



Generative AI and Social Inequality in Education: A Human-Centered Social Science Review of Global and Indonesian Evidence

Siti Nur Aisyah¹, Muhammad Alif Pratama², Dewi Anggraini³

^{1,2,3} Universitas Negeri Malang, Indonesia

Corresponding Email : siti.nur.aisyah@students.um.ac.id

ARTICLE INFO

Article history:

18 August 2026

Received in revised form

19 October 2026

Accepted 19 October 2026

Available online 20 October 2026

Keywords:

generative AI, social inequality, education, digital divide, AI literacy, human-centered AI.

Kata Kunci:

AI generatif, ketimpangan sosial, pendidikan, kesenjangan digital, literasi AI, AI berpusat pada manusia.

ABSTRACT

This review examined how generative artificial intelligence (AI) can widen educational opportunity while also reproducing social inequality. An integrative review synthesized 24 sources: 17 Indonesian publications (2022–2026), identified through national journal portals and Google Scholar, and seven international policy documents and scholarly works (2016–2023) selected as conceptual anchors. Thematic analysis produced four themes: unequal access, unequal digital and AI literacy, unequal capacity to benefit from AI-supported learning, and governance risks involving privacy, bias, academic integrity, and human agency. The reported benefits of generative AI, such as tutoring, writing support, and translation, appear to depend on connectivity, devices, language resources, teacher competence, and students' ability to evaluate machine output. The Indonesian evidence is mostly cross-sectional, self-reported, or based on small single-institution samples, with little evidence on learning outcomes. One national household analysis of school participation (not AI use) found no significant difference between 3T (disadvantaged, frontier, and outermost) and matched non-3T households, suggesting that regional labels may be an imperfect proxy for household conditions. The review proposes a cumulative model of AI inequality that extends van Dijk's sequence of digital inequality with a governance stage and a feedback loop. The model is a conceptual proposition, not a tested result, and recommendations are ordered by the strength of supporting evidence.

ABSTRAK

Kajian ini menelaah bagaimana kecerdasan buatan (AI) generatif dapat memperluas kesempatan pendidikan sekaligus mereproduksi ketimpangan sosial. Tinjauan integratif ini mensintesis 24 sumber: 17 publikasi Indonesia (2022–2026) yang dicari melalui portal jurnal nasional dan Google Scholar, serta tujuh dokumen kebijakan dan karya ilmiah internasional (2016–2023) yang dipilih sebagai jangkar konseptual. Analisis tematik menghasilkan empat tema: ketimpangan akses, ketimpangan literasi digital dan AI, ketimpangan kapasitas memperoleh manfaat dari pembelajaran berbantuan AI, serta risiko tata kelola yang menyangkut privasi, bias, integritas akademik, dan kendali manusia. Manfaat AI generatif yang dilaporkan, seperti bimbingan belajar, bantuan menulis, dan terjemahan, tampak bergantung pada konektivitas, perangkat, sumber daya bahasa, kompetensi guru, dan kemampuan peserta didik menilai keluaran mesin. Bukti Indonesia sebagian besar bersifat potong lintang, dilaporkan sendiri, atau berasal dari sampel kecil pada satu institusi, dengan sedikit bukti tentang hasil belajar. Satu analisis rumah tangga nasional tentang partisipasi sekolah (bukan penggunaan AI) tidak menemukan perbedaan signifikan antara rumah tangga 3T (tertinggal, terdepan, dan terluar) dan non-3T yang dipadankan, sehingga label wilayah mungkin kurang tepat sebagai proksi kondisi rumah tangga. Kajian ini mengusulkan model kumulatif ketimpangan AI yang memperluas urutan ketimpangan digital van Dijk dengan tahap tata kelola dan umpan balik. Model ini merupakan proposisi konseptual, bukan hasil yang telah diuji, dan rekomendasi disusun menurut kekuatan bukti pendukung.

This is an open access article under the [CC BY-SA](#) license.



1. INTRODUCTION

Generative artificial intelligence (AI) reached classrooms faster than the rules, resources, and skills needed to use it fairly. Its arrival has changed public debate about teaching, learning, assessment, and knowledge production. Systems that generate text, translations, explanations, and other content can offer immediate assistance to learners and educators, and they have raised hopes for personalized support and broader access to educational resources (Holmes et al., 2019; Luckin et al., 2016). Indonesian higher education illustrates how quickly this shift has occurred. A survey of 430 respondents (119 lecturers and 311 students) found that awareness of ChatGPT was already high, that lecturers were more likely than students to know and use it, and that a social system governing its academic use had not yet formed (Niyu et al., 2024). At the same time, the diffusion of generative AI has raised a central social question: will these technologies reduce educational inequality, or will they reproduce it in new forms?

The social consequences of educational technology are seldom determined by the technology alone. Benefits depend on the resources available to users and institutions, including reliable internet, devices, supportive learning environments, language resources, prior knowledge, and guidance. Van Dijk (2020) argued that digital inequality unfolds through successive kinds of access, from motivation and material access to skills and usage, and finally to unequal outcomes. Critical scholars likewise caution that educational AI is embedded in institutional and commercial arrangements that shape who benefits (Selwyn, 2019), and UNESCO's (2023) Global Education Monitoring Report asked on whose terms educational technology is being introduced. For generative AI specifically, UNESCO's guidance calls for a human-centered approach that protects privacy, equity, human agency, and cultural and linguistic diversity (Miao & Holmes, 2023).

