

Governing Artificial Intelligence for Educational Justice: A Critical Pedagogy and Policy Analysis in the Philippine Context

John Mark N. Saldivar*

College of Teacher Education, La Salle University, Ozamiz City, Philippines, Email: johnmark.saldivar@lsu.edu.ph

A Review Article

*Corresponding author:

johnmark.saldivar@lsu.edu.ph

ORCID ID: <https://orcid.org/0009-0007-1446-4112>

Received: 07/03/2026

Revised: 28/03/2026

Accepted: 29/03/2026

Published: 04/05/2026

Citation:

Saldivar, J.M.N. (2026). Governing artificial intelligence for educational justice: A critical pedagogy and policy analysis in the Philippine context. *Journal of Teaching and Education for Scholars*, 3(1), 58-72.

<https://doi.org/10.59065/jotes.v3i1.280>



Copyright:

© 2026 by the author(s).

This is an open-access article: the use, distribution, and reproduction of published material are strictly regulated under the [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

Abstract

This study examines how artificial intelligence (AI) is framed and governed within Philippine educational policy, considering its implications for equity, teacher autonomy, and learner agency. The objective is to understand the power dynamics and ideological presuppositions that inform AI-centric reforms. This research employs an integrative analytical methodology that incorporates critical policy analysis, discourse tracing, thematic coding, and ideological critique. Methodological integrity is established through rigorous inclusion and exclusion criteria, adherence to the PRISMA protocol for systematic document review, and the application of the CASP Checklist to assess the credibility and pertinence of 21 significant policy documents published between 2021 and 2025. The findings reveal that prevailing policies primarily depict AI as a tool for modernization, operational efficiency, and data-driven governance. At the same time, critical concerns regarding ethics, human rights, and democratic participation remain insufficiently addressed. These accounts pose a substantial risk of exacerbating existing inequalities between well-resourced and under-resourced educational institutions. In response, the study advocates for the ARISE Framework (AI Rights-based Inclusive Systems for Education). This nascent governance model prioritizes structural preparedness, pedagogical agency, critical scrutiny of AI discourses, and rights-centric oversight. The study emphasizes the imperative of equity impact assessments, transparency requirements, accountability systems, and the continuous development of critical AI literacy among educators and students. These initiatives ensure that AI integration aligns with Sustainable Development Goal 4, thereby reinforcing rather than undermining equitable and inclusive educational opportunities.

Keywords: *Critical Policy Analysis; Critical pedagogy; AI Governance; ARISE Framework; SDG 4; Philippine Educational Policy*

1. Introduction

Artificial intelligence (AI) has emerged as a transformative force in education, reshaping not only instructional practices but also the broader learning ecosystem. It is increasingly exerting its influence over both primary and tertiary education in the Philippines, where enduring issues with learning outcomes, educator workload, and inconsistent digital infrastructure continue to inform educational reform initiatives. Rather than being implemented as an isolated innovation, AI is permeating the Philippine educational landscape through adaptive learning systems, automated feedback mechanisms, learning analytics, and, more recently, generative AI applications that facilitate instructional design and enhance student learning.

Comprehensive global systematic reviews indicate that these technological advancements yield maximum efficacy when they augment pedagogical practices and enhance instructional decision-making processes, rather than supplanting human judgment (Chiu et al., 2023; Wang et al., 2024). Such findings hold particular significance within the Philippine educational context, where transformative efforts rely on tools that can effectively support differentiated instruction while remaining acutely aware of contextual limitations.

In the context of basic education in the Philippines, artificial intelligence is often conceptualized as a viable strategy to address issues of large class sizes and learner heterogeneity. Empirical evidence from both international and regional analyses suggests that formative assessments and individualized

practice, supported by artificial intelligence, can enhance educational outcomes when integrated into educator-led instructional frameworks (Wang et al., 2024). Nonetheless, scholarly investigations underscore that the efficacy of such initiatives depends on factors such as educators' preparedness, curriculum alignment, and the implementation of ethical safeguards, particularly in educational systems marked by disparities in access to connectivity and technological resources. These factors are particularly pertinent to the Philippine K–12 educational system, where there is significant variability in both technological infrastructure and educators' professional development opportunities.

In higher education, universities in the Philippines are experiencing an accelerated integration of AI-driven tools, notably generative AI used for academic writing, evaluative feedback, and content creation. Meta-analytic research substantiates that generative AI can yield beneficial outcomes in learning efficacy, student motivation, and engagement levels, contingent on its application being guided by pedagogical objectives and provisions for metacognitive oversight (Ma & Zhong, 2025; Xia et al., 2025). These findings are congruent with the overarching aims of higher education institutions in the Philippines, which seek to foster adaptable learning environments, enhance research productivity, and implement student-centered instructional methodologies. Notwithstanding, learners' perceptions remain paramount. Empirical investigations indicate that students' acceptance of AI technologies is shaped by factors such as transparency, perceived equity, and the clarity of the intended purpose, with adverse perceptions surfacing when AI is linked to surveillance practices or punitive assessment approaches (Rodway & Schepman, 2023).

Across sectors, research on AI in education has grown, focusing on intelligent tutoring, learning analytics, and conversational agents. Wang et al. (2024) observe that although AI interventions often provide higher engagement and performance, most initiatives do not extend beyond short-term pilots and are not pedagogically consistent in the long term. Building on this view, Mustafa et al. (2024) note in their meta-review of AI-in-education literature that the field is fragmented. It emphasizes that, while research mainly focuses on 'students' and higher education, there is a limited emphasis on teachers and school leaders, or inclusion-oriented applications. Their work highlights the need for a more structured and theory-based research agenda to inform AI integration.

On the other hand, critical pedagogy sees education as the practice of freedom, enabling students to question patterns of power, ideology, and oppression, despite the rapid expansion of AI in educational contexts. It brings tensions that frequently undercut these emancipatory ambitions. Recent work suggests that the deployment of AI systems is closely associated with a technocratic, efficiency-driven logic that both closes off and opens up space for critical questioning. For instance, Kasneci et al. (2023) show that AI technologies potentially stifle intellectual engagement by reinforcing shallow thinking patterns and generating reductive statements. These limit possibilities for dialogic, mediating, and transformative learning—dimensions that are core tenets of critical pedagogy. In addition, Holmberg and Fransson (2022) argue that AI-facilitated assessment and monitoring practices can be seen as contributing to the normalization of behavioral surveillance and governance in classroom settings. This augments the potential consequence of decreased student agency and democratic participation.

