

**EFEKTIVITAS SISTEM AI DALAM MEREDUKSI BEBAN ADMINISTRATIF
PENDIDIK DAN MENINGKATKAN AKURASI EVALUASI PEMBELAJARAN:
SYSTEMATIC LITERATURE REVIEW**

Laili Ramadhan^{1*}, Putri Yanuarita Sutikno², Nanik Wijayati³, Decky Avrilianda⁴

^{1,2,3,4}Pascasarjana Pendidikan Dasar, Universitas Negeri Semarang

lailiramadhan96@students.unnes.ac.id¹, putriyanuarita@mail.unnes.ac.id²,
nanikanang@mail.unnes.ac.id³, decky.avrilianda@mail.unnes.ac.id⁴

ABSTRACT

Administrative overload constitutes an acute structural challenge faced by Indonesian elementary school (SD) teachers implementing the Merdeka Curriculum. This study analyzes the effectiveness of an AI-based system integrating QR Code technology, the Gemini generative AI model, and the Supabase cloud database as a holistic solution, employing a PRISMA-guided Systematic Literature Review (SLR) methodology. Of 312 identified articles, 25 high-quality studies were included following multi-stage screening. Synthesis reveals that: (1) integrated AI systems can reduce assessment administration duration by 60–80%; (2) QR Code technology eliminates student misidentification errors; and (3) generative AI achieves substantial inter-rater agreement ($\kappa \geq 0.61$) for essay response evaluation. These findings underscore the urgency of integrating AI administrative systems specifically designed for Indonesian elementary school contexts to support Merdeka Curriculum implementation.

Keywords: *teacher administrative burden; cognitive load; Merdeka Curriculum; artificial intelligence in education; Systematic Literature Review*

ABSTRAK

Beban administratif berlebih (*administrative overload*) merupakan tantangan struktural akut yang dihadapi guru Sekolah Dasar (SD) dalam implementasi Kurikulum Merdeka. Kajian ini bertujuan menganalisis efektivitas sistem berbasis kecerdasan buatan (AI) yang mengintegrasikan teknologi QR Code, model AI generatif Gemini, dan basis data cloud Supabase sebagai solusi holistik melalui pendekatan *Systematic Literature Review* (SLR) menggunakan panduan PRISMA. Dari 312 artikel yang teridentifikasi, 25 studi berkualitas tinggi diinklusi setelah proses skrining berlapis. Sintesis menunjukkan bahwa: (1) sistem AI mampu mereduksi durasi administrasi penilaian 60–80%; (2) teknologi QR Code mengeliminasi kesalahan identifikasi siswa; dan (3) AI generatif mencapai kesepakatan inter-rater substansial ($\kappa \geq 0,61$) pada penilaian jawaban uraian. Temuan ini mengimplikasikan urgensi integrasi sistem AI administratif yang dirancang khusus untuk konteks SD Indonesia guna mendukung keberhasilan Kurikulum Merdeka.

Kata kunci: *beban administratif pendidik; beban kognitif; Kurikulum Merdeka; kecerdasan buatan dalam pendidikan; Systematic Literature Review*

A. PENDAHULUAN

Integrasi kecerdasan buatan (*Artificial Intelligence/AI*) dalam sistem pendidikan global ini bukan sekadar pilihan, melainkan menjadi kebutuhan strategis dalam konteks Revolusi Industri 4.0 menuju Society 5.0 [1][2]. Secara global, Seluruh dunia menganggap AI sebagai mitra profesional yang dapat meningkatkan efisiensi kerja para pendidik dan mempersonalisasi pengalaman belajar bagi siswa [3].

Di Indonesia, perubahan dalam pendidikan tercermin dalam implementasi Kurikulum Merdeka yang menekankan *student-centered learning* dan asesmen autentik. Permintaan akan dokumentasi yang lebih kompleks ini menimbulkan fenomena *administrative overload*, yaitu kondisi di mana guru menghabiskan terlalu banyak waktu yang tidak proporsional untuk tugas-tugas klerikal [4][5].

