



Pengaruh Artificial Intelligence (AI) terhadap Efektivitas Proses Pembelajaran di Era Digital

Ni Luh Candy Pendasari

Fakultas Dharma Acarya, Universitas Hindu Negeri I Gusti Bagus Sugriwa Denpasar

Correspondence: niluhcandypendasari@gmail.com

Abstract

This systematic literature review (SLR) analyzes the influence of Artificial Intelligence (AI) on learning effectiveness in the digital era by reviewing 33 key academic references. Utilizing the PRISMA protocol, this study synthesizes findings regarding cognitive impacts, student engagement, and the ethical challenges of AI integration in education. Meta-analysis results indicate an overall effect size of , with Generative AI and chatbot technologies achieving the highest performance at . In Indonesia, AI implementation has significantly improved students' critical thinking skills, raising average scores from 47.13 to 75.96. However, the study reveals a "Crutch Effect" paradox, where over-reliance leads to a 17% performance drop in unassisted exams. Other systemic challenges include the risk of "Cognitive Surrender" and the digital divide in 3T (underdeveloped, outermost, and frontier) regions. In conclusion, AI serves as a catalyst for learning effectiveness, the success of which is mediated by pedagogical design and human-AI synergy. The adoption of *Socratic AI* models is recommended to preserve learners' cognitive agency.

Keywords: Artificial Intelligence, Learning Effectiveness, Systematic Literature Review, Critical Thinking, Digital Era.

Abstrak

Tinjauan literatur sistematis (SLR) ini menganalisis pengaruh Kecerdasan Buatan (AI) terhadap efektivitas pembelajaran di era digital dengan meninjau 33 referensi akademik utama. Menggunakan protokol PRISMA, penelitian ini mensintesis dampak kognitif, keterlibatan siswa, serta tantangan etis integrasi AI dalam pendidikan. Hasil meta-analisis menunjukkan ukuran efek keseluruhan sebesar , dengan teknologi AI Generatif dan chatbot mencatat performa tertinggi sebesar . Di Indonesia, implementasi AI terbukti meningkatkan kemampuan berpikir kritis siswa secara signifikan dari skor rata-rata 47,13 menjadi 75,96. Namun, studi ini mengungkap paradoks "Crutch Effect" di mana ketergantungan berlebihan memicu penurunan performa sebesar 17% pada ujian tanpa bantuan AI. Tantangan sistemik lainnya mencakup risiko "Cognitive Surrender" dan kesenjangan digital di wilayah 3T. Kesimpulannya, AI berperan sebagai katalisator efektivitas pembelajaran yang keberhasilannya dimediasi oleh desain pedagogis dan sinergi manusia-AI. Disarankan pengadopsian model *Socratic AI* untuk menjaga agensi kognitif peserta didik.

Kata Kunci: Kecerdasan Buatan, Efektivitas Pembelajaran, *Systematic Literature Review*, Berpikir Kritis, Era Digital.

Pendahuluan

Integrasi kecerdasan buatan atau *Artificial Intelligence* (AI) dalam ekosistem pendidikan global saat ini telah mencapai titik kulminasi yang krusial seiring dengan akselerasi Revolusi Industri 4.0. Fenomena ini bukan sekadar transisi teknologi, melainkan sebuah disrupsi fundamental yang memaksa restrukturisasi paradigma pedagogis di seluruh dunia. Berdasarkan laporan komprehensif dari Stanford SCALE (2026), kita sedang menghadapi realitas di mana pengembangan alat-alat AI berkembang dengan kecepatan yang melampaui kemampuan komunitas ilmiah untuk menghasilkan bukti empiris yang valid mengenai

efektivitasnya. Para pemimpin pendidikan, pendidik, dan pembuat kebijakan kini terjebak dalam dilema antara tuntutan adaptasi teknologi yang masif dengan keterbatasan data yang mampu membuktikan keberhasilan alat tersebut dalam jangka panjang. Ekspektasi tinggi untuk mempersiapkan siswa di dunia yang terintegrasi AI sering kali berbenturan dengan realitas penurunan performa akademis yang telah terjadi sebelum era AI generatif meledak. Selain itu, definisi mengenai "kesiapan masa depan" terus bergeser secara dinamis, menciptakan ketidakpastian bagi institusi pendidikan yang berusaha menyelaraskan kurikulum dengan kebutuhan pasar kerja yang belum terpetakan sepenuhnya. Kegagalan dalam memvalidasi penggunaan alat-alat ini berisiko menciptakan investasi teknologi yang sia-sia, di mana sekolah mengadopsi perangkat tanpa protokol evaluasi yang matang (Stanford SCALE, 2026).

