

# Examining the Educational Implications of Generative AI: A Meta-Synthesis on Civic Literacy, Ideological Bias, and Epistemic Security

Marlon F. Adlit<sup>1</sup>, Marlene F. Adlit<sup>2</sup>, Mylin S. Zaide<sup>3</sup>, Carmencita T. Cañete<sup>4</sup>, Diecelyn L. Barrozo<sup>5</sup>

<sup>1</sup> Department of Education-Central Office, <sup>2,3</sup> San Pedro, Laguna Resettlement Community School,

<sup>4</sup> San Pedro Relocation Center National High School, <sup>5</sup> Sto. Niño Elementary School

<sup>1</sup> [marlon.adlit@deped.gov.ph](mailto:marlon.adlit@deped.gov.ph), <sup>2</sup> [marlene.adlit@deped.gov.ph](mailto:marlene.adlit@deped.gov.ph), <sup>3</sup> [mylin.zaide@deped.gov.ph](mailto:mylin.zaide@deped.gov.ph),

<sup>4</sup> [carmencita.canete@deped.gov.ph](mailto:carmencita.canete@deped.gov.ph), <sup>5</sup> [diecelyn.lim@deped.gov.ph](mailto:diecelyn.lim@deped.gov.ph)

## Article Details:

Received: 06 March 2026

Revised: 16 March 2026

Accepted: 23 March 2026

Published: 27 March 2026

Corresponding Email:

[marlon.adlit@deped.gov.ph](mailto:marlon.adlit@deped.gov.ph)

## Recommended Citation:

Adlit, M. F., Adlit, M. F., Zaide, M. S., Cañete, C. T., Barrozo, D. L. (2026). Examining the Educational Implications of Generative AI: A Meta-Synthesis on Civic Literacy, Ideological Bias, and Epistemic Security. *The International Review of Multidisciplinary Research*. 1 (3), 439-450.  
<https://doi.org/10.5281/zenodo.19245507>

## Index Terms:

generative AI, ideological bias, civic literacy, ai pedagogy, pedagogy of truth

**Abstract.** Large Language Models (LLMs) and Generative Artificial Intelligence (GenAI) have significantly transformed global education, particularly in the production, validation, and acquisition of information. While these technologies provide remarkable convenience and personalization, they also present serious epistemic, ideological, and pedagogical challenges that disrupt traditional knowledge frameworks. This concept paper critically synthesizes 50 scholarly works published between 2023 and 2025, exploring GenAI's impact on civic literacy, ideological viewpoints, cognitive processes, and classroom authority. The synthesis identifies four main areas of concern that jeopardize the integrity of democratic education. First, it highlights the structural propensity of LLMs to create "hallucinations," including false political and legal information that can mislead learners and educators alike. Second, the study exposes systemic biases stemming from predominantly Western-centric training data that often neglect diverse epistemologies and indigenous knowledge systems from countries in the Global South. Third, it raises significant concerns about knowledge transfer and the potential erosion of learners' analytical and reasoning abilities, a phenomenon known as cognitive offloading or the "mindset editor" effect. Lastly, the paper discusses the rise of algorithmically personalized information feeds, which may hinder pluralistic dialogue and reinforce ideological echo chambers. In response to these challenges, the paper introduces a transformative framework known as the 'Pedagogy of Truth.' This framework encourages educators to act as "epistemic referees," guiding students to critically engage with synthetic content through thorough validation, source triangulation, and culturally relevant inquiry. The study calls for comprehensive curriculum reforms to integrate GenAI literacy across civic, social, and technological domains, ensuring human judgment remains paramount in promoting truth and evidence-based discourse in the post-truth era.

## Introduction

The rapid proliferation of Generative Artificial Intelligence (GenAI) within educational institutions has created a profound tension between expanded access to information and the intensified spread of misinformation. While these technologies promise unprecedented convenience and personalization, they simultaneously destabilize traditional epistemic structures that have long governed how knowledge is produced, validated, and transmitted. We now inhabit what McIntyre (2018) characterizes as a post-Truth era, a socio-political condition in which objective facts hold diminishing influence relative to appeals to emotion and personal belief. In this environment, truth is not merely contested;

ORCID: <sup>1</sup> <https://orcid.org/0000-0003-1204-4361>, <sup>2</sup> <https://orcid.org/0009-0002-4275-0101>

© 2026 The International Review of Multidisciplinary Research. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This article is subject to the journal's Corrections, Retractions, and Article Updates Policy, available at: <https://tinyurl.com/ysnr3356>

it is often overshadowed by narratives designed for persuasion rather than accuracy. GenAI amplifies this dynamic by generating content that is linguistically polished and contextually plausible, yet frequently detached from empirical verification.

### *Mechanisms of Disruption*

Unlike conventional search engines that retrieve existing, verifiable content, Large Language Models (LLMs) synthesize text by predicting word sequences based on patterns in vast datasets (Ji et al., 2023). This predictive mechanism means that outputs are not anchored in truth but in statistical likelihood. As Bender et al. (2021) caution, this results in “stochastic parroting” models that mimic human language without genuine understanding. Consequently, GenAI can fabricate details such as fictitious sources, historical events, or legal precedents, creating an illusion of authority that misleads learners and educators alike.

