CONSTRUCTING RESPONSIBLE AI LITERACY FRAMEWORK FOR HIGH-SCHOOL STUDENTS IN SOCIAL DISCOURSE WRITING INSTRUCTION

XÂY DỰNG KHUNG NĂNG LỰC SỬ DỤNG TRÍ TUỆ NHÂN TẠO CÓ TRÁCH NHIỆM CHO HỌC SINH TRUNG HỌC PHỔ THÔNG TRONG DẠY HỌC VIẾT VĂN NGHỊ LUẬN XÃ HỘI

Ho Tran Ngoc Oanh*, Le Thi Thanh Tinh, Nguyen Viet Thuan, Truong Thi Hien, Duong Le Ny, Nguyen Thi Ngoc Quyen

The University of Danang - University of Science and Education, Vietnam

*Corresponding author: htnoanh@ued.udn.vn

(Received: March 16, 2026; Revised: April 12, 2026; Accepted: July 14, 2026)

DOI: 10.31130/ud-jst.2026.24(8).154

Abstract - In the context of the digital transformation of education, the rapid growth of generative AI poses significant challenges to academic integrity and students' independent thinking. This paper systematizes the theoretical foundations of AI literacy, digital ethics, and academic integrity, thereby proposing a Responsible AI Literacy Framework for high school students in teaching social argumentative writing. The competency structure comprises four core components: Transparency; Limitation Control; Cross-checking and Verification; and Output Preservation and Criticism. To ensure practical and legal consistency, the framework is aligned with the Digital Competency Framework for Learners under Circular No. 02/2025/TT-BGDĐT. The findings provide a system of behavioral indicators and pedagogical orientations that help teachers integrate AI as a "collaborative partner" to support thinking development while safeguarding students' agency and humanistic values in the process of text production.

Key words - Artificial Intelligence (AI); Responsible AI; argumentative writing; academic integrity; digital competency

Tóm tắt - Trong bối cảnh chuyển đổi số giáo dục, sự bùng nổ của trí tuệ nhân tạo tạo sinh đang thách thức liêm chính học thuật và tư duy độc lập của người học. Bài báo hệ thống hóa cơ sở lí luận về năng lực sử dụng trí tuệ nhân tạo, đạo đức số và liêm chính học thuật, từ đó đề xuất Khung năng lực sử dụng AI có trách nhiệm cho học sinh trung học phổ thông trong dạy học viết văn nghị luận xã hội. Cấu trúc năng lực gồm bốn thành tố cốt lõi: Minh bạch; Kiểm soát giới hạn; Đối chiếu và kiểm chứng; Bảo toàn và phản biện đầu ra. Nhằm đảm bảo tính thực tiễn và pháp lí, khung này được đối chiếu tương thích với Khung năng lực số dành cho người học theo Thông tư 02/2025/TT-BGDĐT. Kết quả nghiên cứu cung cấp hệ thống chỉ báo hành vi và định hướng sư phạm giúp giáo viên tích hợp AI như một "người cộng sự" hỗ trợ phát triển tư duy, bảo vệ vai trò chủ thể và giá trị nhân văn của học sinh trong hoạt động tạo lập văn bản.

Từ khóa - Trí tuệ nhân tạo (AI); AI có trách nhiệm; viết văn nghị luận; liêm chính học thuật; năng lực số.

1. Introduction

In recent years, the rapid growth of artificial intelligence (AI), particularly generative artificial intelligence (Generative AI - GenAI), has created strategic turning points in global education systems. In Vietnam, the Government has issued several important policies to promote this process, notably Decision No. 749/QD-TTg approving the "National Digital Transformation Program to 2025, with an orientation to 2030" and Decision No. 131/QD-TTg on strengthening the application of information technology and digital transformation in education [1], [2]. These orientations lay the foundation for integrating digital technologies into schools, aiming to improve the quality of teaching and learning in the new era.

The emergence of large language models (LLMs), such as ChatGPT and Gemini, has opened up unprecedented opportunities while also posing serious challenges to academic integrity. GenAI's ability to produce human-like texts is gradually blurring the boundary between learners' intellectual products and machine-generated content. In the context of implementing the 2018 General Education Program, Literature - particularly the teaching of social argumentative writing - plays an important role in developing students' argumentative competence, critical thinking, and social responsibility. Students' overreliance on AI as a substitute for independent thinking not only affects learning outcomes but also undermines creativity and academic ethics.