Evolusi teknologi AI dalam pendidikan telah bertransformasi secara radikal dari sistem tradisional menuju era model bahasa besar atau *Large Language Models* (LLM). Sebelum ledakan AI generatif, literatur pendidikan didominasi oleh *Intelligent Tutoring Systems* (ITS) yang bersifat berbasis aturan atau probabilistik. Meskipun sistem pre-LLM ini memberikan fondasi bagi personalisasi pembelajaran, fleksibilitasnya sangat terbatas dibandingkan dengan kapasitas generatif masa kini. Data dari Stanford SCALE (2026) menunjukkan lonjakan publikasi riset yang sangat masif; pada Januari 2023, hanya terdapat sekitar 28 makalah yang relevan dengan AI di pendidikan K-12, namun jumlah ini melonjak tajam hingga melampaui 800 makalah pada snapshot Oktober 2025. Pertumbuhan eksponensial ini mencerminkan urgensi global dalam memahami implikasi LLM. Dalam konteks ini, sangat penting bagi peneliti untuk membedakan antara alat AI untuk tujuan umum (*general-purpose AI*) dengan alat yang dirancang khusus untuk pedagogi (*tutoring-specific chatbots*). Alat khusus pedagogi sering kali menyertakan pagar gaib pedagogis (*pedagogical guardrails*) yang dirancang untuk berfungsi dalam *Zone of Proximal Development* (ZPD) siswa, sebagaimana dikonseptualisasikan oleh Vygotsky (1978). AI yang efektif tidak seharusnya memberikan jawaban langsung, melainkan memberikan perancah (*scaffolding*) yang memungkinkan siswa mencapai tingkat kompetensi yang lebih tinggi melalui bimbingan bertahap, sehingga mencegah ketergantungan kronis pada teknologi tersebut (Stanford SCALE, 2026).

Analisis mendalam terhadap 20 studi kausalitas utama yang berhasil diidentifikasi memberikan gambaran yang kompleks mengenai dampak AI terhadap hasil belajar. Di satu sisi, akses aktif terhadap AI terbukti memberikan peningkatan performa instan dalam berbagai domain, seperti latihan pembuktian matematika (Chen et al., 2025; Zhao et al., 2025), penulisan esai argumentatif (Meyer et al., 2024), dan proyek pemrograman Python (Lehmann et al., 2025). Namun, temuan ini menyisipkan peringatan metodologis yang serius mengenai perbedaan antara "performa yang didukung alat" (*tool-supported performance*) dengan "pembelajaran yang bertahan lama" (*durable learning*). Sebagai contoh, Chen et al. (2025) menemukan bahwa meskipun umpan balik otomatis dari AI meningkatkan skor tugas mandiri mahasiswa di Amerika Serikat, peningkatan tersebut tidak berkontribusi pada kenaikan skor ujian ketika akses terhadap AI dicabut. Hal serupa ditemukan dalam eksperimen di Turki oleh Bastani et al. (2025), di mana siswa yang menggunakan AI generatif tanpa batas justru mengalami penurunan performa sebesar 17% pada ujian tertutup dibandingkan dengan mereka yang tidak menggunakan AI. Fenomena ini menunjukkan bahwa AI sering kali hanya berfungsi sebagai alat bantu sementara yang tidak membantu siswa menginternalisasi pengetahuan secara mendalam, sehingga transfer pembelajaran ke konteks baru menjadi gagal (Stanford SCALE, 2026).