These concerns are compounded by the opacity of algorithms that determine what counts as valid knowledge or acceptable performance without giving learners or educators access to how these decisions are made. Current scholarship also argues that AI can be a focus of critique for critical pedagogy when it is intentionally repurposed. Atenas et al. (2025) recommend incorporating epistemic data justice frameworks into curricula. With this, learners can interrogate the functioning of AI systems, the values they represent, and how they contribute to social injustices. Complementing this, Veldhuis et al. (2025) put forward a framework to develop critical AI literacy skills that enable learners to analyze bias, transparency, and power relations in the context of AI technologies. When employed in this way, AI becomes a medium of conscientization rather than an instrument of conformity. In the end, the educational impact of AI will depend on whether institutions use it to reinforce efficiency or to encourage critical and justice-oriented inquiry.

Nevertheless, the rapid growth of AI in education raises a significant concern for current schooling. It also poses the risk of undermining critical pedagogy by deepening automation, pervasive data surveillance, and the subtle but powerful impact of algorithmic distortions. As AI shapes more of the

decisions educators make about teaching, assessment, and learner conduct, it risks also restricting pedagogical possibilities and devaluing dialogic, reflective, and justice-oriented practices that critical pedagogy demands.

These studies reveal that while artificial intelligence in education has been widely examined in terms of technological potential, pedagogical applications, and ethical concerns, the literature remains fragmented and uneven in addressing how power, governance, and equity are structured within policy frameworks. More importantly, there is limited interrogation of how these dynamics operate within specific national contexts such as the Philippines. Against this backdrop, Philippine education lacks critical policy analysis on AI reforms and their power implications. Current research primarily emphasizes technology while neglecting ideological, governance, and equity aspects. The absence of critical pedagogy in AI discussions overlooks the effects of automation on teacher autonomy and democratic engagement. Additionally, there is no rights-based framework for AI governance in the Philippines, despite the global call for ethical AI practices. This study seeks to fill these gaps through a critical policy analysis focused on educational justice and principles of human dignity, agency, transparency, equity, and accountability in AI integration.

Significantly, this study has global relevance for SDG 4, which advocates for inclusive and quality education. Ensuring AI fosters equitable learning is essential for enhancing educational quality and inclusivity. This necessitates policies that uphold learner autonomy, honor teacher expertise, and prioritize education as a venue for critical thought and engagement over automated efficiency.

Nevertheless, in response to these gaps, this study sought to answer questions that interrogate AI governance discourses, their implications for equity and educational agency, and the need for a justice-oriented framework aligned with SDG 4.

1. How is artificial intelligence framed and governed in Philippine education policies, and what dominant discourses and ideological assumptions shape this governance?
2. What implications do current AI-in-education policies have for equity, teacher autonomy, and learner agency in Philippine basic and higher education?

How can a rights-based and critical pedagogical governance framework realign AI integration in Philippine education with educational justice and SDG 4?

1.1 Literature Review

This study is anchored in two complementary theoretical traditions: critical pedagogy and critical policy analysis (CPA). Critical pedagogy, most notably advanced by Paulo Freire, conceptualizes education as a practice of freedom that enables learners to question systems of power, inequality, and oppression. As Freire (1970) emphasizes, education should cultivate critical consciousness through dialogue and reflection, positioning learners as active agents rather than passive recipients of knowledge. In the context of artificial intelligence (AI), this perspective raises important questions about whether technology enhances or constrains human agency and critical engagement. In parallel, CPA, as articulated by Stephen J. Ball, extends this critical lens by examining how policies are shaped by ideological struggles, power relations, and institutional interests rather than purely technical considerations. Ball (1993, 2003) argues that policies function both as texts and as practices that structure governance, accountability, and professional identities. When applied to AI in education, CPA enables a deeper interrogation of how policy discourses privilege efficiency, innovation, and performativity, often at the expense of equity and democratic participation. Taken together, these frameworks position AI not merely as a technological innovation but as a socio-political construct embedded in systems of power.

Building on this theoretical grounding, the global expansion of AI in education has been widely documented, with numerous studies emphasizing its transformative potential. For instance, Wang et al. (2024), in their systematic review, found that AI applications such as intelligent tutoring systems and adaptive learning platforms significantly enhance learner engagement and performance when aligned with pedagogical goals. Similarly, Garzón et al. (2025) reported that AI facilitates personalized learning and supports instructional efficiency across diverse educational settings. These findings reflect a techno-optimist perspective, which positions AI as a solution to persistent challenges such as learner diversity

and instructional scalability. However, this optimistic framing is increasingly challenged by critical scholarship. Perrotta and Pangrazio (2023) argue that the growing reliance on digital platforms reflects a process of datafication, wherein educational practices are reorganized around measurable outputs and performance indicators. From this perspective, AI risks reducing education to a technical enterprise focused on efficiency rather than holistic development. Thus, while techno-optimist accounts emphasize innovation and scalability, critical perspectives call attention to the need to interrogate the underlying assumptions and consequences of AI integration, highlighting a tension that runs throughout the literature.

This tension becomes even more pronounced when examining the ethical and equity implications of AI adoption. On the one hand, UNESCO (2021) frames AI as a tool for advancing inclusive and equitable education, particularly in relation to Sustainable Development Goal 4. On the other hand, empirical research demonstrates that AI systems often reproduce existing inequalities. For example, Akgun and Greenhow (2021) emphasize that algorithmic bias and unequal access to technology can disadvantage marginalized learners, while Garzón et al. (2025) further note that AI adoption tends to benefit institutions with greater resources and technical capacity. More importantly, this divergence reveals a gap between policy aspirations and implementation realities. While AI is frequently presented as a democratizing force, its benefits are often concentrated in well-resourced contexts, thereby reinforcing structural inequities. Consequently, ethical considerations must move beyond abstract principles to address how power, access, and resources shape AI deployment in education.