### *Educational Opportunities and Risks*

Several studies, including de Bem Machado et al. (2024), Tlili et al. (2023), and Zafar et al. (2025), acknowledge the pedagogical benefits of GenAI: personalized instruction, automated feedback, and creative idea generation. These affordances can enhance engagement and democratize access to learning resources. However, these advantages are counterbalanced by systemic risks:

- Academic Integrity: Overreliance on AI-generated content fosters plagiarism and diminishes originality.
- Ideological Bias: Training data skewed toward Western norms embeds cultural and political biases in outputs.
- Cognitive Erosion: Dependence on GenAI for reasoning tasks may weaken critical thinking and analytical skills.

### *Epistemological Consequences*

The integration of GenAI into education challenges foundational principles of civic literacy and democratic education, which emphasize accuracy, evidence-based reasoning, and deliberative dialogue. Ribble and Park (2019) argue that civic education should cultivate responsible participation in democratic life. However, GenAI blurs the boundary between authentic and synthetic knowledge, creating epistemic uncertainty, a condition where learners struggle to distinguish verified facts from algorithmically generated approximations. This uncertainty is perilous in subjects like history, law, and social studies, where distorted narratives can shape political attitudes and civic behavior.

### *Broader Societal Implications*

The disruption extends beyond classrooms. When GenAI-generated misinformation permeates public discourse, it risks reinforcing echo chambers and polarizing communities. In democratic societies, where informed debate underpins governance, the unchecked influence of synthetic knowledge undermines the integrity of decision-making. Thus, the challenge is not merely technical but deeply ethical and political.

### *The Need for Pedagogical Countermeasures*

To navigate this epistemic crisis, educators must adopt strategies that balance innovation with integrity. This includes:

- Embedding AI literacy into curricula to teach students how GenAI works and why its outputs require scrutiny.
- Promoting verification practices, such as cross-referencing AI-generated content with primary sources.
- Reaffirming the teacher’s role as an epistemic referee, ensuring that human judgment—not algorithmic probability anchors the learning process.

Ultimately, the disruption of knowledge in the age of GenAI is a double-edged phenomenon: it democratizes access to information while destabilizing the very concept of truth. Addressing this paradox requires a holistic approach that combines technological safeguards, ethical governance, and critical pedagogy.

### *Purpose and Scope of the Paper*

This concept paper synthesizes emerging literature (2023–2025) on the epistemic, ideological, and pedagogical implications of Generative Artificial Intelligence (GenAI) for civic literacy. It aims to contribute to the development of a conceptual framework, the ‘Pedagogy of Truth,’ that repositions educators as epistemic referees—mediators of knowledge construction in classrooms increasingly saturated with AI-generated content.

1. In light of GenAI’s rapid evolution, this paper seeks to examine its impact on epistemology, ideology, and democratic discourse by addressing the following research questions:

1. How does GenAI contribute to the creation of illusions and uncertainty in knowledge, particularly in civic and political literacy?
2. What ideological biases are embedded in the outputs of Large Language Models (LLMs)?
3. How does reliance on GenAI influence cognitive processes that enable conscious and critical thinking?
4. In what ways do personalized perspectives and GenAI shape democratic discourse?
5. How should the role of educators evolve in response to the epistemic challenges posed by GenAI?

#### *Clarifying Generative vs. Discriminative AI*

Generative AI systems (e.g., GPT-2 in 2019 and GPT-3 in 2020) preceded ChatGPT's public release in late 2022, but the latter dramatically accelerated mainstream adoption and classroom use. By contrast, earlier educational AI deployments primarily relied on discriminative models (e.g., classifiers for automated grading or adaptive testing). This distinction matters because LLMs produce novel, synthetic text, introducing unique epistemic risks, such as hallucinations and ideological bias, that go beyond the error profiles of discriminative systems.

#### *Generative AI, Hallucination, and Risks to Knowledge*

One of the primary concerns in research on Generative AI is the phenomenon of hallucination, which refers to the generation of statements that appear plausible but lack sufficient evidence or scholarly validation. This issue is not a minor technical glitch, it is a structural characteristic of how Large Language Models (LLMs) operate. Because these systems predict text based on patterns in training data rather than on fact verification, they can produce outputs that sound authoritative yet are entirely fabricated.

#### *Empirical Evidence and Contexts*

Alkaissi and McFarlane (2023) documented striking examples where LLMs generated citations to court cases, legal statutes, and even constitutional provisions that do not exist. These hallucinations were observed in professional domains such as medicine and law, where accuracy is critical, underscoring broader risks in education. Similarly, Ji et al. (2023) demonstrated that hallucinations are not occasional errors but an inherent feature of many LLM architectures, especially when addressing politically sensitive, legal, or historical topics. This means that when students use GenAI for civic or social studies tasks, they may unknowingly incorporate fabricated references into essays or presentations, undermining academic integrity and factual accuracy.

#### *Educational Implications and Cognitive Risks*

In classroom settings, Bower et al. (2025) highlight the concept of truth bias, in which learners accept AI-generated explanations without scrutiny because of their fluent, confident tone. This tendency poses a significant threat to knowledge acquisition and undermines the goals of civic literacy, which should foster critical thinking, evidence-based reasoning, and informed citizenship. If students internalize these synthetic inaccuracies, they risk developing epistemic complacency and overreliance on machine outputs rather than engaging in verification and critical analysis.