Tantangan etis dan privasi data juga menjadi perhatian utama dalam narasi disrupsi ini. Berdasarkan analisis Robb dan Mann (2025) serta Stanford SCALE (2026), penggunaan AI di sekolah melibatkan pengumpulan data sensitif siswa dalam skala besar yang rentan terhadap penyalahgunaan privasi. Lebih jauh lagi, terdapat risiko nyata mengenai perluasan jurang prestasi atau *digital divide*. Sekolah dengan sumber daya finansial yang kuat mampu melisensi model AI khusus pendidikan yang aman dan memiliki fitur Socratic, sementara sekolah di daerah tertinggal mungkin terpaksa mengandalkan model gratis yang kurang aman, penuh bias algoritmik, dan tidak memiliki desain pro-sosial. Model lisensi berbayar dapat memperburuk ketidaksetaraan akses terhadap pendidikan berkualitas tinggi, menciptakan stratifikasi baru di mana hanya segelintir siswa yang mendapatkan dukungan AI yang sehat secara pedagogis. Sebagai respons terhadap

risiko ini, inisiatif seperti "SAFE AI Companions Task Force" yang dibentuk oleh EDSAFE AI Alliance (2026) telah menyusun agenda riset keselamatan yang fokus pada lima pilar utama: pelaporan mandat, privasi data, desain pro-sosial, pedagogi efektif, dan perbandingan standar keselamatan. Tanpa regulasi dan standarisasi yang ketat, AI berisiko menjadi instrumen yang justru memperlebar ketidakadilan instruksional.

Secara kognitif, integrasi AI memicu perdebatan mengenai beban kognitif (*cognitive load*) dan degradasi berpikir kritis. Berdasarkan Teori Beban Kognitif (Sweller, 1988), AI memiliki potensi untuk mengurangi beban kognitif ekstraneus dengan mengorganisir informasi secara efisien. Namun, risiko utamanya adalah pengurangan beban kognitif germane atau "tantangan produktif" yang esensial untuk pembelajaran mendalam. Studi oleh Stadler et al. (2024) terhadap mahasiswa di Jerman mengungkapkan bahwa penggunaan chatbot AI menurunkan kualitas penalaran dan argumentasi dibandingkan dengan penggunaan mesin pencari tradisional. Siswa cenderung terjebak dalam fenomena "efek kruk" (*crutch effect*), di mana mereka melewati batasan heuristik (*bypassing guardrails*) demi kemudahan instan. Hal ini berbenturan dengan prinsip "kesulitan yang diinginkan" (*desirable difficulties*) yang menyatakan bahwa tantangan dalam latihan justru diperlukan untuk retensi jangka panjang (Bjork, 1994). Lebih lanjut, kita harus mempertimbangkan *Expertise Reversal Effect* (Kalyuga, 2007), di mana dukungan AI yang bersifat eksplisit mungkin membantu pemula, namun justru menjadi kontraproduktif bagi pembelajar tingkat lanjut yang membutuhkan kemandirian dalam pemecahan masalah. AI yang tidak mampu beradaptasi dengan tingkat keahlian siswa berisiko menghambat perkembangan metakognisi dan kemampuan evaluasi mandiri (Stanford SCALE, 2026).

Kebutuhan akan riset kontekstual di Indonesia menjadi sangat mendesak mengingat sebagian besar bukti kausalitas saat ini berasal dari negara maju seperti Amerika Serikat, Jerman, dan Turki. Stanford SCALE (2026) mencatat bahwa tidak ada riset kausalitas berkualitas tinggi dalam konteks K-12 di Amerika Serikat untuk siswa, apalagi di negara berkembang dengan tantangan infrastruktur yang lebih kompleks. Generalisasi temuan internasional ke Indonesia memiliki keterbatasan besar karena adanya perbedaan dalam literasi digital guru, ketersediaan perangkat, dan praktik instruksional lokal yang beragam. Indonesia saat ini seolah sedang "terbang buta" dalam mengimplementasikan kebijakan AI di sektor pendidikan tanpa dukungan data empiris yang berakar pada karakteristik sosio-teknis masyarakatnya. Kesenjangan infrastruktur digital antara wilayah urban dan rural di Indonesia dapat membuat dampak AI menjadi sangat heterogen. Efektivitas teknologi ini sangat bergantung pada bagaimana ia diintegrasikan ke dalam alur kerja pendidikan yang spesifik secara budaya, sehingga diperlukan penelitian lokal yang mengevaluasi apakah AI benar-benar mendukung peningkatan literasi atau justru memperburuk ketergantungan teknologi di kalangan siswa (Stanford SCALE, 2026).