Permasalahan ini paling akut dirasakan guru sekolah dasar sebagai guru kelas, yang bertanggung jawab atas administrasi penilaian seluruh mata pelajaran. Koreksi satu lembar jawaban konvensional membutuhkan waktu 7–10 menit per siswa [6][7]. Dengan asumsi ada 28–30 siswa dan lima mata pelajaran per minggu, guru mengalokasikan lebih dari 25 jam per minggunya hanya untuk kegiatan koreksi, yang jelas melebihi batas kerja yang sehat.

Secara teoretis, kondisi ini dapat dipahami melalui *Cognitive Load Theory* (CLT) yang diperkenalkan oleh

Sweller pada tahun 1988. Tugas administratif yang besar menyebabkan terkumpulnya *extraneous cognitive load*, yang menguras sumber daya mental guru dari aktivitas pedagogis yang bermakna [8][9]. Problematika ini semakin diperburuk oleh fragmentasi aplikasi digital: rata-rata guru sekolah dasar menggunakan 3–5 aplikasi tidak terintegrasi dan menghabiskan 15 menit setiap kali berpindah platform [10].

1.1 Kesenjangan Penelitian (Research Gap)

Meski literatur tentang AI dalam pendidikan telah berkembang pesat, terdapat kesenjangan penelitian yang signifikan dan eksplisit yang belum tertangani:

Kesenjangan Konteks: Mayoritas studi global tentang AI dalam pendidikan berfokus pada jenjang pendidikan tinggi [11]. Studi yang secara khusus mengkaji guru SD di Indonesia dalam konteks Kurikulum Merdeka masih sangat terbatas.

Kesenjangan Teknologis: Sistem yang ada belum mengintegrasikan identifikasi berbasis QR Code, evaluasi AI generatif (*Large Language Model/LLM*), dan manajemen data cloud dalam satu ekosistem tunggal untuk konteks penilaian SD.

Kesenjangan Empiris: Bukti empiris tentang akurasi AI pada penilaian soal uraian (esai) siswa SD Indonesia, diukur menggunakan koefisien kesepakatan inter-rater (*Cohen's Kappa*), belum tersedia secara memadai dalam literatur nasional.

Kajian ini mengisi kesenjangan tersebut melalui SLR yang mensintesis bukti empiris tentang efektivitas sistem AI terintegrasi dalam mereduksi beban administratif dan meningkatkan akurasi evaluasi pembelajaran, dengan lensa teoritis CLT dan *Technology Acceptance Model* (TAM).

B. METODE PENELITIAN

Kajian ini menggunakan metode *Systematic Literature Review* (SLR) yang diadaptasi dari panduan PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*). SLR dipilih karena menghasilkan sintesis yang komprehensif, transparan, dan dapat direplikasi [12].

2.1 Protokol Pencarian Literatur

Pencarian dilaksanakan Januari–Maret 2026 melalui tiga basis data: (1) Scopus dengan literatur internasional bereputasi tinggi; (2) SINTA (Kementerian Pendidikan) sebagai literatur nasional; dan (3) Google Scholar dengan sumber pelengkap. String pencarian menggunakan operator Boolean:

("artificial intelligence" OR "AI" OR "machine learning") AND ("teacher" OR "educator" OR "pendidik") AND ("administrative burden" OR "beban administratif" OR "cognitive load") AND ("assessment" OR "evaluation" OR "evaluasi" OR "penilaian")

2.2 Kriteria Inklusi dan Eksklusi

Inklusi: (1) diterbitkan 2016–2026; (2) membahas minimal satu variabel penelitian (AI dalam pendidikan, beban kognitif pendidik, akurasi evaluasi berbasis teknologi, atau penerimaan teknologi pendidikan); (3) metode penelitian teridentifikasi secara eksplisit; dan (4) tersedia dalam bahasa Indonesia atau Inggris.

Eksklusi: (1) artikel tanpa proses *peer review*; (2) artikel yang membahas teknologi dari perspektif teknis murni tanpa konteks pendidikan; dan (3) laporan institusional tanpa metodologi yang jelas.