Closely related to these concerns are debates surrounding teacher autonomy and learner agency. From a techno-optimist standpoint, AI is seen as augmenting teaching by automating routine tasks and providing actionable insights, thereby enabling educators to focus on higher-order instructional practices. As Giannakos et al. (2024) explain, generative AI has the potential to enhance instructional design and support personalized learning when implemented responsibly. Nevertheless, this perspective is complicated by evidence suggesting that AI systems can also standardize decision-making and constrain professional judgment. At the same time, the implications for learners are equally significant. Williamson, Eynon, and Knox (2023) contend that AI systems are embedded within broader political and economic structures that shape knowledge production and access, often positioning learners as passive data subjects rather than active agents. However, Atenas et al. (2025) offer a countervailing perspective by advocating for critical AI literacy, which empowers both educators and learners to interrogate the assumptions and biases embedded in AI systems. In this sense, the impact of AI on agency is not fixed but contingent on how it is critically engaged within educational contexts.

These competing perspectives are also reflected in policy discourses, where AI is frequently framed as a driver of modernization, competitiveness, and system efficiency. Such narratives align with broader trends in digital governance that prioritize data-driven decision-making and measurable outcomes. However, as Ball (2003) argues, performative cultures in education reshape professional identities and institutional priorities by emphasizing quantifiable results over critical and reflective practice. In the context of AI, this technocratic orientation manifests in the increasing reliance on algorithmic systems to guide decisions related to assessment, resource allocation, and institutional performance. While these developments may enhance efficiency, they also concentrate power within technical and policy elites, thereby limiting democratic participation. Consequently, the governance of AI must be understood not only as a technical issue but also as a political and ethical concern that requires inclusive and participatory approaches.

When situated within the Philippine context, these global dynamics intersect with persistent challenges related to inequality and institutional capacity. Empirical studies by Barrot et al. (2021) and Talidong and Toquero (2021) demonstrate that digital transformation initiatives often exacerbate disparities between urban and rural schools, as well as between well-resourced and under-resourced institutions. These disparities are further compounded by limited internet connectivity, insufficient teacher training, and uneven institutional readiness, all of which constrain the effective adoption of AI technologies. At the same time, national and institutional policies increasingly promote AI as a tool for modernization and innovation. However, the benefits of such initiatives remain unevenly distributed, creating a tension between policy aspirations and implementation realities. Moreover, existing local

scholarship tends to focus on technological adoption and user perceptions, with less emphasis on the underlying power dynamics and ideological assumptions that shape AI governance.

2. Methods

2.1 Research Design

This research employs critical policy analysis (CPA) to frame educational policy as shaped by power dynamics and ideological factors rather than solely by technical considerations. CPA interprets policies as frameworks that influence governance, accountability, and identity within education (Ball, 1993). This approach examines AI-in-education policies as contentious spaces where diverse interpretations of quality education under SDG 4 are debated. The investigation assesses how discourses focused on efficiency and data may normalize surveillance and performative accountability, thereby limiting democratic and humanistic objectives. Drawing on Ball's (2003) critique, CPA examines whether AI-driven reforms compromise learner autonomy, teacher professionalism, and critical pedagogical practices.

The PRISMA flow diagram is also used in this study. Although it is traditionally used for systematic reviews of empirical studies, its structured approach to identification, screening, eligibility, and inclusion provides a transparent and replicable framework for document selection. In this study, PRISMA was adapted to suit policy analysis by treating policy documents as units of analysis rather than empirical studies. This adaptation ensured systematic filtering, minimized selection bias, and enhanced transparency in constructing the policy corpus. Specifically, adaptations included the use of policy relevance, institutional authority, and alignment with AI governance as primary inclusion criteria, rather than methodological quality alone.

2.2 Methodological Process

This study reviewed AI policy documents in Philippine education through well-defined inclusion and exclusion criteria to ensure methodological rigor. It also focused only on texts that had a direct impact on the governance of AI and implications for critical pedagogy. The corpus included documents published during 2021-2025 from authoritative bodies that had a material impact on teaching, equity, and governance. Documents without credible authorship or policy relevance, without any actual AI-specific content, were removed. The research utilized a multi-layered analytical structure of critical discourse analysis and ideological critique within critical policy analysis, underpinned by the PRISMA methodology for selection transparency. To ensure analytical rigor, the study employed the Critical Appraisal Skills Programme (CASP) Qualitative Checklist to assess the credibility, relevance, and clarity of the selected policy documents. The checklist was adapted to evaluate policy texts based on criteria such as clarity of purpose, coherence of arguments, transparency of assumptions, and relevance to AI governance and educational equity. This approach enabled a systematic evaluation of policy quality while remaining sensitive to the interpretive nature of policy analysis.

In addition, to ensure analytical validity, themes underwent an iterative validation process involving comprehensive document review, internal consistency checks of coded segments, and alignment with critical policy analysis. Theme refinement was implemented in response to shifts in coding patterns, the emergence of contradictions, and the revelation of deeper ideological assumptions, thereby enhancing coherence and analytical rigor. An audit trail was established via analytic memos and coding documentation to improve dependability, ensuring transparency in the development and refinement of interpretations and themes.