Educational inequality can therefore emerge at three levels. The first concerns access to tools; the second, the ability to use them effectively; and the third, the capacity to evaluate and challenge their outputs. A student who can open an AI tool but cannot detect hallucinated or biased responses may not obtain the same educational value as a student with stronger critical and disciplinary knowledge.

Indonesia offers an instructive setting for examining these levels. Studies of digital education describe a persistent gap between urban areas and the disadvantaged, frontier, and outermost (3T) regions in infrastructure, devices, connectivity, and the digital competence of teachers and students (Hasdani et al., 2025), and earlier work documents structural barriers to educational access and inconsistent regional targeting of education programs (Falah & Hadna, 2022; Rosmana et al., 2022). A national household analysis found that school participation did not differ significantly between 3T and non-3T households, either before or after matching on connectivity, device ownership, and socioeconomic conditions, which cautions against treating regional labels as a direct measure of exclusion; that study, however, concerned school participation in 2022 and not the use of generative AI (Setyanti et al., 2025). Linguistic

diversity adds another layer: with more than 700 regional languages, Indonesia has many languages that remain low-resource in language technology (Ahmadian, 2026).

Despite this growing body of work, three limitations persist. First, much Indonesian research on generative AI describes perceptions and usage within single institutions and relatively small samples (e.g., Adawiyah et al., 2026; Zaini et al., 2025), so inequality is often treated as a background condition rather than an analytic focus. Second, work on digital inequality and work on the ethics of generative AI have developed largely in parallel. Third, international policy frameworks are rarely read alongside Indonesian evidence. This review addresses these limitations by reading global frameworks alongside Indonesian studies, labeling the strength of each body of evidence, and proposing a cumulative model of AI inequality.

The proposed model does not replace van Dijk's (2020) framework, which already describes how motivational, material, skills, and usage access accumulate into unequal outcomes. It extends that framework in three ways. First, it adds governance and institutional mediation (privacy, bias, academic integrity, language representation, and human agency) as a distinct stage between capability and outcomes, because generative AI reaches learners through institutional rules and commercial platforms that shape who benefits (Miao & Holmes, 2023; Selwyn, 2019). Second, it adds a feedback loop in which outcomes reshape later resources, so that inequality accumulates over time and not only across stages. Third, it labels every stage and link by evidence status, so that readers can see which parts rest on Indonesian empirical evidence, which on literature and policy, and which are the authors' proposals. The contribution is therefore a specification of an existing framework for generative AI in education, offered as a basis for testable hypotheses.

The review addressed three research questions:

1. How can generative AI influence educational opportunity, according to global and Indonesian literature?
2. Which social mechanisms explain why the benefits of generative AI are distributed unequally?
3. What human-centered strategies can support the equitable implementation of generative AI in education?

2. METHOD

Research Design

This study used an integrative literature review to synthesize scholarship on generative AI and educational inequality. An integrative design is suited to topics in which conceptual frameworks, policy documents, and empirical studies of varying design must be combined into a coherent account (Snyder, 2019; Torraco, 2016). The review did not estimate a pooled statistical effect and did not claim to represent all empirical research on generative AI. The review followed the five-stage logic of integrative reviews (problem identification, search, evaluation, analysis, and presentation) and is not a systematic review.

Source Identification and Selection

The analytic corpus comprised 24 sources: 17 publications produced in Indonesia between 2022 and 2026 (16 journal articles and one conference proceedings paper) and seven international policy documents and scholarly works published between 2016 and 2023. All 24

sources are cited in the Results and Discussion. Three further references on review methodology (Braun & Clarke, 2006; Snyder, 2019; Torraco, 2016) guided the design and analysis but were not part of the analytic corpus, so the reference list contains 27 entries.

Indonesian sources were identified through Indonesian- and English-language keyword searches (for example, kecerdasan buatan generatif, ChatGPT, kesenjangan digital, literasi AI, etika AI, and bahasa daerah) on national journal portals and Google Scholar. The seven international sources were not retrieved through these searches. They were selected purposively as foundational frameworks (van Dijk, 2020; Selwyn, 2019; Luckin et al., 2016; Holmes et al., 2019), policy documents (Miao & Holmes, 2023; UNESCO, 2023), and a systematic review (Zawacki-Richter et al., 2019), because they supply the concepts and policy anchors against which the Indonesian evidence is read.

Sources were retained when they (a) addressed generative AI or AI in education, digital inequality, AI literacy, ethics, privacy, or language resources; (b) originated from a scholarly journal, a recognized publisher, or an authoritative policy body; (c) were relevant to the Indonesian educational context or foundational to the analytic framework; and (d) were published between 2022 and 2026 (Indonesian sources) or between 2016 and 2023 (international sources). Sources that did not meet these criteria were excluded.