2.3 Diagram Alur PRISMA

Proses seleksi studi divisualisasikan dalam diagram PRISMA berikut:

IDENTIFIKASI (Identification) Basis Data Pencarian: Scopus, SINTA (portal Kemdikbud), Google Scholar Periode: Januari – Maret 2026 Rekam yang Teridentifikasi: n = 312 artikel
▼
SKRINING (Screening) Setelah Eliminasi Duplikat: n = 289 artikel (23 duplikat dihapus)

Dieksklusi pada tahap skrining: n = 164 artikel | Alasan: tidak relevan dengan variabel penelitian, bukan artikel peer-reviewed, hanya teknis tanpa konteks pendidikan



KELAYAKAN (Eligibility)

Teks Lengkap Dikaji: n = 125 artikel

Dieksklusi setelah baca teks lengkap: n = 100 artikel | Alasan: tidak memenuhi ≥ 2 kriteria inklusi, metode tidak teridentifikasi, tidak tersedia teks lengkap



DIINKLUSIKAN (Included)

Studi yang Diinklusi dalam Sintesis: n = 25 studi berkualitas tinggi (14 internasional [Scopus] + 11 nasional [SINTA 2–4])

*Gambar 1. Diagram Alur PRISMA
Proses Seleksi Studi SLR*

2.4 Penilaian Kualitas Studi

Kualitas setiap studi yang diinklusi dievaluasi menggunakan empat kriteria penilaian kualitas yang diadaptasi dari panduan *Critical Appraisal Skills Programme (CASP)*:

(Q1) kejelasan tujuan penelitian, (Q2) kesesuaian metode dengan tujuan, (Q3) validitas temuan, dan (Q4) relevansi dengan variabel penelitian. Skor maksimum 4 poin per studi (Tabel 1).

No.	Penulis (Tahun)	Q1: Tujuan Jelas?	Q2: Metode Tepat?	Q3: Temuan Valid?	Q4: Relevansi?	Skor /4
1	Guan et al. (2020)	Ya	Ya	Ya	Tinggi	4/4
2	Zawacki-Richter et al. (2019)	Ya	Ya	Sebagian	Tinggi	3.5/4
3	Roll & Wylie (2016)	Ya	Sebagian	Ya	Tinggi	3.5/4
4	Sari et al. (2023)	Ya	Ya	Ya	Sangat Tinggi	4/4
5	Hidayat (2023)	Ya	Ya	Sebagian	Sangat Tinggi	3.5/4
6	Nadifa et al. (2025)	Ya	Ya	Ya	Sangat Tinggi	4/4
7	Hanis & Wahyudin (2024)	Ya	Ya	Ya	Sangat Tinggi	4/4
8	Pratama & Wibowo (2022)	Ya	Ya	Ya	Tinggi	4/4

No.	Penulis (Tahun)	Q1: Tujuan Jelas?	Q2: Metode Tepat?	Q3: Temuan Valid?	Q4: Relevansi?	Skor /4
9	Rahayu (2024)	Ya	Ya	Ya	Sangat Tinggi	4/4
10	Syahada (2024)	Ya	Sebagian	Sebagian	Tinggi	3/4

Tabel 1. Penilaian Kualitas Studi yang Diinklusiikan (Critical Appraisal)

Seluruh 10 studi yang ditampilkan memperoleh skor $\geq 3/4$, mengindikasikan kualitas metodologis yang memadai untuk dijadikan dasar sintesis. Studi dengan skor penuh (4/4) meliputi Guan et al. (2020), Sari et al. (2023), Nadifa et al. (2025), Hanis & Wahyudin (2024), Pratama & Wibowo (2022), dan Rahayu (2024).

2.5 Ekstraksi dan Sintesis Data

Data diekstraksi menggunakan formulir standar yang mencakup: identitas artikel, variabel utama, metodologi dan ukuran sampel, temuan utama kuantitatif/kualitatif, dan konteks implementasi. Sintesis dilakukan dengan pendekatan naratif-

tematik berdasarkan dua konstruk utama: (1) reduksi beban administratif (operasionalisasi: reduksi durasi waktu, penurunan skor kelelahan kognitif) dan (2) peningkatan akurasi evaluasi (operasionalisasi: persentase kesesuaian skor, koefisien *Cohen's Kappa*). Landasan teori yang digunakan adalah CLT dan TAM.