To concretize digital competency standards, the Ministry of Education and Training issued Circular No. 02/2025/TT-BGDDT prescribing the Digital Competency Framework for Learners [3], [4]. However, for students to use AI as a "collaborative partner" that supports thinking rather than as a tool for academic dishonesty, the formation of "Responsible AI Literacy" has become an urgent requirement. This is not merely a technical skill but also a combination of technological awareness, digital ethics, and academic integrity.

However, current studies on AI in education mainly focus on digital competence in general or AI literacy at a broad level, without delving deeply into specific subject areas. In particular, in Literature teaching, few studies have proposed a specialized competency framework associated with the characteristics of social argumentative writing - an activity that requires a high level of critical thinking, personal experience, and social responsibility. In addition, there remains a lack of specific behavioral indicators that help teachers observe and assess students' use of AI.

Challenges related to transparency, the control of technological limitations, and the ability to critically evaluate machine-generated outputs require a scientific set of criteria to guide students throughout the writing process.

Based on this reality, the present study aims to construct and propose a Responsible AI Literacy Framework (R1–R4) to guide the teaching of social argumentative writing at the high school level. The study not only examines the compatibility of this competency framework with the competency domains specified in Circular No. 02/2025/TT-BGDDT but also proposes specific instructional orientations for each stage of the writing process. This study contributes to the field by proposing a clearly structured competency framework (R1–R4) that integrates AI literacy, digital ethics, and academic integrity, while also concretizing it into a system of behavioral indicators and pedagogical orientations suitable for the context of Literature teaching in Vietnam.

2. Approach and research methods

2.1. Approach

The study was conducted using an integrated approach, combining theory, policy, and a human-centered orientation to ensure the scientific validity, legal consistency, and applicability of the proposed competency framework.

Theoretical approach: This approach focuses on establishing the scientific foundation for developing responsible AI literacy in the teaching of social argumentative writing. By synthesizing and systematizing theories of digital literacy, the study clarifies foundational concepts such as AI literacy, digital ethics, and academic integrity in order to identify the core competency components required when students interact with AI during the writing process.

Educational policy approach: The proposed competency framework is directly compared with the Digital Competency Framework for Learners prescribed in Circular No. 02/2025/TT-BGDDT issued by the Ministry of Education and Training. This approach helps concretize national digital competency standards into specific ethical behaviors and skills in the subject of Literature, while also ensuring compatibility and feasibility when applied in the general education environment.

Human-centered approach: Based on UNESCO's orientation (2024), the study approaches technology application in a way that protects students' autonomy. Accordingly, AI is viewed as a tool that supports thinking, while humans retain the central role in making judgments, taking responsibility for content, and preserving personal identity in writing.

2.2. Research objectives and research questions

2.2.1. Objectives

The study aims to establish a competency framework that helps high school students use AI effectively without undermining critical thinking and academic integrity. Specifically, it seeks to:

- Systematize the theoretical foundations of the relationship among AI literacy, digital ethics, and academic integrity in the teaching of social argumentative writing.

- Construct the structure of responsible AI literacy (R1–R4), accompanied by a system of behavioral indicators that can be observed and assessed.

- Propose pedagogical orientations for teachers in guiding students to control the limitations of technology throughout the stages of the writing process.

2.2.2. Research questions

- What core components should the structure of responsible AI literacy in social argumentative writing include to ensure integrity and critical thinking?

- How can the digital competency domains in Circular No. 02/2025/TT-BGDDT be concretized into specific behavioral indicators for students when interacting with GenAI?

2.3. Research methods

The study employed a combination of the following methods to ensure the reliability of the findings.

Method of theoretical analysis and synthesis: The study collected and analyzed scientific literature, policy documents, and domestic and international studies related to AI in education, AI literacy, digital ethics, and academic integrity. The documents were selected based on the following criteria: (i) direct relevance to the research topic; (ii) academic or legal value; and (iii) publication within the past five to ten years. The data were analyzed using a thematic approach to establish foundational concepts and the basis for constructing the competency framework.

Modeling method: This method was used to transform abstract theoretical concepts of ethics and responsibility into a systematic competency structure model, including the following components: Transparency (R1), Limitation Control (R2), Cross-checking and Verification (R3), and Output Critical Thinking (R4).