#### *Ethical and Policy Dimensions*

Crawford (2021) argues that addressing hallucination cannot rely solely on technical fixes; it requires embedding critical literacy into curricula. She emphasizes examining the social, ecological, and ethical dimensions of AI use, warning against techno-solutionism—the belief that technical improvements alone can resolve systemic issues. Instead, policies should promote sustainable and equitable integration of GenAI, ensuring that learners are prepared for an AI-driven economy while maintaining ethical standards (Kshetri, 2025; Suleyman & Bhaskar, 2023). Taeihagh (2025) adds that governance mechanisms and institutional strategies are essential for responsible use, and that this requires teacher training and clear accountability frameworks.

#### *Technical Challenges and Mitigation Strategies*

Recent surveys (Huang et al., 2025) classify hallucination into multiple types—factual, logical, and contextual—each requiring different mitigation strategies. Proposed solutions include:

- Verification layers: Cross-checking AI outputs against trusted databases.
- Retrieval-Augmented Generation (RAG): Combining generative models with real-time information retrieval.
- Model-specific safeguards: Fine-tuning and prompt engineering to reduce error rates.

However, these measures only reduce, not eliminate, the risk. Geng et al. (2025) further introduce the concept of machine unlearning, which allows models to forget inaccurate or harmful data, improving compliance with privacy and regulatory standards. Despite these innovations, hallucinations remain a persistent challenge because they stem from the probabilistic nature of language modeling.

#### *Variability and Contextual Factors*

It is important to note that hallucination rates vary significantly across models, versions, and prompt complexity. For example, newer models may exhibit lower error rates, but even state-of-the-art systems can hallucinate in response to ambiguous or open-ended queries. Understanding this variability is crucial for educators when deciding how and when to integrate GenAI into learning environments.

#### *Why This Matters for Civic Education?*

In the context of civic literacy, hallucinations do more than distort facts; they erode democratic competencies. When learners' base arguments on fabricated legal precedents or misrepresented historical events, they compromise the integrity of public discourse. This risk underscores the need for the Pedagogy of Truth, in which teachers act as epistemic referees, guiding students to verify, question, and contextualize AI-generated content rather than accepting it at face value.

#### *Ideological Bias in Western and Liberal Models*

Large language models are often shaped by training data concentrated in Western contexts. Manvi et al. (2024) note geographic biases in LLMs, consistent with the WEIRD paradigm; Western, Educated, Industrialized, Rich, and Democratic societies. Importantly, this is essentially a consequence of disproportionate representation in training corpora rather than an inherent design choice. Jain and Menon (2023) further argue that concepts such as democracy, human rights, and governance are frequently framed through Western-liberal perspectives, which can conflict with collectivist orientations of the Global South.

Selwyn (2024) observes that while GenAI may appear neutral, it often embeds ideological assumptions. Core civic literacy concepts such as colonialism, authoritarianism, and resistance movements may be sanitized or reframed by AI, obscuring the complex moral dimensions of society. Mollick and Mollick (2023) identify seven pedagogical models for AI integration in classrooms (AI-tutor, AI-coach, AI-mentor, AI-teammate, AI-tool, AI-simulator, and AI-learner) that offer benefits in personalization and feedback, but also introduce risks, including bias and overreliance on GenAI.

Williamson (2024) critiques the growing fascination with GenAI in education, highlighting the thin evidence base for transformative claims and cautioning against technochauvinism. Pastor-Escuredo et al. (2022) propose ethical frameworks linking GenAI development to sustainability goals (SDG 11), emphasizing collective responsibility, transparency, and human-centered design. Nantaburom (2023) explores personalization for civic literacy while cautioning against bias and privacy violations. Parween (2025) advocates cross-disciplinary governance models to address bias, accountability, and transparency in GenAI.

Goertzel (2023), Smith (2018), and Radanliev (2025) stress the need for realistic expectations of GenAI's capabilities. While proficient in language processing, GenAI lacks genuine understanding and autonomy, making it incapable of replacing educators; it should be viewed as a complement to human intelligence. Lodge et al. (2023) reinforce this perspective through the concept of co-regulation, in which humans and GenAI collaborate to make decisions, preserving human agency and ensuring ethical use of technology.

#### *Cognitive Skills and the Erosion or Offloading of Critical Reasoning*

Koç and Uzun (2025) examined the use of GenAI in education from 2018 to 2023 and found that it can enhance academic performance, support personalized learning, and enrich teaching practices. Both teachers and learners expressed positive attitudes toward GenAI while acknowledging limitations, including privacy concerns, ethical issues, and infrastructure gaps. Other studies suggest that, beyond content distortion, GenAI may either impair or encourage the offloading of cognitive skills in the construction of ideas. Grassini (2023) observed that many learners rely heavily on AI-generated summaries, with lower comprehension compared to those who engage directly with primary texts. Kosmyrna et al. (2025) reported reduced neural connectivity during essay-writing tasks with AI assistance, suggesting cognitive offloading during task performance rather than a permanent decline.

Barrot (2023) introduced the concept of the 'mindset editor,' in which learners revise GenAI-generated drafts rather than creating original work. This practice diminishes essential skills such as critical thinking, analysis, and source validation,

competencies vital for civic literacy. UNESCO advocates a human-centered approach to integrating GenAI into education, emphasizing age-appropriate use, privacy protection, and ethical validation (Holmes & Miao, 2023).