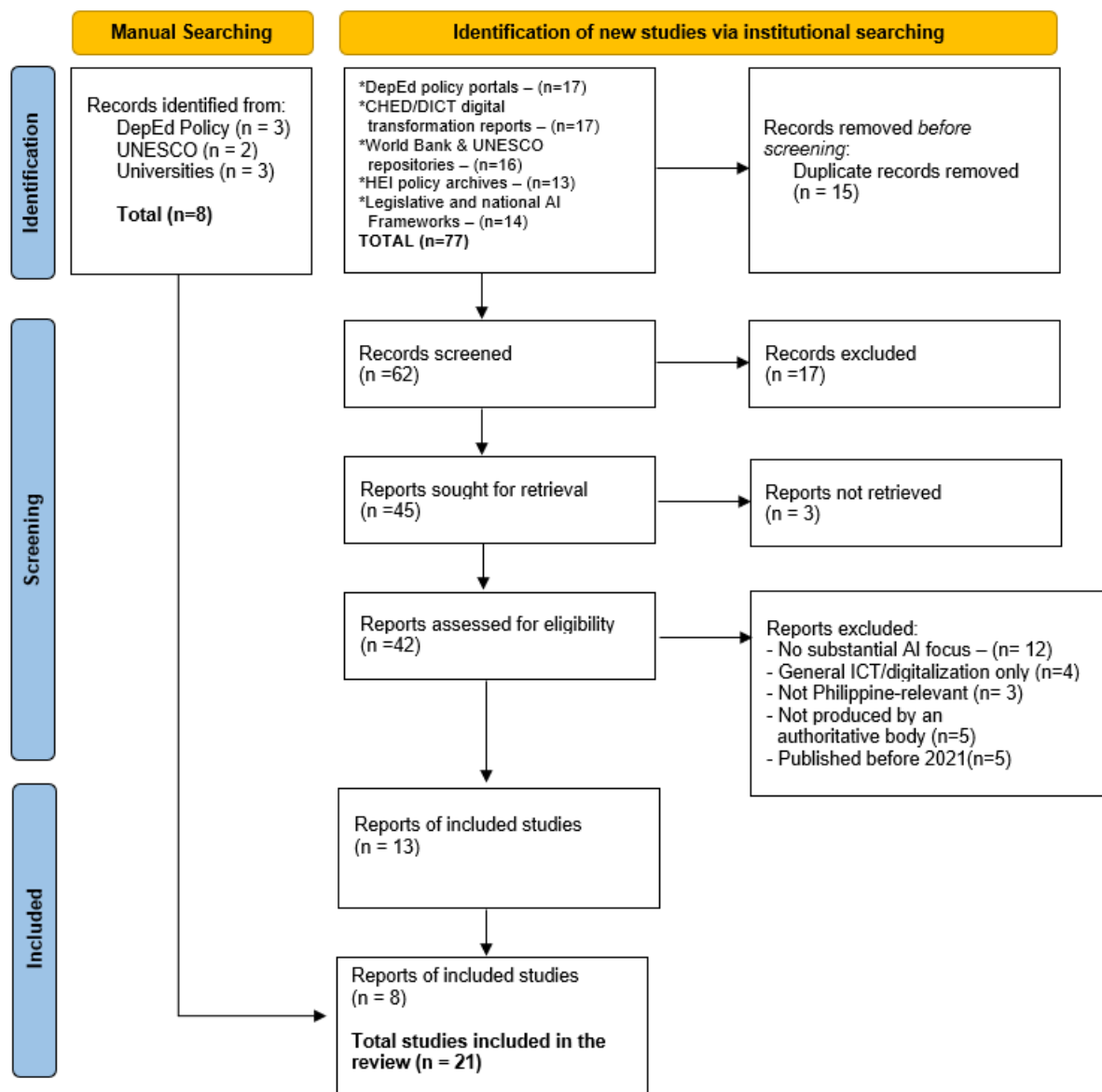
Moreover, using discourse tracing, thematic coding, ideological critique, and critical policy reading, the paper systematically showed how power relations and ideologies evident in policy narratives informed the integration of AI into Philippine education. Collectively, these approaches yielded a coherent and rigorous analytic foundation for delineating the broader effects of AI-mediated reforms on equity, governance, and democratic educational practice.

Figure 1 presents the PRISMA diagram of the processes of identification, screening, and inclusion used to determine which AI-related policy documents were eligible for review. It systematically shows how the criteria were used to screen the initial records and select the relevant studies for the review.

3. Results and Discussion

The PRISMA flow diagram illustrates a systematic and transparent process of document selection (See Figure 1). A total of 85 records were initially identified, comprising 8 documents from manual searches and 77 from institutional databases. After removing duplicates, 62 records remained for screening. Of these, 17 were excluded based on title and abstract review due to lack of relevance to AI governance in education. The remaining 45 documents were assessed for eligibility, during which 3 records were excluded due to inaccessibility. Of the 42 full-text documents reviewed, 34 were excluded for lacking substantive AI focus, 5 for limited applicability to the Philippine context, and 3 for insufficient policy authority. This process resulted in a final corpus of 21 policy documents included in the analysis.

Figure 1
PRISMA flow diagram of Document Selection



3.1 Thematic Coding of Policy Statements

Table 2 synthesizes principal AI-in-education policies in Philippine basic and higher education, including governance, institutional frameworks, and global benchmarks. It elucidates each policy's stance on AI concerning pedagogy, equity, ethical governance, and the progressive digital transformation agenda.

Table 2 illustrates the framing of AI in education across various levels in the Philippines. The DepEd's New Normal Policy prioritizes digital transformation, ethical AI application, and infrastructure advancements through initiatives like Digital Rise and E-CAIR. CHED and the World Bank advocate for AI's role in learning analytics, adaptable delivery, and bridging educational disparities. National programs, including the DOST AI Program and a proposed AI Code, emphasize the importance of ethical governance. Universities interpret these principles into their institutional values concerning transparency and academic integrity, informed by UNESCO and OECD guidelines.

Table 2

AI in Education Policy Table

Category	Policy/Framework	Issuing Body	Statement of AI-Policy in Education
Philippine Basic Education	DepEd Digital Rise Program (2022–present)	Department of Education	Integrates AI and emerging technologies in "Education 4.0" via digital infrastructure, coding, productivity tools, and LMS integration to enhance ICT-supported pedagogy and learner competencies.
Philippine Basic Education	DepEd Order No. 013, s. 2025 (SEAMEO AI Partnership)	Department of Education	Directs DepEd to leverage AI to improve educational frameworks and foster collaboration, indicating a national commitment to systematic AI integration.
Philippine Basic Education	IDE 2025 Statement: "AI and Education – Preserving Human Agency" (2025)	Department of Education	Advocates for ethical AI use in education stress that generative AI should enhance, not undermine, the roles of teachers and learners.
Philippine Basic Education	Education Center for AI Research (E-CAIR) Launch (2025)	Department of Education	Establishes a national center for AI-driven curriculum, assessment, and system management aligned with the DepEd Reform Agenda.
Philippine Basic Education (Regional)	DepEd NCR Regional Memorandum No. 1289, s. 2024	Department of Education – NCR	Establishes a framework for AI integration in education and management to equip stakeholders for a digital future.
Philippine Higher Education/ National Policy	Digital Transformation of Philippine Higher Education (2022)	World Bank / CHED	Advocates for the integration of AI, analytics, and automation in curriculum development, adaptive learning, administrative optimization, and mitigating digital disparities in higher education institutions.
Philippine Higher Education	CHED Digital Integration Initiatives (2021–2025)	CHED / DICT	Advocates for integrating digital instruction, data systems, and AI skills into the modernization of Philippine higher education.
National AI Strategy	Philippine Artificial Intelligence Program Framework (to 2028)	DOST	Prioritizes AI for national development, fostering education via AI-driven innovation, inclusive governance, and rights-based implementation.
National Legislation (Proposed)	House Bill No. 57: National AI Code of the Philippines (2025)	House of Representatives	Mandates the development of ethical and accountable AI frameworks in education by DepEd and CHED.
Philippine Higher Education	UNESCO–CHED AI Readiness Assessment for Philippine HEIs (2023/24)	UNESCO	Evaluates institutional readiness and promotes the incorporation of AI literacy, ethical AI governance, and capability enhancement within higher education institutions.