Expert reference method: The competency framework was refined through comparison with international competency frameworks, particularly UNESCO (2024), and domestic regulations. This approach helps ensure the standardization, currency, and suitability of the behavioral indicators in relation to the psychological characteristics of high school students and the context of Literature teaching in Vietnam.

The research process was implemented through three main steps. First, the research team synthesized and analyzed relevant theoretical and policy documents to establish the conceptual foundation of AI literacy, digital ethics, and academic integrity. Second, based on these theoretical foundations, the R1–R4 framework was constructed through the modeling method, in which competency components were identified and structured into a logical system. Third, the competency framework was reviewed and refined by referencing international standards and comparing it with the Digital Competency Framework under Circular No. 02/2025/TT-BGDDT to ensure its relevance and feasibility in teaching practice.

3. Research content

3.1. The Relationship among AI Literacy, digital ethics, and academic integrity in Literature teaching

In the digital era, the integration of artificial intelligence into education is not merely an addition of tools but a restructuring of cognition and learning behavior. The research team establishes that responsible AI literacy is built upon the sustainable intersection of three components: AI Literacy (Competence), Digital Ethics (Values), and Academic Integrity (Behavior). This interconnection forms a "protective triangle," helping students maintain human agency in the face of the power of algorithms.

First, AI literacy serves as the competency foundation. Previously, digital literacy mainly focused on basic computer skills; however, AI literacy has now become a new measure of digital citizenship. This concept encompasses the ability to understand, use, and, in particular, critically evaluate AI systems. In Literature teaching, AI literacy helps students recognize "AI hallucinations" - the phenomenon in which LLMs generate unreliable information. Without this foundation, students may easily fall into a state of "cognitive dependence," unconditionally accepting machine-generated outputs, which leads to a decline in critical thinking and lifelong self-learning ability. For example, when AI is asked to provide evidence about a social phenomenon, the system may generate unverified statistics or events. Without evaluative competence, students may easily accept and directly use such information in their writing, resulting in content inaccuracies and reduced academic reliability.

Second, digital ethics serves as a value orientation. AI ethics not only regulates the behavior of developers but also places requirements on users regarding transparency and fairness. In the context of social argumentative writing - a genre that requires personal viewpoints and experiences - digital ethics functions as a bridge that transforms technological skills into humanistic behavior. It guides students to use AI as an assistant that opens up perspectives rather than as an entity that writes on their behalf. UNESCO's human-centered approach requires that the use of AI protect learners' autonomy, ensuring that technology does not obscure the writer's personal identity and social responsibility [5].

Third, academic integrity is the concrete behavioral manifestation. It is a vital measure in the digital education environment, expressed through a commitment to complying with standards of honesty and integrity [6]. Academic integrity in the AI era does not mean prohibiting technology; rather, it requires transparency in the interaction process. When students use AI to generate ideas or structure their writing, disclosing and acknowledging sources of reference constitute the practice of integrity. Conversely, presenting algorithm-generated products as one's own intellectual work is not only an act of cheating but also an abandonment of the opportunity to develop one's own argumentative thinking.

It can be affirmed that these three components have a close dialectical relationship. Without AI literacy, students lack the ability to identify risks in order to practice ethics; without digital ethics, AI literacy skills can easily be distorted into tools for cheating; and without a focus on academic integrity, Literature teaching will lose its core humanistic value. Therefore, in teaching social argumentative writing, the synchronous integration of these three elements is a necessary condition for developing responsible AI literacy, enabling students to enter the digital era with the mindset of those who master technology [1].

3.2. Proposing the responsible AI Literacy framework R1–R4

Based on UNESCO's orientation regarding the maintenance of "human agency" [5], the authors establish the R1–R4 competency framework. The R1–R4 framework is designed as a structural model consisting of components connected according to a cognitive process, enabling students to regulate their interaction with AI from awareness to control and critical engagement.

R1. Transparency - The foundation of academic ethics. This component does not stop at "disclosure" but concerns self-awareness of the origin of knowledge. In social argumentative writing, transparency requires students to clearly distinguish between "prompting ideas" from AI and their own "core arguments." Specifically recording the steps of use, or maintaining an AI log, helps prevent technological plagiarism - a subtle form of cheating in which students use AI to generate texts that appear "original" but are, in essence, algorithmic reproductions. Such transparency provides teachers with a "cognitive map" to accurately assess students' actual effort and competence, while also building an honest academic environment.