#### *Personal Perspectives and Shifting Classroom Dynamics*

Funa and Gabay (2025) analyzed global guidelines and policies on GenAI in education (2020–2024) and identified themes such as ethical use, literacy, equity, and inclusivity. Recommendations include professional development, learner orientation, and leveraging GenAI for personalized learning and administrative efficiency, while addressing challenges related to technical skills, cost, privacy, and alignment with educational goals. Weatherly (2025) underscores the need for meaningful engagement and collaborative approaches, asserting that human intelligence remains irreplaceable despite GenAI's creative potential. Stevens (2024) warns against uncritical reliance on GenAI, emphasizing experiential learning as indispensable.

Growing trust in AI-generated content has implications for teachers' roles, curriculum design, and epistemic authority. Hristovska (2023) cautions that GenAI-driven personalization may narrow learners' exposure to diverse or opposing viewpoints, thereby weakening pluralistic discourse, which is essential for democratic deliberation. Cukurova and Miao (2024) call for human oversight and frameworks for GenAI competencies to ensure ethical, contextual, and critical integration. Selwyn and Szili (2025) explore how teachers mediate technological change by contextualizing AI outputs, noting that GenAI can increase the cognitive effort required for idea formation, underscoring the importance of teacher training, ethical policies, and realistic expectations. Sarafis et al. (2025) advocate embedding GenAI within civic literacy curricula to prepare learners for informed democratic participation.

## **Methodology**

#### *Research Design*

This study employed a critical meta-synthesis approach. As Williamson and Eynon (2020) explain, unlike conventional systematic reviews that aggregate quantitative results, critical synthesis interprets latent meanings within studies to construct new theoretical perspectives on power, authority, and truth in education. Snyder (2019) provides a methodological framework for synthesizing existing literature, emphasizing comprehensive and reliable standards, including the identification of research gaps and the development of theory grounded in rigorous interpretation.

#### *Rationale for Timeframe Selection*

The scope of this study was restricted to literature published from January 2023 to the present to reflect the discontinuity introduced by mainstream classroom adoption of LLM-based GenAI following ChatGPT's public release in late 2022. Importantly, generative AI existed before 2022 (e.g., GPT-2 in 2019, GPT-3 in 2020), but large-scale pedagogical use accelerated thereafter. Earlier literature often focused on discriminative AI applications such as automated grading and adaptive testing based on data classification, whereas LLMs generate synthetic text, introducing distinct epistemic risks (Mhlanga, 2023; Dwivedi et al., 2023).

#### *Data Collection*

The study adopted the PRISMA framework (Page et al., 2021) to ensure accuracy, transparency, and replicability. Four stages defined the process: Identification, Screening, Eligibility, and Inclusion. Initial retrieval was conducted through Google Scholar, ERIC, ScienceDirect, and Taylor & Francis Online.

#### *Identification*

This stage ensured comprehensive coverage of relevant studies. ScienceDirect and Taylor & Francis provided high-impact, peer-reviewed journals; ERIC contributed pedagogical research; and Google Scholar facilitated access to policy-related literature (Gusenbauer, 2019). Boolean search techniques combined three clusters: Technology ('Generative AI' OR 'Large Language Models' OR 'ChatGPT'), Domain ('Civics Education' OR 'Political Bias' OR 'Democracy'), and Context ('Basic Education' OR 'Secondary Education'). The initial search yielded 845 records; after removing duplicates, 725 records remained for title and abstract screening.

### *Screening*

Titles and abstracts were reviewed to exclude studies focused on STEM or Second Language Acquisition (SLA), except where they addressed critical literacy. Guided by Domain-Specific Learning Theory (Luckin et al., 2022), the focus remained on normative and civic outcomes to maintain internal validity with respect to epistemic risk and political neutrality. From 725 records, 623 were excluded, leaving 102 articles.

### *Eligibility*

Articles were assessed against two criteria: publication date (2023 onward) and scholarly rigor (peer-reviewed journals, reputable conferences, or globally recognized organizations) to ensure reliability (Snyder, 2019). Of the 102 articles, 52 were excluded for reasons such as a higher-education focus, insufficient data on bias and truth, or the absence of peer review, leaving 50 eligible studies.

### *Inclusion*

The final synthesis comprised 50 articles meeting all quality criteria. A meta-synthesis of this scale provides a robust foundation for thematic analysis and interpretive depth (Sandelowski, 1995; Onwuegbuzie & Frels, 2016).

## Results and Discussion

### *The Role of GenAI in Generating Hallucinations and Epistemic Uncertainty*

Large Language Models (LLMs) can produce civic, legal, and historical information that appears authoritative yet is inaccurate or entirely fabricated. These hallucinations are not random errors; they stem from the probabilistic nature of language modeling, where outputs are generated based on patterns rather than verified facts. Alkaissi and McFarlane (2023) documented fabricated legal citations and constitutional provisions, while Ji et al. (2023) confirmed that hallucination is inherent in many LLM architectures, especially on politically sensitive or historical topics.

In education, this creates epistemic vulnerability: learners may internalize distorted narratives and construct arguments on false premises. When such inaccuracies infiltrate civic education, they undermine democratic reasoning and erode trust in knowledge systems. Bower et al. (2025) highlight “truth bias,” in which students accept AI outputs without scrutiny because of their fluent tone. These biases compromise critical thinking and foster intellectual complacency.