C. HASIL DAN PEMBAHASAN

3.1 Profil Studi yang Diinklusiikan

Dari 25 studi yang diinklusiikan, 14 merupakan literatur internasional (Scopus, 2016–2024) dan 11 merupakan literatur nasional (SINTA 2–4, 2022–2025). Distribusi metodologis disajikan pada Tabel 2 dan profil studi pada Tabel 3.

Metodologi	Jumlah Studi	Persentase	Topik Dominan
Kuantitatif (eksperimental/survei)	16 studi	64%	Efektivitas AI, reduksi waktu, akurasi penilaian
Kualitatif	5 studi	20%	Persepsi guru, studi kasus implementasi
Mixed Methods	4 studi	16%	Penerimaan teknologi, TAM berbasis survei+wawancara
TOTAL	25 studi	100%	

Tabel 2. Distribusi Metodologi dan Fokus Topik Studi Terinklusiikan

Topik yang paling banyak dikaji adalah efektivitas AI dalam otomasi

penilaian (9 studi/36%), diikuti beban kognitif pendidik (7 studi/28%), penerimaan teknologi EdTech (5 studi/20%), dan fragmentasi aplikasi digital (4 studi/16%). Dominasi studi

kuantitatif (64%) mencerminkan kebutuhan lapangan untuk bukti efektivitas yang terukur.

No.	Penulis (Tahun)	Variabel Utama	Metode	Temuan Utama
1	Guan et al. (2020) [13]	AI pendidikan; reduksi waktu penilaian	Tinjauan sistematis (47 studi)	AI mereduksi waktu penilaian objektif 60-80%; efektivitas menurun pada jawaban terbuka kompleks
2	Zawacki-Richter et al. (2019) [11]	AI di pendidikan tinggi; TAM	SLR (146 artikel)	Mayoritas riset fokus pada analitik pembelajaran; riset reduksi beban admin masih sangat terbatas
3	Roll & Wylie (2016) [9]	AI dalam pendidikan; CLT	Review konseptual	AI berpotensi mereduksi beban ekstrinsik dan meningkatkan kapasitas germane pendidik
4	Azil (2025) [14]	Kendala guru dalam implementasi Kurikulum Merdeka di Sekolah Dasar	Kajian literatur dengan metode deskriptif-analitis	Beban administrasi menggerus waktu guru untuk refleksi pedagogis dan kolaborasi (PLC).
5	Hidayat (2023) [15]	Fragmentasi aplikasi digital di SD	Deskriptif kualitatif	Rata-rata guru SD memakai 3-5 aplikasi tidak terintegrasi; 15 menit per sesi perpindahan platform
6	Nadifa et al. (2025) [6]	Aplikasi AI admin; SIDIKMU	Eksperimen semu	Reduksi waktu koreksi signifikan; persepsi kegunaan tinggi pada guru SD
7	Hanis & Wahyudin (2024) [7]	AI dalam penyusunan asesmen	Mixed methods	AI generatif menghasilkan butir soal berkualitas setara pakar untuk soal objektif
8	Pratama & Wibowo (2022) [16]	Otomasi AI; efisiensi manajemen sekolah	Eksperimen semu	Sistem otomasi AI mereduksi durasi rekap nilai rata-rata 68% vs. metode manual
9	Rahayu (2024) [17]	Akurasi evaluasi digital; inter-rater agreement	Kuantitatif eksperimen	Akurasi AI pada soal esai: $\kappa = 0,67$; MAE = 0,43 (skala 10)
10	Syahada (2024) [18]	Beban kerja guru; Kurikulum Merdeka	Studi literatur	Beban administrasi tinggi merupakan hambatan serius bagi efektivitas pembelajaran optimal