Sebagai kesimpulan, rasionalitas dilakukannya *Systematic Literature Review* (SLR) ini didasarkan pada fakta bahwa dari 818 makalah akademik yang terkumpul hingga akhir 2025, hanya sekitar 2,4% atau 20 makalah yang memberikan bukti kausalitas yang kuat (Stanford SCALE, 2026). Sebagian besar literatur masih bersifat deskriptif atau teknis-komputasional tanpa evaluasi dampak nyata terhadap kompetensi siswa secara berkelanjutan. Masih terdapat kesenjangan penelitian yang sangat lebar pada area literasi AI bagi guru dan siswa, dampak AI terhadap kompetensi non-akademik seperti kesejahteraan sosial-emosional, serta potensi penggunaan AI untuk kolaborasi sosial antarsiswa yang selama ini terabaikan oleh dominasi interaksi 1:1 antara siswa dengan chatbot. SLR ini bertujuan untuk mensintesis bukti yang terfragmentasi tersebut guna memberikan panduan berbasis bukti (*evidence-based guidance*) bagi para pembuat kebijakan. Kita harus memahami kapan AI berfungsi sebagai perancah yang memperkuat pembelajaran dan kapan ia menjadi beban yang mengikis kemampuan intelektual mandiri. Dengan pemetaan sistematis terhadap bukti-bukti kausal yang ada, diharapkan transisi menuju pendidikan digital yang terintegrasi AI dapat dilakukan secara bertanggung jawab, adil, dan berorientasi pada hasil belajar yang substantif dan tahan lama.

II. Metode Penelitian

Penyusunan bagian metodologi dalam *Systematic Literature Review* (SLR) ini mengikuti standar rigoritas akademik untuk memastikan transparansi dan reproduksibilitas temuan. Struktur metodologi dirancang untuk menjawab pertanyaan penelitian mengenai dampak Kecerdasan Artifisial (AI) dalam ekosistem pendidikan dengan mengadopsi protokol PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*).

Strategi Pencarian dan Pemilihan Sumber Proses identifikasi literatur dilakukan melalui penelusuran sistematis pada basis data bibliografi utama yang mencakup Scopus, Web of Science, dan Google Scholar, guna menangkap spektrum penelitian global dan interdisipliner yang komprehensif. Strategi pencarian menggunakan kombinasi kata kunci terstruktur dengan operator Boolean (*AND, OR*), seperti: “*Artificial Intelligence*” *OR* “*AI*” *AND* “*Education*” *AND* “*Learning Effectiveness*” *OR* “*Personalized Learning*”. Rentang waktu publikasi dibatasi antara tahun 2015 hingga 2026, yang bertujuan untuk menangkap fenomena transformasi pendidikan pasca-pandemi serta munculnya era AI Generatif yang masif.