Analytical Procedure

The analysis involved three stages. First, sources were read to identify the educational functions attributed to generative AI, the social groups potentially affected, and the design, sample, and data type of each source. Second, findings were coded and grouped into four themes (access, capability, outcomes, and governance) using thematic analysis (Braun & Clarke, 2006). Third, the themes were interpreted through the concept of cumulative digital inequality, in which unequal resources at one stage generate unequal opportunities at later stages (van Dijk, 2020), and this interpretive step produced the model presented in the Results and Discussion. Coding was carried out by the first author and cross-checked by the second and third authors; the theme structure was refined through discussion until consensus was reached.

Conflicting findings were retained and reported as differences by population and setting and were not averaged. For example, students tended to view ChatGPT favorably whereas lecturers were more critical. Because no scoring instrument was used, source credibility was weighed qualitatively. Each source was labeled by design (national household survey, small survey, qualitative study, literature review, or conceptual and policy document), by sample size where reported, and by whether the data were self-reported. Claims resting on small, single-institution, cross-sectional, or self-reported samples are treated as indicative and are described with their limits, and international and conceptual sources are used to explain mechanisms and not as evidence of prevalence in Indonesia. Table 1 reports the evidence base and evidence status of each theme.

Scope and Limitations of the Method

The corpus was assembled purposively, and no formal quality-scoring instrument was applied. The review should therefore be read as a theoretically grounded synthesis of social mechanisms and practical implications for equitable educational policy and practice, not as an exhaustive census of the evidence. Further limitations are discussed after the Results.

3. RESULT AND DISCUSSION

The synthesis produced four themes: unequal access, unequal digital and AI literacy, unequal educational outcomes, and governance risks concerning privacy, bias, integrity, and human agency. Table 1 summarizes the themes together with the type and status of the evidence behind each, and the sections below discuss each in turn before presenting a cumulative model of AI inequality. Overall, the Indonesian studies were mostly cross-sectional surveys, small qualitative studies, and literature reviews, so the findings describe associations and perceptions and do not establish causal effects. Evidence on learning outcomes across social groups was scarce, and Fadlia et al. (2026) similarly noted limited research on the long-term effects of ChatGPT on learning quality. Statements that rest on conceptual or international sources, and not on Indonesian data, are identified as such in the text.

Table 1. Thematic Synthesis, Evidence Base, and Evidence Status of the Reviewed Literature

Theme	Key finding	Sources (type; sample)	Evidence status	Social implication
Access	Connectivity gaps track regional and socioeconomic disadvantage, but one national analysis found no significant 3T versus non-3T difference in school participation, so regional labels are an imperfect proxy.	Setyanti et al. (2025): national household survey, propensity score matching; outcome is school participation, not AI use. Hasdani et al. (2025): descriptive analysis of barriers. van Dijk (2020): conceptual.	Indonesian empirical, indirect (not AI use), plus conceptual	Use household-level indicators alongside regional ones, and measure AI-specific access.
Capability	Benefits appear to depend on AI literacy, ethical awareness, prior knowledge, and verification skills; norms for use remain unsettled.	Zaini et al. (2025): survey of 31 student users and lecturer interviews. Harahap et al. (2026): cross-sectional survey, one university. Niyu et al. (2024): survey of 430 (119 lecturers, 311 students). Salsabila & Sohidin (2024); Fitriani & Arfini (2025); Nugraha (2026): qualitative.	Indonesian empirical; small, self-reported, single-institution	Develop teacher and student AI literacy together and embed it in disciplines.
Outcomes	Unequal starting resources may turn equal access into unequal benefit; evidence on outcomes is scarce and mostly perceptual.	Luthfiyyah et al. (2024): perceptions of students in one program. Adawiyah et al. (2026): 43 students. Fadlia et al. (2026): review of 15 studies. Zawacki-Richter et al. (2019): systematic review.	Perception data; outcome evidence scarce; mechanism is theoretical	Provide differentiated support and evaluate outcomes, not adoption rates.
Governance	Privacy, bias, integrity, language representation, and human agency call for institutional safeguards.	Diman et al. (2025): literature-based ethics discussion. Saputra et al. (2025): review of 34 sources. Ahmadian (2026): review of 13 studies. Miao & Holmes (2023): policy guidance.	Mainly literature and policy based; limited Indonesian empirical evidence	Set transparent rules, keep human oversight, build inclusive language resources, and offer non-AI alternatives.

Note. Evidence status refers to the evidence available in the reviewed corpus; none of the Indonesian sources estimates causal effects of generative AI on learning.

Theme 1: Unequal Access to Generative AI

Generative AI can appear inexpensive because many tools are available through web interfaces. Meaningful use, however, still requires devices, stable connectivity, digital accounts, and sufficient time and support. Van Dijk (2020) argued that digital inequality involves not only physical access but also differences in motivation, skills, usage patterns, and outcomes. This framework is especially useful for generative AI because advanced users can extract far more value from the same tool than inexperienced users.

Indonesian evidence on material preconditions is real but indirect. Hasdani et al. (2025) identified limited devices, weak connectivity, and low digital competence among teachers and students as barriers that keep digitalization from being fully inclusive, particularly in 3T regions. Setyanti et al. (2025) analyzed 2022 national household data on 256,199 households with school-age children, using propensity score matching. In their model, internet access showed the strongest negative association with residing in a 3T region, followed by internet subsidies and mobile phone access, so 3T households were less connected. The outcome studied, however, was school participation and not the use of AI, and the data predate the widespread use of generative AI in classrooms. School participation was already similar in 3T and non-3T households before matching (92.41% and 92.25%), and the matched difference of about 1.2 percentage points was not statistically significant (standard error of about 2.7 percentage points). A non-significant result is not proof of equivalence, and the analysis is observational and limited to measured covariates, so it cannot show that household conditions cause educational exclusion. It supports only a cautious inference for generative AI: spatial classifications may be an imperfect guide to who is excluded, and household-level indicators deserve attention alongside regional ones (Setyanti et al., 2025). Whether the same holds for AI use remains untested.