R2. Limitation Control - Preserving the role of agency. Limitation control is the ability to determine the scope of AI intervention in the structure of a piece of writing. The research team proposes that AI should only support "micro-tasks," such as searching for keywords, suggesting multidimensional approaches, or checking spelling errors. The most important manifestation of this component is that students must independently analyze the prompt and identify the discussion issue before consulting AI. If students leave these core thinking operations to machines, their writing will fall into a state of "formulaic composition," lacking the depth derived from real-life experiences and the writer's emotions - elements that constitute the value of a social argumentative essay.

R3. Cross-checking and Verification - The capacity to evaluate digital knowledge. Because LLMs operate based on statistical probability, which leads to the phenomenon of "AI hallucinations," cross-checking competence serves as a safety filter for writing. Students must develop the habit of scientific skepticism and verify all social evidence, including statistics, events, and figures, provided by AI through validated sources, such as textbooks, reputable journalism, and research documents [7]. This component not only helps ensure the authenticity of writing but also develops students' information management competence

in a digital environment filled with fake news and algorithmic bias.

R4. Output Critical Thinking - Higher-order critical thinking. This component represents the highest level of maturity in responsible AI literacy. At this level, students do not passively receive AI-generated products but engage in "dialogue with the machine." Students need to ask counter-questions: Is this AI-generated argument one-sided? Is this evidence appropriate to the Vietnamese cultural context? From there, students restructure, adjust, and personalize the content according to their own style and standpoint. Preserving control over outputs helps transform reference content into products of independent thinking, ensuring that the final product is the result of students' own innovation and critical thinking.

The establishment of the four-component R1–R4 system not only supports students in using AI effectively but also directly contributes to developing autonomy and self-learning competence - one of the core requirements of the 2018 General Education Program. This competency framework creates a solid barrier that protects the humanistic values of Literature against the tendency toward excessive technologization. The novelty of the R1–R4 framework lies in its simultaneous integration of three dimensions: technological competence (AI literacy), value orientation (digital ethics), and behavioral standards (academic integrity), while concretizing them into behavioral indicators that can be observed and assessed in the context of Literature teaching.

3.3. Behavioral indicator system and alignment with Circular No. 02/2025/TT-BGDDT

The R1–R4 competency framework does not operate in isolation but is developed based on the concretization of the digital competency domains prescribed in Circular No. 02/2025/TT-BGDDT of the Ministry of Education and Training [4]. This comparison not only ensures legal consistency but also demonstrates that responsible AI literacy is the in-depth application of national digital competence to the practice of Literature teaching.

From the table above, it can be seen that the alignment between the R1–R4 framework and Circular No. 02/2025 reflects a shift from "knowing how to use" to "knowing how to master."

First, in the domain of AI Application, while the Circular requires learners to understand AI ethics in general, R1, Transparency, concretizes this requirement into the acts of noting and explaining AI use in Literature writing. This transforms an abstract ethical principle into a daily habit of academic integrity.

Second, the Problem Solving domain in the Circular is concretized by the research team into R2, Limitation Control. In social argumentative writing, problem solving is not only about which tool to use but also about the extent to which it should be used so as not to weaken autonomous thinking. Students' identification of AI as merely a "micro-task assistant" reflects an advanced level of digital competence: the ability to regulate technology according to personal needs without being led by it.

Table 1. *Behavioral indicator system and alignment with the national digital competency framework*