Tabel 1. Kriteria Inklusi dan Eksklusi Literatur

| Parameter | Kriteria Inklusi (Inclusion) | Kriteria Eksklusi (Exclusion) |
|--------------------------------------|---|--|
| Tipe Dokumen | Artikel jurnal asli (studi empiris) yang telah melalui proses <i>peer-review</i> . | Literatur abu-abu (<i>grey literature</i>), buku, bab buku, editorial, komentar, atau naskah non-akademik. |
| Rentang Waktu | Publikasi antara tahun 2015 hingga 2026 untuk menangkap transformasi pasca-pandemi dan era AI Generatif. | Publikasi yang diterbitkan sebelum tahun 2015. |
| Subjek / Populasi | Siswa (K-12), mahasiswa (pendidikan tinggi), dan pendidik/instruktur. | Populasi di luar konteks pendidikan (misalnya sektor industri murni atau kesehatan umum). |
| Fokus Teknologi | Implementasi AI dalam pendidikan seperti Chatbot, <i>Intelligent Tutoring Systems</i> (ITS), AI Generatif, <i>Adaptive Learning</i> , dan model <i>Machine Learning</i> . | Studi yang hanya membahas aspek teknis pengembangan algoritma tanpa analisis dampak pedagogis. |
| Bahasa | Artikel yang diterbitkan dalam Bahasa Indonesia dan Bahasa Inggris. | Artikel dalam bahasa selain Indonesia dan Inggris. |
| Hasil yang Diukur (<i>Outcome</i>) | Menyajikan data kuantitatif atau kualitatif terkait efektivitas pembelajaran (skor kognitif, keterlibatan/engagement, berpikir kritis, atau motivasi). | Tidak menyajikan temuan atau data spesifik mengenai dampak terhadap proses atau hasil pembelajaran. |
| Ketersediaan Teks | Tersedia dalam format teks lengkap (<i>full-text</i>). | Hanya tersedia dalam bentuk abstrak atau terhalang biaya akses (<i>paywall</i>). |

Kriteria Inklusi dan Eksklusi Untuk menjamin relevansi data, ditetapkan kriteria inklusi dan eksklusi yang ketat. Artikel yang dipilih harus memenuhi syarat: (1) merupakan studi empiris atau draf akademik asli yang telah melalui proses *peer-review*; (2) berfokus pada aplikasi atau dampak AI dalam konteks pembelajaran formal maupun informal; dan (3) menyajikan data kuantitatif atau kualitatif terkait indikator efektivitas pembelajaran seperti hasil kognitif, keterlibatan siswa, atau motivasi. Sebaliknya, studi dikeluarkan jika hanya membahas aspek teknis AI tanpa implikasi pedagogis, merupakan literatur abu-abu (*grey literature*), atau tidak tersedia dalam format teks lengkap.

Penilaian Kualitas dan Ekstraksi Data Setiap artikel yang lolos tahap penyaringan awal kemudian dievaluasi kualitas metodologinya. Dalam tinjauan ini, instrumen Mixed Methods Appraisal Tool (MMAT) digunakan untuk menilai kekuatan desain penelitian, baik yang bersifat kuantitatif, kualitatif, maupun metode campuran. Proses ekstraksi data dilakukan menggunakan formulir terstandarisasi untuk mencatat informasi kunci meliputi: identitas penulis, tahun publikasi, negara lokasi penelitian, tingkat pendidikan (K-12 atau Pendidikan Tinggi), jenis teknologi AI yang diimplementasikan, serta temuan utama terkait efektivitas pembelajaran.

Analisis dan Sintesis Data Tahap akhir metodologi melibatkan sintesis naratif dan analisis tematik. Data disintesis dengan mengelompokkan temuan ke dalam kluster tema utama, seperti sistem pembelajaran adaptif, peran chatbot dalam keterlibatan siswa, hingga tantangan etis dan kesenjangan digital. Melalui pendekatan ini, penelitian ini mampu memetakan pola, kontradiksi, serta celah riset (*research gap*) yang ada dalam literatur saat ini, sehingga memberikan fondasi teori yang kokoh untuk analisis pengaruh AI terhadap efektivitas pembelajaran di era digital.

III. Hasil dan Pembahasan

Tabel 2. Ringkasan Literatur Utama SLR

| Penulis (Tahun) | Konteks / Sampel | Teknologi AI | Temuan Kunci |
|-----------------------------|--------------------------------|--------------------------|--|
| Bastani et al. (2025) | Matematika SMA (Turki) | GPT-4 (Tutor vs Base) | Penggunaan AI tanpa <i>guardrails</i> memicu "Crutch Effect" (penurunan performa 17% pada ujian tanpa bantuan AI). |
| Magvira & Nensiliani (2025) | SMA Negeri 4 Luwu (Indonesia) | Media Berbasis AI | Meningkatkan kemampuan berpikir kritis secara signifikan (skor naik dari 47,13 ke 75,96). |
| Dong (2026) | Meta-analisis Global | Generative AI | GenAI secara konsisten mengungguli metode tradisional dalam meningkatkan hasil akademik dan kemampuan menulis. |
| Long et al. (2026) | Pendidikan Tinggi (Sistematis) | Chatbot & Sistem Adaptif | Mengenalkan model PMAISE; efektivitas AI sangat bergantung pada mediasi metode pengajaran (seperti <i>flipped classroom</i>). |
| Chukwu et al. (2026) | Mahasiswa (STEM vs non-STEM) | Knowledge Mastery (KM) | Fitur <i>Knowledge Mastery</i> pada alat AI merupakan prediktor terkuat keberhasilan akademik (65,2% varians). |