Institutional (State University)	UP System Principles for Responsible AI (2022–2023)	University of the Philippines	Establishes 11 principles that mandate transparency, fairness, human oversight, and safeguarding of learner rights in AI-enhanced education and evaluation.
Institutional (Open University)	UPOU Guidelines on the Use of AI in Teaching and Learning (CMDPB Memo 2024-001)	University of the Philippines Open University	Establishes guidelines for generative AI use in online education, mandates transparency, and emphasizes coherence with educational objectives and ethical standards.
Institutional (Private HEI – Catholic)	Generative AI in Education Policy (AY 2025–2026)	De La Salle University	Outlines a responsible GenAI application in education, emphasizing transparency, integrity, and ethical alignment.
Institutional (State University)	MSU-IIT AI Policy (BOR Resolution No. 11, s. 2024)	MSU-IIT	Establishes frameworks for ethical AI governance in academia.
Institutional (Private HEI)	Silliman University AI Integration Framework (2023–2024)	Silliman University	Offers frameworks for ethical AI implementation in educational settings, prioritizing community engagement.
Institutional (Private HEI)	LPU-Cavite Policy on AI Use for Student Outputs (2023/24)	Lyceum of the Philippines University–Cavite	Regulates student use of AI tools, defining appropriate applications and accountability for academic integrity.
Institutional (Private HEI)	FEU Academic Integrity and Generative AI Guidelines (2024)	Far Eastern University	Integrates GenAI regulations into academic integrity frameworks, highlighting attribution, ethical limits, and student literacy.
Global Framework	UNESCO Recommendation on the Ethics of AI (2021)	UNESCO	Establishes global ethical standards for fairness, transparency, data governance, and human-centric AI in education.
Global Framework	UNESCO Guidance for Generative AI in Education and Research (2023)	UNESCO	Advocates for suitable GenAI usage based on age, development of teacher competencies, ethical incorporation in evaluations, and robust data management frameworks.
Global Framework	OECD Digital Education Outlook (2023)	OECD	Identifies risks and governance in AI-enhanced education, emphasizing algorithmic transparency and bias reduction.
Global Framework	UNESCO AI in Education Orientation (2020s)	UNESCO	Frames AI as a mechanism for promoting SDG 4, emphasizing inclusivity, protecting rights, and guaranteeing equitable advantages of AI-driven education.

3.2 Thematic Clusters from the AI-in-Education Policy Analysis

Figure 2 shows the generated themes from clustered policy statements that depict the conceptualization and implementation of AI in Philippine and global education frameworks. These themes facilitate understanding of the priorities, assumptions, and governance strategies that influence AI integration in both basic and higher education.

3.3 Discussion

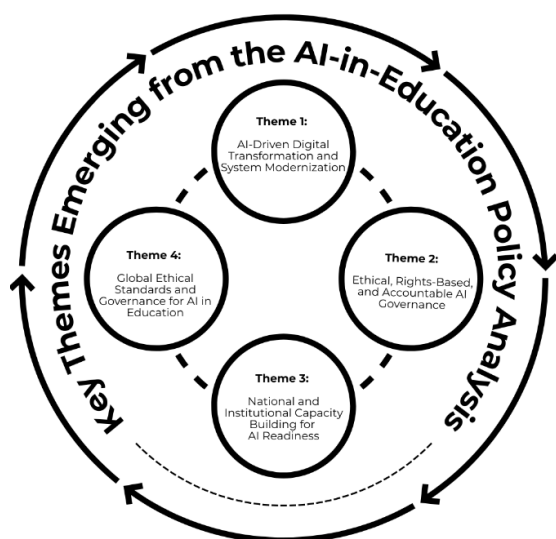
3.3.1 Policy Analysis

Critical Policy Analysis (CPA) indicates that global automation discourses significantly shape the formulation of educational AI policies in the Philippines. This analysis employs a complex analytical framework to examine the creation of meaning and the dissemination of power in these policies.

Utilizing discourse tracing, thematic coding, ideological critique, and CPA, the analysis identifies underlying assumptions affecting equity, governance, and educational agency.

Figure 2

Thematic Clusters from the AI-in-Education Policy Analysis



3.3.2 Discourse Tracing

Across the Department of Education (DepEd), the Commission on Higher Education (CHED), the World Bank, and institutional policies, AI is consistently framed as vital to broader digital transformation agendas. Their emphasis on governance, ethics, and equity differs between these entities. At the basic education level, DepEd integrates AI into broader information and communication technology (ICT) initiatives, emphasizing digital content, connectivity, and 21st-century skills. However, evidence from the Philippines suggests that technology-focused reform efforts exacerbate inequality between well-resourced and poorly-resourced schools. This means that without robust equity policies built into the AI or used in conjunction with it, it may continue to perpetuate these oppressive biases (Barrot et al., 2021; Talidong & Toquero, 2021).