R1–R4 Components	Specific Behavioral Indicators / Manifestations	Alignment with Competency Domains under Circular 02/2025
R1. Transparency	- Clearly note the purpose of AI use at the end of the essay. - Describe the steps of use and the role of AI when requested to provide an explanation. - Clearly distinguish between machine-suggested content and student-developed content.	Domain VI: AI Application. Emphasizes "understanding, using, and evaluating AI systems ethically and responsibly" [4, p. 5].
R2. Limitation Control	- Use AI only to support micro-tasks, such as idea generation and error checking. - Independently analyze the prompt and identify the discussion issue before consulting AI. - Add arguments and evidence beyond AI's suggestions.	Domain V: Problem Solving. Focuses on "the ability to think critically and creatively to identify needs and appropriate technological solutions" [4, p. 4].
R3. Cross-checking and Verification	- Verify social evidence through textbooks and official documents. - Detect and remove inaccurate or ambiguous information. - Clearly cite references for data obtained from AI.	Domain I: Data and Information Literacy. Requires learners to "analyze, compare, and evaluate the reliability and authenticity of digital content" [4, p. 2].
R4. Output Critical Thinking	- Identify the strengths and limitations of AI-generated arguments. - Restructure and rewrite paragraphs in a personal writing style. - Revise the text based on dialogue with AI's suggestions.	Domain III: Digital Content Creation. Emphasizes the ability to "refine and reconstruct content to create new and unique knowledge" [4, p. 3].

Third, R3, Cross-checking and Verification, is the realization of the Data and Information Literacy domain in response to the phenomenon of "AI hallucinations." In the context where high school students can easily be overwhelmed by the vast and fluent flow of information from chatbots, this indicator requires students to return to foundational sources of knowledge, such as textbooks and academic materials, thereby developing information evaluation skills - a vital component of digital competence in the age of fake news.

Finally, R4, Output Critical Thinking, is the highest expression of the Digital Content Creation domain. If the Circular emphasizes the creation of new content, then in Literature teaching, such "new content" must bear a strong personal imprint and critical thinking. Students do

not merely integrate information from AI; they must use it as a counterpoint to challenge their own arguments, thereby constructing texts with depth and individual character.

In short, the R1–R4 behavioral indicator system successfully transforms the national digital competency domains into a set of specific action standards for high school students. This alignment ensures that AI-supported teaching of social argumentative writing does not deviate from the trajectory of national educational innovation, while also creating a safe corridor for the development of learners' character and thinking in the digital environment.

3.4. Orientations for teaching social argumentative writing through an AI-supported process

Process-oriented writing emphasizes thinking operations and practical activities throughout the process of producing a written product [3]. When AI is integrated, this process does not change in structure but changes in the mode of interaction. The research team proposes orientations for teaching AI-supported social argumentative writing through three strategic stages, in which teachers act as "pedagogical referees" who coordinate the relationship between students and technology.

(1) Pre-writing stage: Stimulating thinking and establishing ideas. This is the stage in which students often encounter the greatest difficulty in determining approaches and collecting evidence for complex social issues. AI can function as an "argumentation support tool," providing multidimensional perspectives, suggesting counterarguments, and helping learners overcome the initial "bottlenecks" in idea formation [8], [9].

Students' actions: Students use prompts to ask AI to suggest outlines or opposing streams of opinion on the discussion issue.

Teachers' intervention: Teachers guide students to implement R2, Limitation Control, by requiring them to independently analyze the prompt and establish their personal stance before consulting AI. Teachers need to emphasize that AI-generated outlines are only raw materials; students themselves must select the main arguments in order to preserve their thinking agency.

(2) While-writing stage: Drafting and developing arguments. At this stage, students must transform outlines into complete texts. AI participates as a "technical assistant" to improve expression quality and coherence [10].

Students' actions: Students use tools such as QuillBot or Wordtune to rephrase less fluent passages, search for suitable vocabulary, or check grammar errors instantly. For example, students may ask AI to rephrase a paragraph and then compare it with the original version to choose an expression that fits their personal writing voice.

Teachers' intervention: Teachers supervise the process to ensure that students do not fall into a "copy-and-paste" pattern, which makes their writing "formulaic" and lacking in personal identity. Teachers need to encourage students to implement R4, Output Critical Thinking, by adjusting machine-suggested sentences so that they match the writer's authentic voice and emotions. Students must be the ones who directly construct paragraphs connected to real-life contexts and their own distinctive experiences.

(3) Post-writing stage: Feedback, revision, and completion. This is a key stage for developing self-learning competence and academic integrity. AI supports students in identifying logical errors in arguments or insufficiencies in evidence.

Students' actions: Students compare their writing with AI-generated feedback while also implementing R3, Cross-checking and Verification, to verify statistics or quotations suggested by AI through official sources [7].

