inggris: Sebuah Kajian Literatur Sistematis," Pendas J. Ilm. Pendidik. Dasar, vol. 10, no. 04, pp. 425–439, 2025.
[37] S. E. Eaton, "Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology," Int. J. Educ. Integr., vol. 19, no. 1, p. 23, 2023.
[38] N. Hadinejad, K. Sperling, and C. McGrath, "Generative AI chatbots in higher education: Student experiences and perceived ethical challenges," Comput. Educ. Open, vol. 9, p. 100311, 2025, doi: https://doi.org/10.1016/j.caeo.2025.100311.
[39] A. Mazaherian and E. Nourbakhsh, "Beyond the Hype: Critical Analysis of Student Motivations and Ethical Boundaries in Educational AI Use in Higher Education," arXiv Prepr. arXiv2511.11369, 2025.
[40] A. Pulungan, R. Ramadhan, F. Windiarti, M. Ariani, and A. Hasibuan, "Examining The Determinants of AI Misuse Among Students in Indonesia: An Analysis Through The Fraud Diamond Framework," J. Kaji. Akunt., vol. 9, pp. 136–161, Jun. 2025, doi: 10.33603/jka.v9i1.10442.
[41] A. Cheng, A. Calhoun, and G. Reedy, "Artificial intelligence-assisted academic writing: recommendations for ethical use," Adv. Simul., vol. 10, no. 1, p. 22, 2025.
[42] M. Smit, R. F. Wagner, and T. J. Bond-Barnard, "Ambiguous regulations for dealing with AI in higher education can lead to moral hazards among students," Proj. Leadersh. Soc., vol. 6, p. 100187, 2025, doi: https://doi.org/10.1016/j.plas.2025.100187.
[43] M. Bearman, J. Ryan, and R. Ajjawi, "Discourses of artificial intelligence in higher education: A critical literature review," High. Educ., vol. 86, no. 2, pp. 369–385, 2023.
[44] J. M. Lodge, K. Thompson, and L. Corrin, "Mapping out a research agenda for generative artificial intelligence in tertiary education," Australas. J. Educ. Technol., vol. 39, no. 1, pp. 1–8, 2023.
[45] U. L. Snis, L. Svensson, and R. Säljö, "Learning in and for a changing work life," Studies in Continuing Education, vol. 47, no. 3. Taylor & Francis, pp. 491–495, 2025.
[46] L. Markauskaite et al., "Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI?," Comput. Educ. Artif. Intell., vol. 3, p. 100056, 2022.
[47] J. P. Miranda *et al.*, "Erosion of Critical Academic Skills Due to AI Dependency Among Tertiary Students: A Path Analysis," 2025, pp. 25–48. doi: 10.4018/979-8-3373-0122-8.ch002.
[48] B. D. Lund and T. Wang, "Chatting about ChatGPT: how may AI and GPT impact academia and libraries?," Libr. hi tech news, vol. 40, no. 3, pp. 26–29, 2023.
[49] J. Rudolph, M. F. bin Mohamed Ismail, and S. Popenici, "Higher education's generative artificial intelligence paradox: The meaning of chatbot mania," J. Univ. Teach. Learn. Pract., vol. 21, no. 6, pp. 14–48, 2024.
[50] C. Gonsalves, "Addressing student non-compliance in AI use declarations: implications for academic integrity and assessment in higher education," Assess. Eval. High. Educ., vol. 50, pp. 1–15, Oct. 2025, doi: 10.1080/02602938.2024.2415654.
[51] O. P. Barus, A. N. Hidayanto, E. Y. Handri, D. I. Sensuse, and C. Yaiprasert, "Shaping generative AI governance in higher education: Insights from student perception," Int. J. Educ. Res. Open, vol. 8, p. 100452, 2025, doi: https://doi.org/10.1016/j.ijedro.2025.100452.
[52] X. Liu and B. Zhong, "Integrating generative Artificial Intelligence into student learning: A systematic review from a TPACK perspective," Educ. Res. Rev., vol. 49, p. 100741, 2025, doi: https://doi.org/10.1016/j.edurev.2025.100741.

## BIOGRAPHIES OF AUTHORS


Nio Awandha Nehru He is a Postgraduate student in the English Language Education (*Tadris Bahasa Inggris*) Program at the Postgraduate Department of UIN Sayyid Ali Rahmatullah Tulungagung. He obtained his Bachelor of Education degree in English Language Education from the Faculty of Teacher Training and Education, Universitas Islam Balitar. He is also an alumnus of the National Resilience Institute of the Republic of Indonesia/*Lembaga Ketahanan Nasional Republik Indonesia* (Lemhannas RI) having completed the Training of Trainers *Pemantapan Nilai-Nilai Kebangsaan* (ToT PPNK) Batch 222 in 2025. His research interests include linguistics, artificial intelligence in education, curriculum development, and Teaching English as a Foreign Language (TEFL). He can be contacted via email at nioawandha25@gmail.com

**Nik Haryanti** was born in Blitar on October 1, 1983. She completed her formal education at SDN Pojok 1 (1997), MTsN Kunir Wonodadi Blitar (1999), MA Ma'arif Bakung Udanawu Blitar (2002), and earned her bachelor's and master's degrees from STAIN Tulungagung in 2006 and 2012, respectively. She later obtained her doctoral degree from IAIN Tulungagung in 2020. Her non-formal education includes studies at several Islamic boarding schools, namely Mahaijatul Qurro Blitar, Al-Sakdiyah Blitar, and Sirojut Tholibin Tulungagung. Her academic interests focus on education, management, and economics. She has authored numerous scholarly works, including books on Islamic education, educational management, research methodology, digital learning strategies, organizational behavior, and related topics, alongside publications in national and international journals. She can be contacted via email at nikharyanti1983@gmail.com