Access also carries hidden costs. Students who share devices or rely on unstable connections may spend more effort overcoming technical barriers and less on higher-order learning, so a free interface does not guarantee free educational support. This point is an inference from the access literature rather than a finding of the Indonesian AI studies reviewed here, and it deserves direct empirical testing.

Theme 2: Unequal Digital and AI Literacy

Effective use of generative AI requires users to formulate suitable prompts, provide relevant context, recognize uncertainty, verify claims, protect personal information, and understand the limits of automated systems. UNESCO's guidance stresses that institutions should build human capacity and ethical validation rather than simply deploy new tools (Miao & Holmes, 2023). AI literacy should therefore be treated as a social and educational competence, not as a technical accessory.

Indonesian studies suggest that use has outpaced this competence. In a survey of 31 student ChatGPT users combined with interviews with lecturers at one university, Zaini et al. (2025) found that most students used the tool for assignments and to understand course material but did not fully grasp the importance of ethical use. Qualitative work by Salsabila and Sohidin (2024) reached a similar conclusion: students implicitly understood academic ethics, yet some still committed violations, showing that understanding was not consistently applied. In a cross-sectional survey of economics education students at one university ($n = 214$), Harahap et al. (2026) reported that AI literacy and ethical awareness were each positively associated with learning integrity, together accounting for 46.2% of its variance. Because these data are self-reported and correlational, they show association rather than causation, but they support the view that literacy is a precondition for responsible benefit.

Students also expressed a need for support. Fitriani and Arfini (2025) found that students felt they needed lecturer guidance so that AI use remained appropriate and ethical, and Niyu et al. (2024) observed that a social system governing academic ChatGPT use had not yet formed. Small studies also report benefits: Adawiyah et al. (2026) found that 43 primary-

teacher-education students who already used ChatGPT perceived it as an effective medium for self-directed learning that supported their digital literacy. Such positive self-assessments call for caution, however, because they come from a small, single-institution sample of existing users and cannot show whether less-resourced learners would gain equally.

Teacher capacity is part of the same problem. Hasdani et al. (2025) point to limited digital competence among teachers as a barrier to inclusive digital education, and views of ChatGPT among academics remain ambivalent. A review of 15 studies found that students tended to view it positively for its efficiency and accessibility, whereas lecturers were more critical, emphasizing academic integrity, plagiarism, and possible erosion of critical thinking (Fadlia et al., 2026; see also Kusumaningrum et al., 2023, on lecturers' perceptions in Indonesia).

AI literacy is most useful when embedded in disciplinary learning. Students need to learn not only how to operate a system but also how to question its outputs against the knowledge standards of a discipline. In social science, for example, learners can compare AI-generated explanations with empirical evidence, identify unsupported causal claims, and ask whose perspectives are missing. This approach turns AI use into an opportunity for critical inquiry.

Theme 3: Unequal Educational Outcomes

Generative AI can provide explanations, examples, feedback, language support, and opportunities for practice, functions that may matter most for learners with limited access to private tutoring. Students of a primary-teacher education program at Universitas Pendidikan Indonesia reported that ChatGPT made assignments easier and gave them an overview of how to approach tasks, although they held both favorable and unfavorable views of its presence (Luthfiyyah et al., 2024).

The benefits are unlikely to be evenly distributed. Students with stronger prior knowledge are better positioned to identify weak answers and request useful revisions, whereas students with less prior knowledge may accept plausible but incorrect outputs. AI can therefore lower some barriers while reinforcing others. This is a theoretical expectation consistent with the reviewed literature rather than an established empirical result: a systematic review of AI in higher education found that research gave limited critical attention to risks and had weak links to pedagogical theory (Zawacki-Richter et al., 2019). Luckin et al. (2016) emphasize the potential of AI to support learners, whereas Selwyn (2019) urges scrutiny of who gains and who loses. The present synthesis suggests that both positions are defensible, depending on how access, capability, and mediation are distributed.

These mechanisms create the possibility of a double gap for disadvantaged learners: a gap in access to reliable tools and a gap in the capacity to use them productively. Providing every student with identical access may appear fair, yet identical provision can produce unequal outcomes when students begin with different resources and competencies. Equity may therefore require differentiated support, such as instruction in prompting and verification for some learners, language support or accessibility features for others, and alternative learning resources where connectivity is unreliable.

For this reason, AI adoption should be evaluated through outcomes rather than adoption rates. A university may report high AI use while failing to improve equity. The more relevant questions concern who benefits, who is excluded, what competencies are required, and whether the technology changes opportunities for disadvantaged learners.