Teachers' intervention: Teachers combine AI-generated results with multi-source assessment tools, such as rubrics, checklists, and self-assessment forms. Teachers guide students to use the self-assessment form to review the entire process: *Have I been transparent (R1) in my use of AI? Have I mastered my argument (R4), or have I been led by AI?*

Thus, teaching according to this orientation shifts the teacher's role from a provider of knowledge to a process advisor. Teachers do not only grade the final text but also assess how students "dialogue" with technology. AI does not make writing easier in a passive sense; in fact, it imposes higher requirements for critical competence: students must have sufficiently solid background knowledge to control and coordinate the tool. This orientation helps protect the core humanistic value of Literature, transforming AI into a lever that enables students to attain a higher level of thinking in digital academia.

3.5. Orientations for and construction of an assessment toolkit for responsible AI Literacy

To realize the assessment of responsible AI literacy, the study establishes the view that assessment should not focus solely on the final written product but should primarily assess the interaction process. The assessment toolkit is designed according to the principle of evidence-based assessment, in which behaviors related to AI use are transformed into observable and measurable indicators [5].

3.5.1. Operational orientations of the toolkit

The assessment of responsible AI literacy does not stop at checking outcomes; rather, it is a continuous interactive process that requires close coordination among data sources and assessment moments. The research team establishes the operational orientation of the toolkit based on the following pillars:

- *Multi-source and multi-stage assessment:* This competence is measured through a combination of teacher assessment and students' self-reflection, or self-assessment. This process is conducted throughout the pre-writing, while-writing, and post-writing stages in order to shift the assessment philosophy from "behavior control" to "supporting the development of reflective competence." This approach helps fully identify changes in learners'

awareness of digital ethics throughout the process of text production.

- *Ensuring objectivity through behavioral indicators:* To overcome assessors' subjective biases toward the application of technology in academic activities, the toolkit focuses on specific and observable behavioral indicators. Quantifying manifestations such as source transparency, control over the scope of support, and data authenticity helps make assessment results fair, scientific, and transparent.

- *Assessment for progress and humanistic values:* The study emphasizes process assessment instead of focusing only on the formal completeness of the final product. Thinking operations such as how students formulate prompts, filter information, and critically respond to machine-generated outputs are considered the most important evidence of competency maturity. This affirms the human-centered philosophy, in which AI serves as a lever for students to enhance independent academic thinking.

3.5.2. Integrated assessment toolkit

The study designed a synchronized toolkit to fully reflect the R1–R4 components.

Analytical rubric: This is the central tool used to classify students according to four proficiency levels: Proficient, Satisfactory, Needs Practice, and Not Yet Competent. The rubric focuses on qualitative criteria such as the level of transparency (R1), the ability to maintain thinking agency (R2), and personalization in argumentation (R4). Publicizing the rubric from the beginning helps students recognize their "learning map" and consciously adjust their AI-use behaviors.

Checklist for teachers: This is used as a classroom observation tool. Teachers record the presence or absence of specific responsible behaviors during lessons. For example: Do students verify evidence from AI? Do students independently analyze the prompt beforehand? The checklist helps supplement the rubric score with practical evidence, ensuring the authenticity of process assessment.

Self-assessment form for students: This tool plays an important role in developing self-reflection and digital ethics. Students examine their own level of proficiency through five-level survey questions. This is the stage in which students reconsider their own integrity: *Have I mastered AI, or am I being led by AI?*

3.5.3. Guidelines for coordinating tools in writing instruction

The coordinated use of the three types of tools - rubric, checklist, and self-assessment form - establishes a closed assessment ecosystem, in which data are continuously collected to accurately reflect students' growth in responsible AI literacy. This process not only measures the final product but also intervenes deeply in learners' ethical behavior through three stages.

(1) Pre-writing stage - Orientation and commitment building: Instead of serving only as a scoring scale, the analytical rubric is introduced by teachers from the outset as an "ethical guide." Students' early access to criteria concerning Transparency (R1) and Limitation Control (R2) helps them proactively establish the boundary between consulting tools and thinking independently. This is an important step in transforming principles of academic integrity from theory into concrete commitments to action.

(2) While-writing stage - Process monitoring and timely adjustment: While students interact with GenAI, teachers use the checklist to record actual behaviors. Close monitoring of how students formulate prompts or cross-check data (R3) allows teachers to detect early signs of technological dependence or information inaccuracy. Immediate feedback based on the checklist helps students adjust their actions, ensuring that technology always serves the goal of developing thinking rather than replacing it.