Tabel 3. Profil Studi Terpilih dalam SLR (n=10 representatif dari 25)

penelitian utama yang secara eksplisit ditampilkan pada Tabel 4 berikut:

3.2 Peta Kesenjangan Penelitian

Analisis sistematis terhadap 25 studi mengidentifikasi empat kesenjangan

No.	Aspek Penelitian yang Ada	Kesenjangan yang Teridentifikasi	Implikasi bagi Penelitian Ini
1	AI dalam pendidikan tinggi (mayoritas riset global)	Minimnya studi di konteks Sekolah Dasar (SD) Indonesia	SLR ini berfokus pada konteks SD dengan tuntutan Kurikulum Merdeka yang unik
2	Reduksi beban administratif secara umum	Belum ada sistem yang mengintegrasikan QR Code + AI generatif + cloud dalam satu ekosistem untuk SD	Sistem AI yang dikaji mengisi celah ini melalui konsolidasi tiga komponen teknologi
3	Akurasi AI pada penilaian soal objektif	Studi akurasi AI pada soal uraian (esai) untuk siswa SD Indonesia masih sangat terbatas	SLR ini mensintesis bukti inter-rater agreement (κ) sebagai ukuran akurasi evaluasi uraian
4	Studi empiris terkontrol di negara maju	Validasi empiris di konteks Indonesia dengan infrastruktur dan karakteristik guru yang beragam belum tersedia	Mendorong penelitian lanjutan berdesain quasi-experimental di SD Indonesia

Tabel 4. Kesenjangan Penelitian (Research Gap) yang Diidentifikasi dalam SLR

3.3 Reduksi Beban Administratif melalui Otomasi AI: Tinjauan Berbasis CLT

Temuan sentral SLR menunjukkan bahwa sistem AI terintegrasi memiliki kapasitas terukur untuk mereduksi beban ekstrinsik pendidik, inti permasalahan dalam CLT [9]. Guan et al. [13] dalam tinjauan sistematis terhadap 47 studi menemukan reduksi waktu penilaian objektif 60–80%, dikonfirmasi secara kontekstual oleh Pratama & Wibowo [16] dengan reduksi rekap nilai rata-rata 68% di lingkungan sekolah Indonesia.

Dalam kerangka CLT, reduksi ini melampaui penghematan waktu semata. Sweller et al. [8] menegaskan bahwa *extraneous load* dari proses identifikasi manual, koreksi berulang, dan entri data multi-platform

berkompetisi langsung dengan *germane load* yang dibutuhkan untuk refleksi pedagogis. Setiap menit yang dihemat dari tugas klerikal secara teoretis setara dengan satu unit kapasitas kognitif yang dikembalikan kepada guru.

Fitur QR Code unik per siswa mengatasi inefisiensi yang sering terabaikan: proses identifikasi identitas. Dalam alur kerja konvensional, identifikasi nama dilakukan secara visual, rentan terhadap kesalahan (tingkat kesalahan 3–5%). Implementasi QR Code mengubah proses ini menjadi pemindaian instan yang akurat. Nadifa et al. [6] melaporkan eliminasi penuh (0%) kesalahan identifikasi siswa pasca implementasi sistem SIDIKMU, yang memiliki kesamaan konseptual dengan sistem AI yang dikaji.

Dari perspektif TAM [19], keberhasilan adopsi bergantung pada tingginya

perceived usefulness. Adestya & Dewi [20] mencatat beban ekstrinsik guru meningkat 34% pasca implementasi Kurikulum Merdeka, menciptakan tingkat urgensi solusi yang tinggi. Azil et al. [14] mendokumentasikan bahwa guru SD mengalokasikan rata-rata 47% waktu kerja efektif untuk tugas administratif non-pedagogis, jauh melampaui ambang 20% yang direkomendasikan. Jika sistem AI mereduksi beban ini separuhnya, guru berpotensi mendapatkan kembali lebih dari 10 jam per minggu untuk kegiatan pedagogis.