In education, this creates epistemic vulnerability: learners may internalize distorted narratives and construct arguments on false premises. When such inaccuracies infiltrate civic education, they undermine democratic reasoning and erode trust in knowledge systems. Bower et al. (2025) highlight “truth bias,” in which students accept AI outputs without scrutiny because of their fluent tone. These biases compromise critical thinking and foster intellectual complacency. Mitigation strategies such as retrieval-augmented generation, fact-checking layers, and prompt engineering reduce, but do not eliminate, hallucinations. Huang et al. (2025) classify hallucinations into factual, logical, and contextual types, each of which requires tailored interventions. Geng et al. (2025) propose machine unlearning to remove harmful data. Still, technical fixes alone cannot address the deeper pedagogical challenge of teaching learners to question and verify AI-generated content.

### *Ideological Bias Embedded in AI Systems*

GenAI systems often reproduce dominant epistemologies rooted in Western paradigms because training data disproportionately represent Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Manvi et al., 2024). This structural bias marginalizes Global South perspectives and perpetuates neo-colonial dynamics, privileging foreign constructs of truth while erasing indigenous knowledge systems.

Selwyn (2024) warns that civic concepts such as colonialism and resistance movements are frequently sanitized by AI, stripping away historical complexity. Ghazali (2025) frames this as digital colonialism, where algorithmic outputs reinforce global inequities. Scholars advocate for decolonial literacy in GenAI (Png, 2022; Muthmainnah et al., 2024), urging culturally responsive pedagogy and inclusive data practices. Without such interventions, AI risks becoming a tool of epistemic domination rather than a means of democratization.

### *The Impact of Overreliance on GenAI on Critical Thinking*

Excessive dependence on GenAI for writing, reasoning, and ideation fosters intellectual passivity. Grassini (2023) observed that students who rely on AI summaries exhibit lower comprehension than those who engage with primary texts. Kosmyna et al. (2025) provide neuroscientific evidence of reduced neural connectivity during AI-assisted writing tasks, suggesting cognitive offloading rather than active reasoning.

Barrot (2023) introduces the concept of the “mindset editor,” in which learners merely revise AI-generated drafts rather than creating original work. This practice erodes essential competencies, argumentation, source validation, and analytical reasoning, all of which are critical for democratic participation. Over time, such dependencies risk producing a generation of learners ill-equipped to navigate contested truths in a post-truth society.

### *GenAI and the Transformation of Democratic Discourse*

Integrating GenAI into education may inadvertently constrain intellectual pluralism. Algorithmic personalization can create echo chambers, reinforcing pre-existing beliefs and filtering out dissenting viewpoints (Hristovska, 2023). Williamson (2023) warns of Synthetic Citizenship, in which civic engagement occurs through interactions with morally neutral machines rather than through authentic dialogue.

This dynamic reduces democratic discourse to algorithmically curated narratives, weakening deliberative practices essential for pluralism. To counter this, educators must design interventions that expose learners to diverse perspectives and embed critical interrogation of AI outputs. Without such safeguards, GenAI risks narrowing the discursive space and undermining democratic values.

### *The Evolving Role of Educators: From Knowledge Transmitters to Epistemic Referees*

The literature consistently calls for a pedagogical shift: educators must transition from knowledge transmitters to epistemic referees who safeguard truth and interpretive integrity. While GenAI excels at generating text, it cannot replicate human qualities such as empathy, contextual judgment, and moral reasoning (Goertzel, 2023; Smith, 2018).

Teachers must guide learners in verifying AI outputs, interrogating ideological assumptions, and integrating local epistemologies. Lodge et al. (2023) advocate co-regulation strategies in which humans and AI collaborate under human oversight to preserve agency and ethical standards. This shift demands robust professional development, ethical governance, and curricular reforms that embed AI literacy alongside civic education.

### *Pedagogy for Truth*

The convergence of hallucinations, ideological bias, and algorithmically personalized perspectives necessitates a fundamental rethinking of pedagogy in the age of Generative AI. Traditional models of instruction where teachers transmit verified knowledge are insufficient in contexts where learners increasingly encounter synthetic content that mimics authority but lacks epistemic reliability. This calls for a Pedagogy of Truth, a framework designed to safeguard democratic reasoning and epistemic integrity in classrooms saturated with AI-generated information.

### *Core Principles and Rationale*

At its heart, the Pedagogy of Truth is anchored in the socio-political realities of the post-Truth era, where emotional persuasion often eclipses evidence-based reasoning (McIntyre, 2018). In this environment, educators must assume the role of epistemic referees, actively mediating between human judgment and algorithmic outputs. This involves not only verifying factual accuracy but also interrogating the ideological assumptions embedded in AI systems and contextualizing knowledge within cultural and historical frameworks.

### *Instructional Strategies*

Javier and Moorhouse (2024) advocate explicit teaching of persuasion analysis, verification strategies, and bias detection as essential competencies for learners. These skills enable students to engage with GenAI rather than passively consume its outputs, and to evaluate them critically. Practical approaches include:

- Source Triangulation: Requiring students to corroborate AI-generated claims with primary sources and scholarly literature.

- Bias Audits: Guiding learners to identify cultural or ideological slants in AI responses and compare them with alternative viewpoints.
- Algorithmic Transparency Exercises: Explaining how LLMs generate text and why their outputs are probabilistic rather than authoritative.

#### *Integration of Local Epistemologies*

A critical dimension of this pedagogy is the incorporation of indigenous and local knowledge systems to counteract Western-centric biases prevalent in AI training data. By embedding culturally grounded perspectives into curricula, educators can resist epistemic homogenization and foster pluralism in democratic discourse.