(3) *Post-writing stage - Synthesis and value reflection:* At this stage, assessment reaches the intersection between teacher assessment and student self-assessment. Teachers use the rubric to evaluate the degree of output criticism (R4) in the writing, while students complete the self-assessment form. The comparison between teachers' judgments and students' self-reflective perspectives helps students recognize competency gaps, thereby forming an awareness of self-regulation in subsequent interactions with technology.

Establishing this orientation for coordinating the toolkit ensures that artificial intelligence is positioned properly as a "means of promoting pedagogical innovation," while strictly protecting the central role of humans and the value of academic integrity [1]. Consistency in the assessment process not only supports students in improving their social argumentative writing skills but, more importantly, helps them gradually develop the capability and ethics of responsible digital citizens in the future.

4. Conclusion

The study on constructing a responsible artificial intelligence literacy framework for high school students in teaching social argumentative writing demonstrates the urgent need to establish an ethical and competency-based frame of reference in the context of educational digital transformation. Based on the systematization of the relationship among AI literacy, digital ethics, and academic integrity, the study proposed a four-component competency framework (R1–R4): Transparency, Limitation Control, Cross-checking and Verification, and Output Criticism. The novelty of the study lies in integrating these three dimensions into a unified competency structure, while concretizing them into a system of behavioral indicators that can be observed and assessed in the context of Literature teaching.

The research findings affirm that, in Literature teaching, AI should be positioned as a cognitive tool rather than as an entity that replaces learners' thinking. The implementation of the R1–R4 framework not only helps students use technology effectively but also contributes to

protecting the central role of critical thinking and personal identity in text production. This is particularly meaningful in realizing the requirements of the 2018 General Education Program and the Digital Competency Framework for Learners under Circular No. 02/2025/TT-BGDDT.

For this competency framework to be effective in practice, the Ministry of Education and Training needs to soon issue specific guidelines on academic integrity in the AI era. High schools need to organize in-depth digital competency training for Literature teachers, focusing on skills for managing the AI-supported writing process. At the same time, it is necessary to develop digital learning resource repositories and academic ethics checklists as a foundation for students to form the habit of using technology in a humanistic, creative, and responsible manner, thereby gradually becoming outstanding digital citizens in the future.

REFERENCES

[1] Prime Minister of Vietnam, Decision approving the Scheme "*Strengthening the application of information technology and digital transformation in education and training for the period 2022-2025, with a vision to 2030*", No. 131/QD-TTg, 2022.

[2] Prime Minister of Vietnam, Decision approving the "*National Digital Transformation Program to 2025, with an orientation to 2030*", No. 749/QD-TTg, 2020.

[3] Ministry of Education and Training, *Circular Promulgating the General Education Program - Literature Program,* Circular No. 32/2018/TT-BGDDT, 2018.

[4] Ministry of Education and Training, *Circular Promulgating the Digital Competency Framework for Learners,* No. 02/2025/TT-BGDDT, 2025.

[5] UNESCO, "AI competency framework for students," *UNESCO Digital Library*, 2024. https://unesdoc.unesco.org/ark:/48223/pf0000391105 [Accessed March 03, 2026].

[6] P. D. L. Nguyen *et al*., "Academic integrity in the education and scientific research environment," *Vietnam Journal of Social Sciences*, no. 4, pp. 3-20, 2024. DOI: 10.56794/KHXHVN.4(196).3-20

[7] N. T. Nguyen and T. T. H. Nguyen, "Applying Artificial Intelligence (AI) to enhance self-study skills in Literature for students," *Journal of Education and Society*, no. 2, pp. 49–53, 2025.

[8] T. H. P. Tran and T. K. N. Chau, "Applying Artificial Intelligence (AI) in teaching Literature in high schools," *Journal of Educational Equipment*, vol. 1, no. 308, pp. 176–178, 2024.

[9] W. Holmes *et al*., *Artificial Intelligence in Education: Promise and Implications for Teaching and Learning*, Boston, MA, USA: Center for Curriculum Redesign, 2019.

[10] E. Kasneci *et al*., "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, p. 102274, 2023. https://doi.org/10.1016/j.lindif.2023.102274