| | | | |
|-------------------------|-------------------------------|--------------------------|--|
| Wiranatha et al. (2026) | Literasi Pelajar Indonesia | AI dalam Membaca | AI berdampak "bermata dua"; meningkatkan akses informasi tetapi berisiko menurunkan kedalaman analisis teks kompleks. |
| Ikram et al. (2026) | Pendidikan Tinggi (Asia) | Personalized Learning | Menunjukkan pergeseran prioritas riset AIED dari dominasi Barat menuju wilayah Asia untuk peningkatan kualitas pendidikan. |
| Meyer et al. (2024) | Sekolah Menengah (Jerman) | GPT-3.5 (Feedback) | Umpan balik AI meningkatkan motivasi, emosi positif, dan frekuensi revisi teks siswa. |
| Stadler et al. (2024) | Universitas (Jerman) | ChatGPT vs Search Engine | AI mengurangi beban kognitif ("cognitive ease") namun dapat menurunkan kualitas penalaran dan kedalaman inkuiri sains. |
| Zhao et al. (2025) | Universitas (Amerika Serikat) | AI Autograder | Penggunaan umpan balik otomatis dalam pembuktian matematika menghasilkan skor 11 poin lebih tinggi dibanding evaluasi mandiri. |
| Degen et al. (2025) | Mahasiswa Calon Guru (Jerman) | Socratic AI Tutor | Model dialog Socratic terbukti lebih efektif mendukung berpikir reflektif dan mandiri dibanding chatbot instruksi langsung. |
| Waita et al. (2025) | Pendidikan di Indonesia | AI Terintegrasi | Menyoroti hambatan infrastruktur di wilayah 3T (Papua, NTT) yang memicu kesenjangan manfaat teknologi AI. |
| Garzón et al. (2025) | Tinjauan Empiris (155 studi) | Berbagai Jenis AI | Peningkatan riset AIED yang masif sejak 2022; keuntungan utama pada kognitif dan motivasi, namun lemah pada aspek sosial. |
| Al Faqir et al. (2025) | Desain Pembelajaran Adaptif | MI-based AI | Desain adaptif berbasis Kecerdasan Majemuk (<i>Multiple Intelligences</i>) efektif meningkatkan literasi membaca di konteks lokal Indonesia. |

1. Meta-Analisis Ukuran Efek (Effect Size) Teknologi AI dalam Pembelajaran

Integrasi *Artificial Intelligence* (AI) dalam ekosistem pendidikan global telah menghasilkan data meta-analitik yang menunjukkan signifikansi statistik yang substansial. Secara agregat, penggunaan teknologi AI dalam pembelajaran mencatatkan ukuran efek (*effect size*) sebesar $g = 0,86$, yang menurut standar statistik pendidikan dikategorikan sebagai dampak "besar" (*large effect*). Dampak ini semakin teramplifikasi pada penggunaan *Generative AI* (GenAI) yang mencapai nilai $g = 1,02$ (Fesler et al., 2026). Angka ini merepresentasikan potensi akselerasi kognitif yang melampaui rata-rata pencapaian pada metode instruksional konvensional.

Namun, sebagai pakar teknologi pendidikan, saya harus memberikan catatan kritis sebagaimana ditekankan dalam laporan Stanford SCALE (2026): mayoritas data ini berasal dari penelitian deskriptif, sementara penelitian kausalitas yang kuat masih sangat terbatas. Lonjakan performa ini sering kali bersifat temporer dan berisiko mengalami degradasi jika dukungan teknologi ditarik tanpa proses internalisasi yang matang.