In higher education, the convergence of CHED with the World Bank's digital transformation agenda positions analytics and AI as powerful resources for enabling greater flexibility, system efficiency, and institutional resilience. Nevertheless, international and Philippine studies suggest that the attainment of these dividends is conditioned on a strong governance framework. Studies show that, in the absence of adequate digital readiness, training, and organizational support, higher education institutions (HEIs) face challenges when adopting advanced technologies. Therefore, these activity lines also become ethical, fair, and sustainable (Irie, 2021; Villaseñor, 2024). Notwithstanding, there are global reviews that argue the narrow focus on efficiency is likely to hide important issues regarding bias, transparency, and preservation of teacher and learner agency (Wang et al., 2024; Garzón et al., 2025; Mustafa et al., 2024).

At the institutional level, institutions such as the University of the Philippines and Far Eastern University (FEU) take a more normative position, focusing on responsible use, transparency, and accountability. These positions are consistent with global discourses on the (in)equities inherent in the need for AI in education to be underpinned by human-centered rather than merely technical or performative concerns. Academics argue that meaningful human oversight, strong privacy rights, and equity-based design are key to avoiding AI that reinforces systemic inequities (Holstein & Doroudi, 2021; Karpouzis, 2024). Evidence from the Philippines also bolsters this argument, showing that new technologies tend to exacerbate inequalities unless institutions put in place tight controls and an inclusive framework.

Hence, national agency discourse positions AI as a potential modernizing tool, whereas empirical investigations and university-level policies emphasize the necessity for ethical, equity-driven oversight. Connecting these disparate viewpoints requires creating a national AI-in-education framework that puts

values such as trust, support for educators, data governance, and inclusive design at its core, alongside building digital infrastructure so that AI supports rather than undermines the right to quality education.

Across Philippine education, from basic to higher learning, artificial intelligence (AI) is framed as part of the exciting start to digital transformation and the responsible governance of this domain. DepEd and CHED policies highlight AI's ability to augment digital infrastructure, facilitate administrative processes, and enhance adaptive learning. This aligns with global outcomes that AI can be used to improve educational pathways and organizational processes (Wang et al., 2024; Garzón et al., 2025). However, evidence from the Philippines suggests that digital transformation tends to further entrench existing inequalities in contexts where technological development outpaces local capacity. This implies that AI may also exacerbate disparities unless targeted efforts to address equity are implemented (Barrot et al., 2021). This tension highlights the importance of system preparedness, including infrastructure, teacher preparation, and organizational leadership. This study confirmed that unprepared institutions struggle to embed AI effectively, highlighting the importance of continuing professional development and governance structures (Mustafa et al., 2024; Irie, 2021).

Apart from modernization, ethical governance and rights-based community care are core issues. Researchers posit that AI should transmit some of its principles of openness, fairness, and the adequacy of human oversight to address bias, protect learner autonomy, and foster trust (Holstein & Doroudi, 2021; Karpouzis, 2024). Those discoveries are consistent with Philippine institutional policies that require transparency on how AI will be used, articulate academic integrity procedures, and prioritize student rights. This directional focus is supported by global frameworks, such as UNESCO's ethical guidelines, which position AI governance as a critical means of protecting equity and align with the vision of inclusive-quality education articulated in Sustainable Development Goal 4 (SDG 4).

An integrative analysis of these narratives unveils a consistent narrative. It shows that while national-level policies have incentivized the view that AI should play a boosting role in modernization, empirical analyses and global frameworks emphasize that technological integration must be both ethically motivated and capacity-driven. AI adoption that really works requires robust digital policies, effective governance systems, strengthened teacher capacity, and protection of students. These views indicate that the Philippines needs to moderate innovation with responsibility and ensure that AI promotes educational equity and human agency.

3.3.3 Ideological Analysis

Philippine AI-in-education policy promotes a modernization narrative that views digitalization and automation as inherently beneficial. AI is presented as central to "Education 4.0," which emphasizes efficiency and competitiveness, reflecting global trends towards datafication in education (Williamson et al., 2023; Perrotta & Pangrazio, 2023). This aligns with Pischetola's (2021) concept of technosolutionism, which frames systemic issues as solvable through technology rather than through social reforms. Consequently, modernization is perceived as an inevitable path, constraining the conception of "good" education.

Similarly, governance reflects a managerial paradigm. AI is utilized for prediction and optimization, corresponding with emerging "AI governance" frameworks that concentrate authority within expert bodies (Filgueiras, 2024). Discussions on ethics often result in compliance measures that prioritize institutional risk management over public engagement. Research indicates governance is increasingly reliant on opaque systems, altering authority dynamics through data-driven metrics (Perrotta & Pangrazio, 2023; Williamson et al., 2023). This promotes a technocratic system in which data governs educational stakeholders rather than fostering collaborative educational development.

Learners are generally viewed as beneficiaries of AI efficiencies, rather than as active political agents. They are often reduced to data points or potential workers, lacking the agency to shape AI's role in education. Scholarship on AI literacy argues for empowering learners to critically engage with data systems rather than viewing them as passive recipients (Atenas et al., 2025). Additionally, critical perspectives emphasize that educational technologies embody specific values and power dynamics, and overlooking these aspects perpetuates existing socio-economic interests (Pischetola, 2021).

This body of literature reveals the ideological foundations of AI-in-education policies, which equate modernization with automation, governance with institutional oversight, and learners with managed entities. Failing to challenge these premises may exacerbate existing power disparities rather than promote equitable and inclusive educational futures.

3.3.4 Critical Policy Interpretation

AI education reforms in the Philippines redistribute benefits and power within the system. A critical analysis shows that well-resourced institutions benefit most from AI, leveraging it for efficiency and research productivity. Technology providers and policymakers gain influence as AI shapes educational priorities and resource distribution. Their vision of innovation limits the scope of legitimate reforms. Marginalized groups, including rural learners and underfunded schools, face increased risks under AI reforms. AI systems favor those who align with dominant cultural norms, potentially exacerbating inequities for disadvantaged students. Teachers may also be marginalized as AI reduces their autonomy and heightens performance expectations.

Centralized structures and partnerships govern power dynamics in these reforms. Decision-making becomes technocratic, favoring policymakers and data scientists over educators. AI reframes educational issues as technical challenges, diminishing democratic discourse on educational priorities. This shift alters school relationships, positioning learners as data subjects and teachers as implementers, while administrators depend on algorithms for decision-making. Such changes indicate that AI not only alters pedagogy but also redistributes control over knowledge and policy.