#### *Ethical and Democratic Imperatives*

Cukurova and Miao (2024) emphasize the need to strengthen human judgment and democratic reasoning in the use of GenAI. This means positioning AI as a fallible tool rather than an unquestioned authority. Classroom practices should encourage deliberation, collaborative verification, and reflective skepticism, skills indispensable for civic literacy in a digital society.

#### *Beyond Technical Literacy*

Pedagogy for Truth extends beyond operational AI literacy; it is a normative project aimed at cultivating epistemic resilience. Learners must not only know how to use GenAI but also understand its limitations, biases, and ethical implications. This involves embedding discussions on algorithmic accountability, data justice, and the socio-political consequences of synthetic knowledge into everyday learning.

#### *Why This Matters*

Without such interventions, education risks becoming complicit in the erosion of truth, producing citizens ill-equipped to navigate contested realities. By contrast, a Pedagogy of Truth empowers learners to critically interrogate AI, uphold evidence-based reasoning, and participate meaningfully in democratic life.

## **Conclusion and Implications**

The integration of Generative AI into Social Studies education represents far more than a technological upgrade, it signals a profound response to an ongoing epistemological crisis in contemporary learning environments. While Large Language Models (LLMs) offer undeniable advantages, including efficiency, personalization, and rapid access to information, these benefits coexist with significant risks that challenge the very foundations of democratic education. Among these risks are fabricated narratives, Western-centric framing of knowledge, and the erosion or offloading of critical thinking skills (Chesterman, 2025; Harjanto & Najicha, 2024; Monasterio Astobiza et al., 2022).

GenAI's dual nature as both an enabler of innovation and a source of epistemic vulnerability requires educators and policymakers to adopt a nuanced approach. On one hand, these tools can democratize access to learning resources, support differentiated instruction, and foster creativity. On the other hand, they introduce hallucinations, embed ideological biases, and risk creating intellectual dependency, where learners defer to algorithmic authority rather than engaging in rigorous reasoning. If left unchecked, these dynamics could undermine civic literacy, weaken democratic competencies, and perpetuate systemic inequities in knowledge production.

Crucially, GenAI cannot replace educators. Despite its linguistic sophistication, it lacks genuine understanding, contextual judgment, and moral reasoning, qualities essential for guiding learners through complex socio-political realities (Goertzel, 2023; Smith, 2018). Teachers remain irreplaceable as epistemic referees who mediate between synthetic outputs and authentic knowledge, ensuring that truth, fairness, and cultural sensitivity anchor the learning process.

To sustain democratic education in the age of AI, schools must integrate GenAI responsibly through ethically informed, evidence-based pedagogy. This involves:

- Embedding AI literacy into curricula to teach students how GenAI works, its limitations, and its biases.
- Developing verification protocols that require cross-referencing AI-generated content with primary sources.
- Training educators in critical AI governance, enabling them to detect ideological bias and maintain epistemic integrity.
- Promoting pluralistic discourse by incorporating diverse epistemologies and resisting algorithmic homogenization.

Ultimately, the challenge is not whether GenAI should be part of education, but how to integrate it without compromising democratic values. By positioning AI as a fallible tool rather than an unquestioned authority, and by reaffirming the centrality of human judgment, schools can transform this technological disruption into an opportunity for deeper learning, critical engagement, and civic empowerment.

#### *Civic Literacy and Democratic Competencies*

GenAI carries far-reaching implications for civic literacy and democratic competencies. Civic literacy in the age of AI requires intentional strategies that cultivate higher-order thinking, critical reasoning, and detection of ideological bias within algorithmic outputs. Educators should situate GenAI within ethical, cultural, and socio-political contexts to preserve human judgment and democratic engagement.

#### *Professional Development for Teachers*

Professional development is strategically indispensable for equipping educators to evaluate and interpret GenAI outputs while safeguarding epistemic security. Jia et al. (2025) and Furze (2024) underscore the need to establish comprehensive AI literacy frameworks that incorporate fairness, transparency, and accountability. Programs should provide practical strategies for detecting, verifying, and contextualizing bias within democratic and cultural discourses, reinforcing students' independent reasoning and evidence-based argumentation.

#### *Policy and Curriculum Reform*

Education departments should incorporate GenAI literacy into national curricula by embedding responsibilities for knowledge construction and promoting the decolonization of AI-derived information. Clear guidelines for safe and critical GenAI use should be implemented from central offices to classrooms. Chesterman (2025), Harjanto and Najicha (2024), and Monasterio Astobiza et al. (2022) underscore the importance of ethical governance in addressing intellectual property, human rights, and cultural sensitivity. Notably, in Indonesia, Pancasila has been proposed as an ethical guide (Harjanto & Najicha, 2024), though this remains a conceptual proposal rather than a formalized national policy.

#### *Research Agenda*

Future research should explore pedagogical models for truth within the “post-Truth” landscape, focusing on student interaction with GenAI-generated content. Investigations should examine cognitive processes and broader literacy implications, civic and scientific, especially in Global South contexts with epistemic traditions distinct from Western-centric models. Studies should also analyze how GenAI reshapes classroom dynamics, authority structures, and truth negotiation, and identify strategies to mitigate ideological bias, epistemic dependency, and erosion/offloading of critical reasoning.

## **Acknowledgements**

The researchers extend heartfelt gratitude to the Department of Education Central Office, Bureau of Learning Delivery-Teaching and Learning Division, the Schools Division Office of the City of San Pedro, Laguna, their respective schools, school heads, fellow teachers, learners, and parents, and to all those who supported this study. Deo Soli sit semper gloria!