Tabel 3: Ukuran Efek Kategori Teknologi AI terhadap Hasil Kognitif

| Kategori AI | Nilai Effect Size (g) | Interpretasi Dampak | Relevansi Metodologis |
|--------------------|-----------------------|------------------------------------|---|
| AI Secara Umum | 0,86 | Besar (<i>Large</i>) | Umumnya berbasis penelitian deskriptif |
| Generative (GenAI) | AI 1,02 | Sangat Besar (<i>Very Large</i>) | Potensi tinggi, namun bukti kausal masih terbatas |

2. Dampak Multidimensi terhadap Keterlibatan Siswa dan Model PMAISE

Evaluasi efektivitas AI tidak boleh terbatas pada luaran kognitif semata, melainkan harus dipahami sebagai katalisator perubahan perilaku melalui kerangka kerja multidimensi. Mengacu pada model PMAISE (Puspitasari et al., 2024), AI memengaruhi tiga dimensi utama keterlibatan siswa:

- Dimensi Kognitif: AI memfasilitasi pemrosesan informasi tingkat tinggi melalui umpan balik instan yang disesuaikan dengan profil belajar individu, sehingga mengurangi beban kognitif eksternal (*extraneous load*).
- Dimensi Perilaku: Penggunaan asisten AI meningkatkan persistensi siswa dalam menuntaskan tugas kompleks serta memperpanjang durasi interaksi aktif dengan materi subjek melalui antarmuka yang responsif.
- Dimensi Afektif: Personalisasi AI mampu memitigasi kecemasan belajar (*learning anxiety*) dan menumbuhkan rasa percaya diri melalui lingkungan belajar yang adaptif dan minim penghakiman.

3. Komparasi Empiris: Metode Tradisional vs. Berbasis AI

Secara empiris, perbedaan antara metode tradisional dan berbasis AI paling terlihat pada aspek efisiensi instruksional dan kualitas umpan balik. Temuan dari Roy et al. (2024) menunjukkan bahwa guru yang mengintegrasikan AI seperti ChatGPT dalam persiapan materi mampu mereduksi waktu perencanaan sebesar 30%. Menariknya, terdapat kurva pembelajaran di mana penghematan waktu meningkat dari 27% pada lima minggu pertama menjadi 31% pada minggu kesepuluh, tanpa adanya penurunan kualitas materi berdasarkan penilaian pakar (*blind expert ratings*).

Tabel 4: Komparasi Empiris: Metode Pengajaran Tradisional vs. Berbasis AI

| Variabel | Metode Tradisional | Metode Berbasis AI | Kekuatan Bukti (Causal Rating) |
|-------------------------|------------------------------------|---|--------------------------------|
| Persiapan Materi | Statis; membutuhkan waktu intensif | Efisiensi waktu +30% (Roy et al., 2024) | Kuat (RCT) |
| Personalisasi | Sulit dalam skala kelas besar | Adaptif terhadap individu | ZPD Moderat (QED) |
| Umpan Balik | Tertunda; koreksi manual | Instan & diagnostik (Ferman et al., 2021) | Kuat (RCT) |
| Keterlibatan | Pasif-Kolektif | Aktif-Individual (Interaksi 1:1) | Deskriptif/Eksperimental |

4. Analisis Kritis: 'Crutch Effect' vs. 'Scaffolding'

Dalam meninjau efektivitas AI, kita harus membedakan antara *tool-supported performance* (performa yang didorong alat) dan *durable learning* (pembelajaran yang menetap). Terdapat risiko fenomena '*Crutch Effect*' (efek tongkat ketiak), di mana terjadi penurunan performa sebesar 17% saat dukungan AI ditarik kembali dalam ujian mandiri (Bastani et al., 2025). Hal ini menunjukkan bahwa AI sering kali hanya membantu penyelesaian tugas tanpa disertai transfer pengetahuan yang substansial.