3.3.5 Emergent Framework

Figure 3 presents the ARISE Framework (AI Rights-based Inclusive Systems for Education) is proposed as a normative and analytical response to the key tensions identified in the policy analysis, particularly the dominance of efficiency-driven narratives, the insufficient attention to equity, and the marginalization of teacher and learner agency. While current AI-in-education policies in the Philippine context frequently frame technological integration as a pathway to modernization and system optimization, the findings of this study reveal that such framings often obscure deeper concerns related to power, access, and democratic participation. In response, the ARISE Framework seeks to reorient AI governance toward principles of educational justice, aligning explicitly with the core mandate of Sustainable Development Goal 4 (SDG 4), which calls for inclusive, equitable, and quality education for all. By foregrounding inclusivity, human dignity, and critical engagement, the framework ensures that AI integration supports—not undermines—the transformative vision of SDG 4.

The framework is composed of four interrelated domains that collectively address the limitations identified in existing policy discourses. First, Structural Conditions emphasize the need for equitable infrastructure, institutional readiness, and participatory policymaking. This domain directly supports SDG 4's commitment to equitable access by addressing disparities in digital resources and institutional capacity. The findings reveal that policy documents often assume uniform readiness across institutions, thereby overlooking structural inequalities. By embedding equity considerations into infrastructure development and policy design, this domain ensures that AI contributes to expanding access rather than reinforcing exclusion.

Second, Pedagogical Agency highlights the importance of preserving teacher autonomy and learner rights in AI-mediated environments, which is central to SDG 4's emphasis on quality education. The analysis indicates that current policies tend to position educators and students as passive recipients of technological systems. In contrast, this domain promotes active participation, critical thinking, and meaningful engagement, ensuring that AI enhances rather than diminishes the human dimensions of teaching and learning. In doing so, it reinforces SDG 4's goal of fostering relevant and empowering educational experiences.

Third, Critical Governance focuses on interrogating the power relations embedded in AI systems and policy frameworks. The study finds that many policies adopt a technocratic orientation, prioritizing efficiency and data-driven decision-making while neglecting transparency and democratic participation. This domain aligns with SDG 4's broader vision of inclusive and participatory education systems by advocating for governance structures that are transparent, accountable, and responsive to diverse

stakeholders. By challenging technocratic dominance, it ensures that AI governance remains socially grounded and ethically informed.

Figure 3
AI-Rights Inclusive Systems for Education (RISE) Framework



Finally, Rights-based Oversight ensures that ethical principles are operationalized through concrete accountability mechanisms. This includes transparency requirements, algorithmic audits, data protection policies, and continuous evaluation of AI's impact on equity and inclusion. These mechanisms are essential for safeguarding the rights of learners and educators, directly supporting SDG 4's commitment to safe, inclusive, and effective learning environments. Without such safeguards, AI risks undermining the very principles that SDG 4 seeks to advance.

These four domains are mutually reinforcing. Achieving SDG 4 in the context of AI integration requires simultaneous attention to access, quality, governance, and accountability. For example, expanding digital infrastructure without ensuring ethical oversight may reproduce inequities, while promoting pedagogical innovation without addressing structural constraints may limit impact. As such, the ARISE Framework functions as both an evaluative lens and a strategic guide for aligning AI policies with the multidimensional goals of SDG 4.

Operationally, the framework can inform policy and practice across multiple levels. At the national level, it can guide the development of AI strategies that incorporate equity impact assessments and participatory governance mechanisms. At the institutional level, it can support the design of policies that promote transparency, protect data rights, and strengthen teacher capacity in critical AI literacy. Through

these applications, the ARISE Framework provides a concrete pathway for ensuring that AI integration advances the inclusive and transformative agenda of SDG 4 rather than reinforcing existing inequalities.

3.3.6 Limitations of the Study

This study is subject to several limitations. First, the analysis is confined to policy documents within the Philippine context, which may limit the generalizability of the findings to other national settings with different governance structures and levels of technological development. Second, the study focuses on policy texts rather than their actual implementation, and therefore does not capture how AI policies are enacted, interpreted, or experienced in practice by educators and learners. Third, the use of critical policy analysis and ideological critique, while valuable for uncovering power relations and assumptions, involves an interpretive dimension that may introduce researcher subjectivity despite efforts to ensure analytical rigor. Finally, the document corpus is limited to publicly available and accessible materials, which may exclude informal or emerging policy discourses. These limitations suggest the need for future research that incorporates empirical investigations, such as case studies, interviews, and longitudinal analyses, to examine how AI governance unfolds in real educational contexts.

4. Conclusion

This study demonstrates that AI governance in Philippine education is predominantly shaped by modernization narratives that prioritize efficiency, data-driven decision-making, and system optimization. Through critical policy analysis, the findings reveal that such orientations risk reinforcing structural inequalities, marginalizing teacher autonomy, and limiting learner agency, despite increasing references to ethics and fairness in policy discourse. By examining the ideological underpinnings of these policies, the study highlights how AI integration is not merely a technical process but a deeply political one that redistributes power within educational systems. The proposed ARISE Framework contributes to existing scholarship by offering a theoretically grounded and justice-oriented lens for rethinking AI governance in alignment with the principles of inclusive and equitable education under SDG 4. In doing so, the study underscores the importance of critically engaging with emerging technologies to ensure that they support, rather than undermine, democratic and human-centered educational goals.

The findings of this study have several implications for policy and practice. First, national agencies such as DepEd and CHED should institutionalize equity impact assessments to ensure that AI initiatives do not exacerbate disparities between well-resourced and under-resourced schools. Second, governance structures must be strengthened through the creation of inclusive AI oversight committees composed of educators, students, policymakers, and community stakeholders to promote participatory decision-making. Third, transparency requirements should be enforced, including the disclosure of algorithmic processes, data usage, and potential biases, alongside regular ethical audits of AI systems. Fourth, sustained investment in infrastructure, teacher training, and localized resources is necessary to support equitable implementation, particularly in underserved areas. Finally, the integration of critical AI literacy across educational levels is essential to empower both teachers and learners to engage with AI technologies in informed and reflective ways. Collectively, these measures can help align AI governance with the principles of educational justice and the broader goals of SDG 4. Nevertheless, future research may extend this study by examining how AI policies are enacted in practice through classroom-based studies, institutional case analyses, and longitudinal designs that capture evolving governance dynamics.