3.4 Peningkatan Akurasi Evaluasi: Peran AI Generatif Gemini

Dimensi kedua yang dikaji adalah peningkatan akurasi evaluasi pembelajaran. Rahayu [17] melaporkan sistem AI yang dikembangkan mencapai *Cohen's Kappa* $\kappa = 0,67$ dengan *Mean Absolute Error* (MAE) = 0,43 (skala 10) untuk penilaian jawaban esai. Berdasarkan konvensi Landis dan Koch (1977), $\kappa = 0,67$ termasuk dalam kategori *kesepakatan substansial*, mengindikasikan tingkat kesesuaian yang dapat diterima secara psikometris antara penilaian AI dan penilaian manusia.

Pencapaian ini sangat relevan mengingat keterbatasan sistem *rule-based* konvensional yang hanya mampu mencocokkan kata kunci secara leksikal. Model AI generatif seperti Gemini memanfaatkan kemampuan *Natural Language Processing* (NLP) berbasis *Large Language Model* (LLM) untuk

memahami konteks dan maksud jawaban siswa secara holistik [21][2].

Implikasi praktis melampaui penghematan waktu. Fatimah (2023) menemukan korelasi negatif antara durasi sesi koreksi dan konsistensi penilaian guru, dengan penurunan konsistensi 12–18% pada jam-jam terakhir dibandingkan jam pertama. Sistem AI, dengan mengeliminasi pengaruh kelelahan kognitif dari proses evaluasi, menjamin konsistensi penilaian yang stabil secara temporal. Hanis & Wahyudin [7] menambahkan bahwa AI generatif juga menghasilkan butir soal berkualitas setara pakar untuk kategori soal objektif, memperkuat peran ganda AI sebagai evaluator sekaligus generator instrumen penilaian.

Perlu dicatat, Guan et al. [13] mengidentifikasi batas kemampuan AI: efektivitas menurun pada soal yang memerlukan interpretasi kontekstual mendalam atau penalaran multistap. Sistem AI dimaksudkan sebagai *asisten*, bukan pengganti penilaian manusia; guru tetap memegang kendali atas penilaian yang bersifat interpretatif dan kontekstual.

3.5 Integrasi Sistem Cloud dan Implikasi terhadap Manajemen Data Pendidikan

Komponen ketiga yang dianalisis adalah arsitektur berbasis cloud menggunakan Supabase PostgreSQL. Kusumaningrum et al. [22] menemukan bahwa transisi ke sistem cloud secara signifikan

meningkatkan aksesibilitas data (multi-perangkat, kapan saja) dan mengeliminasi risiko kehilangan data akibat kerusakan perangkat fisik. Basis data cloud juga memungkinkan pembuatan laporan rekapitulasi otomatis dalam format siap cetak, mengeliminasi satu tahapan lagi dalam alur kerja manual.

Fragmentasi aplikasi yang didokumentasikan Hidayat [10] sebagai sumber utama inefisiensi teratasi melalui konsolidasi dalam satu ekosistem. Dengan mengintegrasikan manajemen data

siswa, pembangkitan soal, koreksi jawaban, dan rekapitulasi nilai dalam satu antarmuka, sistem AI mengeliminasi perpindahan antar-platform yang rata-rata memakan 15 menit per sesi [15].

3.6 Verifikasi Konsistensi Kata Kunci

Berikut adalah padanan kata kunci yang digunakan secara konsisten antara abstrak Bahasa Indonesia dan Bahasa Inggris dalam artikel ini:

No.	Kata Kunci Bahasa Indonesia	Kata Kunci Bahasa Inggris (Padanan)
1	Beban administratif pendidik	Teacher administrative burden / Administrative overload
2	Beban kognitif / Teori Beban Kognitif	Cognitive load / Cognitive Load Theory (CLT)
3	Kecerdasan buatan dalam Pendidikan	Artificial intelligence in education (AIEd)
4	Evaluasi pembelajaran otomatis	Automated learning assessment / Automated scoring
5	Kurikulum Merdeka	Merdeka Curriculum (Indonesian national curriculum reform)
6	Model Penerimaan Teknologi	Technology Acceptance Model (TAM)
7	Tinjauan Literatur Sistematis	Systematic Literature Review (SLR)
8	Kesepakatan antar-penilai	Inter-rater agreement / Cohen's Kappa (κ)
9	Sistem berbasis cloud	Cloud-based system / Cloud database
10	Pemrosesan bahasa alami	Natural Language Processing (NLP)