Language and Cultural Representation

Many generative AI systems perform strongly in dominant languages, while support for minority languages and culturally specific knowledge varies. UNESCO identifies linguistic and cultural diversity as a dimension of responsible AI (Miao & Holmes, 2023). In Indonesia the issue is concrete. Ahmadian (2026), reviewing 13 studies on natural language processing for regional languages from 2020 to 2025, showed that research has moved from fragmented studies of relatively large languages, such as Sundanese and Madurese, toward multilingual benchmarks that also cover low-resource languages, yet gaps in data standardization and large-scale pretraining resources remain. The same review recommended evaluating models for sociolinguistic sensitivity, including their handling of formal and informal speech registers.

Representation matters beyond language as well. Diman et al. (2025) illustrate how a model trained mainly on data from urban students may fail to capture the learning styles and difficulties of rural students. If educational AI primarily reflects dominant linguistic and cultural patterns, learners from less represented communities may receive less relevant support. Equity therefore requires attention to language resources and contextual appropriateness, together with open educational resources and multilingual AI resources that do not leave linguistically diverse communities behind.

Theme 4: Privacy, Bias, Integrity, and Human Agency

Educational use of generative AI can involve the processing of students' prompts, writing, learning difficulties, and personal information. Diman et al. (2025) warn that collecting and using student data without clear transparency can violate privacy. Saputra et al. (2025), reviewing 34 academic and policy sources, concluded that massive data collection without strict control risks privacy violations, that algorithms trained on unrepresentative data can reinforce educational injustice, and that unclear regulation creates accountability gaps. From a social science perspective, privacy is therefore not simply a technical matter. It concerns power relationships among students, institutions, technology providers, and data systems, and learners often have limited ability to understand or negotiate how their data are collected and used.

Academic integrity is a second concern. Generative AI can support drafting and revision, but unrestricted use may blur the boundary between student-generated and machine-generated work. Indonesian studies document uncertainty rather than simple misconduct (Salsabila & Sohidin, 2024; Zaini et al., 2025). In a qualitative descriptive study, Nugraha (2026) identified three patterns of student interaction with generative AI (full, adaptive, and critical reproduction) and found that students framed academic ethics through honesty, fairness, and responsibility, which suggests that practice varies even where values are shared. Lecturers, for their part, tend to emphasize plagiarism risks (Fadlia et al., 2026). A purely punitive response could widen inequality, because students with more access to private guidance may learn to use AI more strategically while others are penalized under unclear norms. Zaini et al. (2025) recommend formal regulation, sustained guidance on digital ethics, and greater AI literacy. A complementary step is to redesign assessment so that students demonstrate process, reflection, source evaluation, and oral or contextual application rather than relying only on final products.

A third concern is automation bias. Users may treat fluent AI output as authoritative, which is especially problematic while learners are still developing disciplinary knowledge. Teachers therefore remain essential as interpreters, mentors, and evaluators. Diman et al. (2025) argue that educators must be involved in AI-based decisions and given a sound understanding of ethical and privacy practice, and Fadlia et al. (2026) describe a shift in the lecturer's role from primary source of knowledge to facilitator of learning technology. AI should augment human judgment rather than replace it.

A Cumulative Model of AI Inequality

Taken together, the themes suggest a cumulative model of AI inequality (Figure 1 and Table 2). The model extends van Dijk's (2020) sequence of access in the three ways described in the Introduction, and it has four stages. Resources and access refers to the household and institutional resources, such as connectivity, devices, affordability, time, and language resources, that determine whether a learner can use generative AI at all. Capability refers to the digital, AI, and disciplinary competence needed to prompt, verify, and critique outputs. Governance and institutional mediation refers to teacher competence, institutional rules, data protection, bias safeguards, and assessment design, which shape whether use is reliable, ethical, and permitted on fair terms. Educational outcomes refers to the learning, qualifications, and opportunities that result.

Inequality accumulates in two ways. First, advantage or disadvantage at one stage raises or lowers the chance of benefiting at the next, so that a learner with limited resources faces a double gap, in access to reliable tools and in the capacity to use them productively. Second, outcomes feed back into later resources, because learners who benefit more tend to obtain better access, support, and credentials (the dashed arrow in Figure 1). Governance is treated as a distinct stage because the reviewed literature presents privacy, bias, integrity, and agency as conditions under which the benefits of AI are distributed. It may act as a safeguard, or as an additional source of inequality when rules are unclear or unevenly enforced (Diman et al., 2025; Saputra et al., 2025; Zaini et al., 2025).