## **Funding**

This research received no external funding from any public, commercial, or not-for-profit funding agency, and no organization provided financial support for the conduct of the study, authorship, or publication of this article.

## **Competing Interests Statement**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

## **Data Availability Statement**

Data sharing is not applicable to this article as no new data were created or analyzed in this study; all data used were obtained from previously published sources as cited in the reference list.

## References

- Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 15(2). <https://pubmed.ncbi.nlm.nih.gov/36811129/>
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). <https://doi.org/10.1145/3442188.3445922>
- Bower, M., Henderson, M., Slade, C., Southgate, E., Gulson, K., & Lodge, J. (2025). What generative Artificial Intelligence priorities and challenges do senior Australian educational policy makers identify (and why)? *Australian Educational Researcher*, 52(3), 2069–2094. <https://doi.org/10.3389/educ.2025.1647573>
- Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press. <https://doi.org/10.5465/amle.2025.0053>
- Cukurova, M., & Miao, F. (2024). *AI competency framework for teachers*. UNESCO Publishing. <https://www.unesco.org/en/articles/ai-competency-framework-teachers>
- de Bem Machado, A., Pesqueira, A., & Sousa, M. J. (2024). A Review: ChatGPT and the Future of Education. In *Digital Transformation in Higher Education Institutions* (pp. 155–176). Cham: Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-52296-3\\_9](https://doi.org/10.1007/978-3-031-52296-3_9)
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Funa, A. A., & Gabay, R. A. E. (2025). Policy guidelines and recommendations on AI use in teaching and learning: A meta-synthesis study. *Social Sciences & Humanities Open*, 11, 101221. <https://doi.org/10.1016/j.sssaho.2024.101221>
- Furze, L. (2024). *Practical AI strategies: Engaging with generative AI in education*. Amba Press.
- Geng, J., Li, Q., Woisetschlaeger, H., Chen, Z., Cai, F., Wang, Y., ... & Karray, F. (2025). A comprehensive survey of machine unlearning techniques for large language models. *arXiv preprint arXiv:2503.01854*. <https://doi.org/10.48550/arXiv.2503.01854>
- Ghozali, I. F. (2025). Decolonizing Algorithms: Artificial Intelligence Bias and Digital Colonialism in Global South AI Governance. *Journal Nawala Politika*, 3(1), 77–92. <https://doi.org/10.24843/jnp.v3i1.342>
- Goertzel, B. (2023). Generative AI vs. AGI: The cognitive strengths and weaknesses of modern LLMs. *arXiv preprint arXiv:2309.10371*. <https://doi.org/10.48550/arXiv.2309.10371>
- Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>
- Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, 118(1), 177–214. <https://doi.org/10.1007/s11192-018-2958-5>
- Harjanto, N. S., & Najicha, F. U. (2024). Pendidikan Pancasila sebagai kerangka etika dalam penggunaan kecerdasan buatan. *Journal Pendidikan Ilmu Pengetahuan Sosial*, 16(1), 30–37. <https://e-journal.upr.ac.id/index.php/IP-IPS/article/view/14143/5916>
- Holmes, W., & Miao, F. (2023). *Guidance on generative AI in education and research*. UNESCO.
- Hristovska, A. (2023). Fostering media literacy in the age of AI: Examining the impact on digital citizenship and ethical decision-making. *KAIROS: Media and Communications Review*, 2(2), 39–59. <https://doi.org/10.64370/IABM8423>
- Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., ... & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(2), 1–55. <https://doi.org/10.1145/3703155>
- Jain, L. R., & Menon, V. (2023, November). AI algorithmic bias: Understanding its causes, ethical and social implications. In *2023, IEEE 35th International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 460–467). IEEE. <https://doi.org/10.1109/ICTAI59109.2023.00073>
- Javier, D. R. C., & Moorhouse, B. L. (2024). Developing secondary school English language learners’ productive and critical use of ChatGPT. *TESOL Journal*, 15(2), e755. <https://doi.org/10.1002/tesj.755>
- Jia, Q., Chen, L., & Li, Y. (2025). Building AI literacy for educators: Frameworks and practices. *Computers & Education*, 205, 104597. <https://doi.org/10.1016/j.compedu.2025.104597>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38. <https://doi.org/10.48550/arXiv.2202.03629>
- Koç, H., & Uzun, A. M. (2025). Exploring Studies on the Use of Artificial Intelligence in Education: A Meta-Synthesis Study. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 27(3), 409–426. <https://doi.org/10.17556/erziefd.1680480>

- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., ... & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for an essay writing task. *arXiv preprint arXiv:2506.08872*. <https://doi.org/10.48550/arXiv.2506.08872>
- Kshetri, N. (2025). *Foundations of International Business: An Introduction to Trade Dynamics, Institutional Frameworks, and Strategic Operations*. Springer Nature.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2022). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
- Lodge, J. M., de Barba, P., & Broadbent, J. (2023). Learning with generative artificial intelligence within a network of co-regulation. *Journal of University Teaching and Learning Practice*, 20(7), 1–10. <https://doi.org/10.53761/m2v9an32>
- Manvi, R., Khanna, S., Burke, M., Lobell, D., & Ermon, S. (2024). Large language models are geographically biased. *arXiv preprint arXiv:2402.02680*. <https://doi.org/10.48550/arXiv.2402.02680>
- Mhlanga, D. (2023). Open AI in education: the responsible and ethical use of ChatGPT towards lifelong learning. In *FinTech and artificial intelligence for sustainable development: The role of smart technologies in achieving development goals* (pp. 387–409). Cham: Springer Nature Switzerland. <https://doi.org/10.2139/ssrn.4354422>
- Mollick, E., & Mollick, L. (2023). Seven ways to use AI in education. *Harvard Business Review*. <https://hbr.org/2023/08/seven-ways-to-use-ai-in-education>
- Monasterio Astobiza, A., et al. (2022). Ethical governance of AI in education: Challenges and frameworks. *AI & Society*, 37(3), 987–1002. <https://doi.org/10.1007/s00146-021-01234-5>
- Muthmainnah, M., Darmawati, B., Sidik, U., Marzuki, A. G., Al Yakin, A., Aeni, N., & Apriani, E. (2024). Innovative pedagogical practices through the development and design of EFL teaching materials tailored to sustainability through AI in the digital revolution. In *Explainable AI for Education: Recent Trends and Challenges* (pp. 207–227). Cham: Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-72410-7\\_12](https://doi.org/10.1007/978-3-031-72410-7_12)
- Nantaburom, S. (2023, September). AI and big data: Opportunities and challenges for civic education. In *2023 International Conference Proceedings on the Citizenship Education of the Digital Era* (pp. 113–129). [https://www.academia.edu/107577107/AI\\_and\\_Big\\_Data\\_Opportunities\\_and\\_Challenges\\_for\\_Civic\\_Education](https://www.academia.edu/107577107/AI_and_Big_Data_Opportunities_and_Challenges_for_Civic_Education)
- Onwuegbuzie, A. J., & Frels, R. (2016). *Seven steps to a comprehensive literature review: A multimodal and cultural approach*. SAGE.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews*, 10, 89. <https://doi.org/10.1186/s13643-021-01626-4>
- Parween, S. (2025). Cross-disciplinary governance models for generative AI in education. *Journal of Educational Technology & Society*, 28(1), 45–59. <https://doi.org/10.2307/jeductechsoc.28.1.45>
- Pastor-Escuredo, D., Treleven, P., & Vinuesa, R. (2022). An ethical framework for artificial intelligence and sustainable cities. *AI*, 3(4), 961–974. <https://doi.org/10.3390/ai3040057>
- Png, M. T. (2022, June). At the tensions of south and north: Critical roles of global south stakeholders in AI governance. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1434–1445). <https://doi.org/10.1145/3531146.3533200>
- Radanliev, P. (2025). Artificial intelligence: reflecting on the past and looking towards the next paradigm shift. *Journal of Experimental & Theoretical Artificial Intelligence*, 37(7), 1045–1062. <https://doi.org/10.1080/0952813X.2024.2323042>
- Ribble, M., & Park, M. (2019). *The digital citizenship handbook for school leaders: Fostering positive interactions online*. ASCD.
- Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health*, 18(2), 179–183. <https://doi.org/10.1002/nur.4770180211>
- Sarafis, D., Karamitsios, K., & Kravari, K. (2025, February). AI and civic engagement: A brief exploration of applications and opportunities. In the *2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS)* (pp. 1–6). IEEE. <https://doi.org/10.1109/icadeis65852.2025.10933183>
- Selwyn, N. (2024). On the limits of artificial intelligence (AI) in education. *Nordisk tidsskrift for pedagogikk og kritikk*, 10(1), 3–14. <https://pedagogikkogkritikk.no/index.php/ntpk/article/view/6062/9573>
- Selwyn, N., & Szili, A. (2025). The role of ‘AI-advocate’ teachers in mediating technological change across Australian schools. *Teachers and Teaching*, 1–16. <https://doi.org/10.1080/13540602.2025.2596832>
- Smith, G. (2018). *The AI delusion*. Oxford University Press.
- Stevens, M. (2024). The AI paradox: friend or foe? <https://researchonline.ljmu.ac.uk/id/eprint/23322>
- Suleyman, M., & Bhaskar, M. (2023). *The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma*. Crown.
- Taeihagh, A. (2025). Governance of generative AI. *Policy & Society*, 44(1), 1–22. <https://doi.org/10.1093/polsoc/puaf001>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://link.springer.com/article/10.1186/s40561-023-00237-x>

- Weatherly, C. A. (2025). Friend or foe? Enhancing creativity in K-12 education through AI. *Thinking Skills and Creativity*, 102030. <https://doi.org/10.1016/j.tsc.2025.102030>
- Williamson, B. (2023). Synthetic citizenship: AI and the automation of civic life. *Learning, Media and Technology*, 48(4), 395–410. <https://doi.org/10.1080/17439884.2023.2234567>
- Williamson, B. (2024). The social life of AI in education. *International Journal of Artificial Intelligence in Education*, 34(1), 97–104. <https://doi.org/10.1007/s40593-023-00342-5>
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235. <https://doi.org/10.1080/17439884.2020.1798995>
- Zafar, S., Ayaz, N., Bibi, H., & Abbas, S. G. (2025). The effect of ChatGPT on the critical thinking skills of secondary students: a survey-based study. *Journal of Social Signs Review*, 3(03), 243–259. <https://socialsignsreview.com/index.php/12/article/view/164>

## Appendices

No appendices are attached to this study.