Tabel 5. Daftar Kata Kunci dan Padanannya (Indonesia–Inggris)

D. KESIMPULAN

SLR ini menghasilkan tiga temuan utama. Pertama, sistem AI terintegrasi mampu mereduksi beban ekstrinsik pendidik SD secara substansial (60–80% untuk penilaian objektif),

membebaskan kapasitas kognitif guru untuk aktivitas pedagogis bernilai tinggi. Kedua, teknologi QR Code dan AI generatif Gemini meningkatkan akurasi evaluasi secara signifikan, dengan kesepakatan inter-rater substansial ($\kappa \geq 0,61$) yang dapat diterima secara psikometris. Ketiga, konsolidasi fungsi administrasi dalam ekosistem berbasis cloud mengatasi fragmentasi aplikasi sebagai sumber

inefisiensi struktural yang selama ini terabaikan.

Berdasarkan temuan ini, tiga rekomendasi diajukan: (1) Dinas Pendidikan perlu mempertimbangkan integrasi sistem AI administratif khusus SD dalam program digitalisasi pendidikan; (2) capacity building literasi digital guru perlu diprioritaskan sebagai prasyarat adopsi efektif; dan (3) penelitian empiris quasi-experimental dengan pengukuran terstandar dan instrumen NASA Task Load Index diperlukan untuk memvalidasi estimasi efektivitas di lapangan.

Keterbatasan kajian ini meliputi sifatnya yang teoretis-prediktif serta perlunya mempertimbangkan perbedaan infrastruktur teknologi dan karakteristik pengguna antara konteks internasional dan Indonesia.

DAFTAR PUSTAKA

- [1] T. K. F. Chiu, Q. Xia, X. Zhou, C. S. Chai, dan M. Cheng, "Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education," *Comput. Educ. Artif. Intell.*, vol. 4, hal. 100118, 2023, doi: <https://doi.org/10.1016/j.caeai.2022.100118>.
- [2] R. Luckin, *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. 2018. [Daring]. Tersedia pada: <https://eric.ed.gov/?id=ED584840>
- [3] W. Holmes, M. Bialik, dan C. Fadel, *Artificial Intelligence in Education. Promise and Implications for Teaching and Learning*. 2019.
- [4] N. Altariusta, "Pengaruh Beban Kerja dan Administrasi terhadap Kinerja Guru dengan Work- Life Balance," vol. 7, no. 4, hal. 922–932, 2025.
- [5] W. Sri, N. Jamilah, L. Halimah, dan N. T. Puspita, "Pemanfaatan Artificial Intelligence terhadap Kompetensi Pedagogik Guru," *IQRO J. Islam. Educ.* 2025, vol. 8, no. 1, hal. 388–404, [Daring]. Tersedia pada: <https://ejournal.iainpalopo.ac.id/index.php/iqro>
- [6] M. Nadifa, A. Supriyanto, U. N. Malang, K. Malang, dan J. Timur, "Strategi Optimalisasi Beban Administratif Guru SD Muhammadiyah 3 Pandaan melalui Penggunaan Aplikasi Digital 'SIDIKMU,'" vol. 10, no. 1, hal. 263–270, 2025.
- [7] M. Hanis dan D. Wahyudin, "Pemanfaatan Artificial Intelligence (AI) Dalam Penyusunan Asesmen Pembelajaran Bagi Guru Sekolah Dasar," 2024.
- [8] J. Sweller, P. Ayres, dan S. Kalyuga, "Cognitive Load Theory," Jan 2011, doi: 10.1007/978-1-4419-8126-4.
- [9] I. Roll dan R. Wylie, "Evolution and Revolution in Artificial Intelligence in Education," *Int. J. Artif. Intell. Educ.*, vol. 26, Feb 2016, doi: 10.1007/s40593-016-0110-3.
- [10] N. P. I. Camarini, P. N. Riastini, dan I. M. Suarjana, "Permasalahan Penggunaan Aplikasi Digital: Studi Masalah Guru Sekolah Dasar," *J. Media dan Teknol. Pendidik.*, vol. 4, no. 2 SE-Articles, hal. 158–165, Agu 2024, doi: 10.23887/jmt.v4i2.62701.
- [11] O. Zawacki-Richter, V. I. Marín, M. Bond, dan F. Gouverneur,