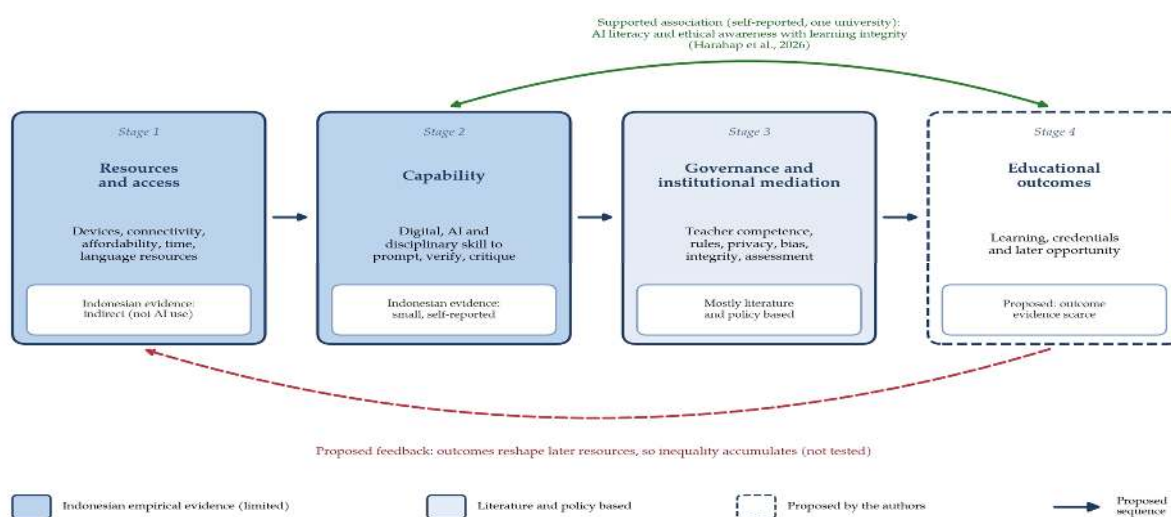


Figure 1. Cumulative Model of AI Inequality in Education. Source: authors' synthesis, extending van Dijk (2020). Shading shows the evidence status of each stage, the green arrow shows the one supported association, and dashed elements are proposed and untested.

Figure 1 and Table 2 show which parts of the model are supported. Stage 1 rests on Indonesian evidence that is empirical but indirect, because it concerns connectivity and school participation and not AI use. Stages 2 and 3 rest on small, mostly self-reported, single-institution Indonesian studies that report limited ethical understanding, unsettled norms, and a need for guidance, together with literature-based discussion of privacy and bias. The only link with direct support is an association between AI literacy and learning integrity (Harahap et al., 2026), which is not an achievement outcome. The proposed sequence, the effect of governance on outcomes, the accumulation across stages, and the feedback loop are the authors' propositions, based on conceptual work (Selwyn, 2019; van Dijk, 2020), and none has been tested in the reviewed corpus. The model should therefore be read as a framework for generating hypotheses.

Table 2. Cumulative Model of AI Inequality in Education: Stages, Evidence, and Evidence Status

Stage	Mechanism	Illustrative evidence	Evidence status	Priority intervention
1. Resources and access	Socioeconomic and infrastructural resources shape access to devices, connectivity, and time.	Hasdani et al. (2025); Setyanti et al. (2025)	Indonesian empirical, indirect (not AI use)	Connectivity, devices, and subsidies targeted by household need.
2. Capability	Prior digital and academic competence shapes the quality of AI use.	Harahap et al. (2026); Fitriani & Arfini (2025); Zaini et al. (2025); Nugraha (2026)	Indonesian empirical; small, self-reported, single-institution	AI literacy and verification training embedded in disciplines.
3. Governance and institutional mediation	Teacher competence, rules, and data governance shape the reliability, ethicality, and fairness of use.	Zaini et al. (2025); Diman et al. (2025); Saputra et al. (2025); Miao & Holmes (2023)	Mainly literature and policy based	Teacher development, transparent policy, privacy by design, and assessment redesign.
4. Outcomes	Accumulated differences shape learning and opportunity and may feed back into Stage 1.	Luthfiyyah et al. (2024); Fadlia et al. (2026); Zawacki-Richter et al. (2019)	Proposed; outcome evidence scarce	Outcome-based evaluation and differentiated support.

Intervention at the first stage alone is therefore unlikely to be sufficient. Providing access without developing capability may produce a new form of inequality: equal access but unequal benefit. The Indonesian studies reviewed here point in the same direction, since they identify infrastructure, teacher competence, literacy, and regulation as interdependent needs (Hasdani et al., 2025; Saputra et al., 2025; Zaini et al., 2025).

Implications for Policy, Institutions, and Pedagogy

The following recommendations are the authors' inferences from the synthesis and not findings of the reviewed studies. They are ordered by the strength and convergence of the supporting evidence and not by assumed importance, and each carries the limits of that evidence.

Priority 1: clear institutional guidance, AI literacy, and teacher development. This is the recommendation with the most convergent Indonesian support, although every supporting study is small or self-reported. Students report unsettled norms and a need for lecturer guidance (Fitriani & Arfini, 2025; Niyu et al., 2024; Zaini et al., 2025), understanding of academic ethics is not consistently applied (Salsabila & Sohidin, 2024), and limited teacher competence is a recognized barrier (Hasdani et al., 2025). Institutions should therefore state which assistance is acceptable and which is substitution, build AI literacy into disciplinary learning, include ethical reasoning and assessment redesign in teacher education, and have

students verify AI outputs against reliable sources. Because a purely punitive response could widen inequality when norms are unclear, assessment should let students show process, reflection, and individual reasoning.

Priority 2: household-level targeting of connectivity and devices. This rests on a single large national dataset whose outcome was school participation (Setyanti et al., 2025), together with descriptive evidence of infrastructure gaps (Hasdani et al., 2025). Public policy should integrate AI strategies with broader digital inclusion and, where data allow, target support by household indicators such as connectivity, device ownership, and social protection and not by regional classification alone. Whether this reduces gaps in AI use has not yet been tested.