Funding

The author declares that no funding was received during the study.

Conflict of Interest

There was no competing interest in the conduct of this study.

Declaration of AI Use

The research employed advanced language processing models, including ChatGPT (OpenAI), QuillBot, and SciSpace, to refine, elucidate, and elevate the quality of the written work. These tools facilitated the organization, coherence, and readability of the manuscript, thereby ensuring adherence to academic standards while preserving the author's original insights and arguments. References were managed and formatted using a reference management tool (e.g., Zotero) and manually cross-checked to ensure compliance with APA 7th edition guidelines.

References

- Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K–12 settings. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/s43681-021-00096-7>
- Atenas, J., Havemann, L., & Nerantzi, C. (2025). Critical and creative pedagogies for artificial intelligence and data literacy: An epistemic data justice approach for academic practice. *Research in Learning Technology*, 32. <https://doi.org/10.25304/rlt.v32.3296>
- Ball, S. J. (1993). What is policy? Texts, trajectories, and toolboxes. *Discourse: Studies in the Cultural Politics of Education*, 13(2), 10–17. <https://doi.org/10.1080/0159630930130202>
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228. <https://doi.org/10.1080/0268093022000043065>
- Barrot, J. S., Llenares, I. I., & del Rosario, L. S. (2021). Students' online learning challenges during the pandemic in a developing country. *Education and Information Technologies*, 26(6). <https://doi.org/10.1007/s10639-021-10589-x>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Filgueiras, F. (2024). Artificial intelligence and education governance. *Education, Citizenship and Social Justice*, 19(3), 349–361. <https://doi.org/10.1177/17461979231160674>
- Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction*, 9(8), 84. <https://doi.org/10.3390/mti9080084>
- Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernández-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative artificial intelligence in education. *Behaviour & Information Technology*. Advance online publication. <https://doi.org/10.1080/0144929X.2024.2394886>
- Giray, L., De Silos, P. Y., Adornado, A., Buelo, R. J. V., Galas, E., Reyes-Chua, E., & Ulanday, L. (2024). Use and impact of artificial intelligence in Philippine higher education: Reflections from instructors and administrators. *Internet Reference Services Quarterly*, 28(3), 315–338. <https://doi.org/10.1080/10875301.2024.2352746>
- Holstein, K., & Doroudi, S. (2021). Equity and artificial intelligence in education. *arXiv*. <https://doi.org/10.48550/arXiv.2104.12920>
- House of Representatives. (2025). *House Bill No. 57: National AI Code of the Philippines*. Congress of the Philippines.
- Irie, M. A. C. (2021). E-learning as a key player of Philippine education in light of COVID-19. *ICoESE 2021 Proceedings*. <https://doi.org/10.1145/3450148.3450149>
- Karpouzis, K. (2024). Artificial intelligence in education: Ethical considerations. *arXiv*. <https://doi.org/10.48550/arXiv.2409.15296>
- Ma, N., & Zhong, Z. (2025). A meta-analysis of generative AI on learning outcomes. *Journal of Computer-Assisted Learning*. Advance online publication. <https://doi.org/10.1111/jcal.70117>
- Mustafa, M. Y., Tlili, A., Lampropoulos, G., Chen, N.-S., & Burgos, D. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments*, 11, 59. <https://doi.org/10.1186/s40561-024-00350-5>

- Perrotta, C., & Pangrazio, L. (2023). The critical study of digital platforms and infrastructures: New directions for education technology research. *Education Policy Analysis Archives*, 31, 1–20. <https://doi.org/10.14507/epaa.31.7952>
- Pischetola, M. (2021). Re-imagining digital technology in education through critical and neo-materialist insights. *Digital Education Review*, 40(2), 154–171. <https://doi.org/10.1344/der.2021.40.154-171>
- Rodway, P., & Schepman, A. (2023). The impact of adopting AI educational technologies on projected course satisfaction in university students. *Computers and Education: Artificial Intelligence*, 5, 100150. <https://doi.org/10.1016/j.caeai.2023.100150>
- Talidong, K. J. B., & Toquero, C. M. D. (2021). Emergency online education anchored in Khan's framework. *European Journal of Interactive Multimedia and Education*, 2(1). <https://doi.org/10.30935/ejimed/9683>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. <https://doi.org/10.54675/PCSP7359>
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252, 124167. <https://doi.org/10.1016/j.eswa.2024.124167>
- 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., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation, and datafication in education. *Learning, Media and Technology*, 48(1), 1–5. <https://doi.org/10.1080/17439884.2023.2167830>
- Williamson, B., Eynon, R., & Knox, J. (2023). Critical perspectives on AI in education: Political economy, discrimination, commercialization, governance, and ethics. In W. Holmes et al. (Eds.), *Handbook of Artificial Intelligence in Education* (pp. 553–570). Edward Elgar. <https://doi.org/10.4337/9781800375413.00037>
- Xia, Q., Li, W., Yang, Y., Weng, X., & Chiu, T. K. F. (2025). A systematic review and meta-analysis of the effectiveness of generative artificial intelligence on students' motivation and engagement. *Computers and Education: Artificial Intelligence*, 9, 100455. <https://doi.org/10.1016/j.caeai.2025.100455>

Biography

John Mark Navarette Saldivar is an Assistant Professor 2 at La Salle University–Ozamiz City, where he also serves as Chair of the University Research Ethics Board and faculty member of the School of Graduate Studies. He is currently completing his Doctor of Philosophy in Education major in Research and Evaluation at Cebu Normal University. With a strong background in English education, research ethics, and teacher education, his scholarly work spans leadership, resilience, gamification, constructivism, values education, and educational technology. He has published extensively in international peer-reviewed journals and has received multiple Best Paper and Best Research awards in national and international conferences. Mr. Saldivar is actively involved in academic service as an editorial manager, editorial board member, scientific committee member, and journal peer reviewer across reputable international and Scopus-indexed journals, including *Teaching and Teacher Education*, *Frontiers in Education*, and *Frontiers in Psychology*. He holds professional affiliations with PAFTE Region 10, the National Research Council of the Philippines, UNESCO-APNIEVE, and several international research organizations.