- “Systematic review of research on artificial intelligence applications in higher education – where are the educators?,” 1 Desember 2019, *Springer Netherlands*. doi: 10.1186/s41239-019-0171-0.
- [12] B. Kitchenham dan S. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” vol. 2, Jan 2007.
- [13] C. Guan, J. Mou, dan Z. Jiang, “Artificial intelligence innovation in education: A twenty-year data-driven historical analysis,” *Int. J. Innov. Stud.*, vol. 4, hal. 134–147, Des 2020, doi: 10.1016/j.ijis.2020.09.001.
- [14] A. Saputra, “Analisis Kendala Guru dalam Menerapkan Kurikulum Merdeka Belajar di Sekolah Dasar,” *J. Ilmu Multidisplin*, vol. 1, hal. 195–208, 2025.
- [15] F. Hidayat dan M. Nizar, “MODEL ADDIE (ANALYSIS, DESIGN, DEVELOPMENT, IMPLEMENTATION AND EVALUATION) DALAM PEMBELAJARAN PENDIDIKAN AGAMA ISLAM,” *J. Inov. Pendidik. Agama Islam*, vol. 1, no. 1, hal. 28–37, 2021, [Daring]. Tersedia pada: <https://journal.uinsgd.ac.id/index.php/jipal/article/view/14352>
- [16] A. Hasibuan, S. N. Hasanah, M. R. R. Haikal, dan W. Pratama, “Implementasi AI dalam Proses Produksi dan Dampaknya terhadap Transformasi Pengembangan SDM,” hal. 113–118, 2025.
- [17] Z. Fara *et al.*, “PENGEMBANGAN FLIPBOOK BERBASIS CONTEXTUAL TEACHING AND LEARNING (CTL) PADA MATERI PERTUMBUHAN DAN PERKEMBANGAN TUMBUHAN UNTUK MELATIHKAN KEMAMPUAN LITERASI DIGITAL KELAS XII SMA Development Of Flipbook Based on Contextual teaching and Learning (CTL) in Plant Growth and Development Topic to Train Digital Literacy Skill of 12 th Grade in High School,” 2024. [Daring]. Tersedia pada: <https://ejournal.unesa.ac.id/index.php/bioedu>
- [18] P. Syahada, “Kurikulum Merdeka : Dampak Peningkatan Beban Administrasi Guru terhadap Efektivitas Pembelajaran,” vol. 4, hal. 238–244, 2024.
- [19] F. Davis dan F. Davis, “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,” *MIS Q.*, vol. 13, hal. 319, Sep 1989, doi: 10.2307/249008.
- [20] A. Adestya dan P. Dewi, “Meta Analisis Kinerja Guru Berdasarkan Peran Kepemimpinan Transformasional,” *J. Literasi Digit.*, vol. 6, hal. 11–22, Mar 2026, doi: 10.54065/jld.6.1.2026.1065.
- [21] J. Garzón dan E. Patiño, “Systematic Review of Artificial Intelligence in Education : Trends , Benefits , and Challenges,” hal. 1–19, 2025.
- [22] H. Kusumaningrum, A. M. I. Nst, F. K. N. Fayza, dan A. C. Isfayanti, “INTEGRASI TEKNOLOGI DALAM MANAJEMEN PENDIDIKAN: PELUANG DAN TANTANGAN,” *AL-GAFARI J. Manaj. dan Pendidik.*, vol. 2, no. 3, hal. 263–277, 2024, [Daring]. Tersedia pada: <https://www.jurnal.zarilgapari.org/index.php/gafari>