Priority 3: privacy, governance safeguards, and inclusive language resources. The evidence here is mainly literature-based and policy-based, with limited Indonesian empirical testing (Ahmadian, 2026; Diman et al., 2025; Saputra et al., 2025; Miao & Holmes, 2023). Institutions should keep governance transparent and periodically reviewed, provide channels for reporting errors, bias, and privacy concerns, and preserve the right to limit AI use for legitimate privacy, cultural, or accessibility reasons, with non-AI alternatives for essential educational functions. Public institutions should consider open educational resources and multilingual AI resources so that linguistically diverse communities are not excluded.

Priority 4: pedagogical design and outcome evaluation. This is the least evidenced area, because no reviewed study measured learning outcomes across social groups. Educators can embed generative AI in pedagogical relationships, with students generating questions, comparing explanations, and obtaining feedback while teachers facilitate verification, reflection, and ethical judgment, and AI use should be evaluated through outcomes and equity effects and not adoption rates. Translation, adaptive explanation, accessible text transformation, and low-cost tutoring are plausible supports for learners who lacked individualized assistance, but their benefits remain a hypothesis for Indonesian settings.

Overall, generative AI may be best understood as a social infrastructure whose benefits are shaped by institutions and inequalities, which is an interpretation of the synthesis and not an empirical finding. The technology does not itself determine whether education becomes more equal; equity is more likely when access, competence, governance, cultural relevance, and human support are developed together.

Limitations and Directions for Future Research

Several limitations qualify these conclusions. First, the corpus was purposive, and no formal quality scoring was applied, so studies of different rigor were weighed by judgment. Second, the evidence base is uneven in quality and recent. Fourteen of the 17 Indonesian sources were published in 2024–2026, in a field that is changing quickly, so the tools, practices, and policies they describe may already differ from current conditions, and the sources vary in design, peer-review rigor, and methodological detail. Many were cross-sectional, self-reported, or based on small single-institution samples, and few measured learning outcomes across social groups. Third, no longitudinal study was identified, so the review cannot say whether short-term gains in productivity translate into durable learning or whether inequalities widen over time, which the feedback loop of the model assumes. Fourth, the search was limited to Indonesian- and English-language sources on national portals and Google Scholar, and publication bias is likely, because studies reporting benefits or high adoption may be more likely to be published; the seven international sources were few and

chosen as conceptual anchors, and Indonesian findings cannot be generalized to other Global South settings.

These limits constrain the claims made here. The cumulative model and the ordered recommendations are propositions for testing, not demonstrated mechanisms or proven interventions, and they should be applied with caution in specific institutions and regions. Future research should investigate whether AI-supported learning narrows or widens achievement gaps, using longitudinal and multi-institution designs, and compare AI use across socioeconomic status, disability, geographic location, and language background. It should also test whether AI-literacy interventions reduce differences in outcomes, examine the hidden costs of use for learners with shared devices or unstable connectivity, and measure access at the household level in studies of AI use.

4. CONCLUSION

This review asked three questions. Regarding the first, how generative AI may influence educational opportunity, global and Indonesian sources report that it can offer explanations, feedback, translation, and writing support, and Indonesian students and lecturers show growing awareness and use (Niyu et al., 2024). The Indonesian evidence, however, consists mostly of perceptions and small surveys, and it does not show whether generative AI has improved or worsened outcomes for any social group. Regarding the second, the social mechanisms that may explain unequal benefit are unequal enabling conditions, uneven AI literacy and unsettled norms, unequal institutional mediation, and under-resourced regional languages, which the proposed model describes as cumulative. The evidence for each mechanism is partial and mostly cross-sectional, and the accumulation itself is a proposition. Regarding the third, human-centered strategies should combine the following, in order of evidential support: clear institutional guidance with AI literacy and teacher development; household-level targeting of connectivity and devices; privacy and governance safeguards with inclusive language resources; and pedagogical redesign with outcome-based evaluation.

The review contributes a cumulative model that extends van Dijk's sequence of digital inequality with a governance stage, a feedback loop, and explicit evidence labeling, and it positions educational equity as the guiding principle for adoption decisions. The decisive question is not whether learners can open an AI tool but whether they have the resources, literacy, institutional support, and critical capacities to benefit from it. The model is offered as a basis for testable hypotheses. Priority for future work is research that compares AI use across socioeconomic and linguistic groups, measures household-level access, uses longitudinal designs, and tests whether carefully designed AI-literacy interventions reduce differences in educational outcomes.

ACKNOWLEDGMENT

The authors acknowledge the researchers and institutions whose publicly available studies and policy documents informed this review. No external funding was received for the preparation of this article.

5. REFERENCES

- Adawiyah, R., Yulinda, E., Mubarak, A. F., & Setyawan, A. (2026). Analisis peran ChatGPT sebagai media pembelajaran mandiri dalam meningkatkan literasi digital mahasiswa pendidikan guru sekolah dasar. *Elementary School: Jurnal Pendidikan dan Pembelajaran ke-SD-an*, 13(1), 195–209.
- Ahmadian, H. (2026). A systematic literature review of natural language processing for Indonesian regional languages. *JINTECH: Journal of Information Technology*, 7(1), 62–74.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Diman, P., Purwaka, A., & Linarto, L. (2025). Kecerdasan buatan dan etika pendidikan. *ENGANG: Jurnal Pendidikan, Bahasa, Sastra, Seni, dan Budaya*, 6(1), 230–239. <https://doi.org/10.37304/enggang.v6i1.24259>
- Fadlia, N., Aziz, S., & Sonni, A. F. (2026). Persepsi dosen dan mahasiswa terhadap penggunaan ChatGPT: Sebuah studi literatur. *NUSANTARA: Jurnal Ilmu Pengetahuan Sosial*, 13(6), 1580–1588. <https://doi.org/10.31604/jips.v13i6.2026.1580-1588>
- Falah, A. I., & Hadna, A. H. (2022). Problematika pendidikan masa pandemi di Indonesia pada daerah 3-T (terluar, tertinggal, dan terdepan). *Jurnal Pendidikan dan Kebudayaan*, 7(2), 164–185. <https://doi.org/10.24832/jpnk.v7i2.2997>
- Fitriani, F., & Arfini, B. D. (2025). Pemanfaatan artificial intelligence (AI) untuk meningkatkan kemampuan literasi mahasiswa. *Transformasi*, 11(1), 141–149. <https://doi.org/10.33394/jtni.v11i1.16739>
- Harahap, E. S., Khaira, I., & Handayani, R. (2026). Integritas pembelajaran di era generative AI: Peran literasi AI dan kesadaran etika mahasiswa pendidikan ekonomi. *Journal of Education Technology and Civic Literacy*, 6(2), 153–163. <https://jurnal.uisu.ac.id/index.php/jetcivil/article/view/13959>
- Hasdani, S. A., Widana, I. D. K. K., & Cahaya, Y. F. (2025). Pendidikan digital di Indonesia: Menjawab tantangan kesenjangan dan mewujudkan transformasi inklusif. *NUSANTARA: Jurnal Ilmu Pengetahuan Sosial*, 12(12), 4692–4696. <https://doi.org/10.31604/jips.v12i12.2025.4692-4696>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Kusumaningrum, S. R., Sukma, R., Dewi, I., & Pristiani, R. (2023). Persepsi dosen di Indonesia terhadap penggunaan ChatGPT di lingkup akademik. *Community Development Journal: Jurnal Pengabdian Masyarakat*, 4(6), 11898–11905.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Luthfiyyah, K., Zhafira, L., Nurani, S., & Giwangsa, S. F. (2024). Analisis peran artificial intelligence (AI): ChatGPT dalam perkuliahan di kalangan mahasiswa PGSD Universitas Pendidikan Indonesia. *Jurnal Pendidikan Tambusai*, 8(1), 5282–5290.

- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Niyu, Dwihadiah, D., Gerungan, A., & Purba, H. (2024). Penggunaan ChatGPT di kalangan mahasiswa dan dosen perguruan tinggi Indonesia. *Coverage: Journal of Strategic Communication*, 14(2), 130–145. <https://doi.org/10.35814/coverage.v14i2.6058>
- Nugraha, W. (2026). Perilaku mahasiswa dalam memanfaatkan generative AI: Tinjauan etika akademik. *Manifestasi: Jurnal Pendidikan dan Pembelajaran*, 6(2), 13–18. <https://ojs.unm.ac.id/manifestasi/article/view/87170>
- Rosmana, P. S., Iskandar, S., Fadilah, N., Azhar, N., Oktavini, D., & Munte, A. C. (2022). Upaya pemerataan pendidikan berkelanjutan di daerah 3T. *Attadib: Journal of Elementary Education*, 6(2), 405–418. <https://doi.org/10.32507/attadib.v6i2.1212>
- Salsabila, S., & Sohidin. (2024). Pemahaman etika akademik mahasiswa dalam penggunaan artificial intelligence (AI). *Journal of Education Research*, 5(4), 6671–6680. <https://doi.org/10.37985/jer.v5i4.1944>
- Saputra, R. F., Zamsuri, A., & Amri, A. D. (2025). Tantangan etika teknologi dalam pemanfaatan artificial intelligence: Privasi, bias, dan akuntabilitas dalam pendidikan di Indonesia. *Prosiding Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMASTER)*, 4(1), 65–71. <https://doi.org/10.31849/r0b2fm79>
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity.
- Setyanti, A. M., Prestianawati, S. A., & Fawwaz, M. (2025). From isolation to inclusion: Does digital adoption enhance educational access in Indonesia's 3T areas? *Salus Cultura: Jurnal Pembangunan Manusia dan Kebudayaan*, 5(1), 14–26. <https://doi.org/10.55480/saluscultura.v5i1.368>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review*, 15(4), 404–428. <https://doi.org/10.1177/1534484316671606>
- UNESCO. (2023). *Global education monitoring report 2023: Technology in education – A tool on whose terms?* UNESCO.
- van Dijk, J. (2020). *The digital divide*. Polity.
- Zaini, A. N., Arisanty, D., & Rini, S. (2025). Pengendalian dalam penggunaan chatbot AI (ChatGPT) dalam dunia pendidikan di kalangan mahasiswa teknologi pendidikan Universitas Lambung Mangkurat. *J-INSTECH*, 6(2), 93–99. <https://doi.org/10.20527/j-instech.v6i2.15624>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the

educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>