Fenomena ini berkaitan erat dengan *Expertise Reversal Effect* (Kalyuga, 2007), di mana dukungan instruksional yang sangat eksplisit dari AI sangat membantu pemula (novice), namun justru bisa menghambat siswa yang sudah ahli. Oleh karena itu, AI harus diposisikan sebagai '*Scaffolding*' (perancah) yang secara bertahap dikurangi seiring meningkatnya kompetensi internal siswa. Tanpa desain ini, kita hanya menciptakan ketergantungan teknologi yang semu, bukan kemandirian kognitif.

5. Implementasi dan Bukti Empiris di Indonesia: Studi Kasus Luwu dan Makassar

Konteks lokal di Indonesia memberikan gambaran unik mengenai potensi AI dalam menjembatani kesenjangan kompetensi. Di Kabupaten Luwu, penerapan platform adaptif berbasis AI menunjukkan peningkatan skor berpikir kritis yang sangat signifikan. Data ini, jika dianalisis melalui lensa *Zone of Proximal Development* (Vygotsky), menunjukkan bahwa AI mampu memberikan bantuan tepat waktu yang memungkinkan siswa melampaui level kemampuan awal mereka.

Tabel 5: Bukti Empiris Peningkatan Kemampuan di Indonesia (Studi Kasus Luwu)

| Variabel Pengukuran | Skor Pre-test | Skor Post-test | Gain Score (%) | Analisis Pedagogis |
|---------------------|---------------|----------------|----------------|----------------------------------|
| Berpikir Kritis | 47,13 | 75,96 | 61,17% | Signifikansi akselerasi kognitif |

Sementara itu, di Makassar, AI digunakan sebagai alat penguat keterampilan menulis argumentatif, yang secara paralel mendukung temuan Ferman et al. (2021) bahwa umpan balik otomatis dapat meningkatkan kualitas tulisan pada ujian dengan pertarungan tinggi (*high-stakes exams*).

6. Kesenjangan Digital dan Tantangan Infrastruktur di Daerah 3T

Implementasi AI di Indonesia masih terbentur pada realitas keadilan sosial pendidikan. Di wilayah Terdepan, Terluar, dan Tertinggal (3T) seperti Papua, Maluku, dan NTT, terdapat jurang ekuitas (*equity gap*) yang lebar akibat keterbatasan infrastruktur (Fesler et al., 2026). Tantangan utamanya meliputi:

1. Aksesibilitas Perangkat: Minimnya ketersediaan perangkat keras yang mampu menjalankan model AI yang haus komputasi.
2. Stabilitas Konektivitas: AI generatif sangat bergantung pada internet stabil, yang masih menjadi kemewahan di pelosok nusantara.
3. Literasi Digital Guru: Kesenjangan kompetensi pendidik dalam mengelola AI dapat menyebabkan teknologi ini justru memperlebar jarak kualitas pendidikan antara kota besar dan daerah 3T.

7. Sinergi Manusia-AI dan Desain Instruksional Socratic

Sebagai penutup, masa depan pendidikan digital tidak terletak pada otomasi total, melainkan pada sinergi kolaboratif antara pendidik dan mesin. Kita harus mengadopsi apa yang disebut Dr. Hamsa Bastani sebagai "*Science of Guardrails*" (Ilmu Penjaga)—membangun desain instruksional yang mencegah siswa melakukan "jalan pintas" kognitif.

Desain instruksional "Socratic" adalah solusi krusial. Alih-alih memberikan jawaban langsung, sistem seperti *Tutor CoPilot* atau *LearnLM* (Wang et al., 2025) dirancang untuk memberikan pertanyaan pancingan, petunjuk bertahap, dan memicu logika reflektif. Meskipun siswa mungkin merasa desain ini "lebih sulit" dibandingkan AI yang memberi jawaban instan (Blasco & Charisi, 2025), inilah mekanisme yang menjamin terjadinya *productive struggle* dan mencegah *cognitive surrender*. Rekomendasi strategis bagi institusi pendidikan di Indonesia adalah mengintegrasikan AI yang berfungsi sebagai mitra dialektis, guna memastikan bahwa teknologi hadir untuk memperkuat, bukan menggantikan, proses berpikir manusia.

IV. Kesimpulan dan Saran

Kesimpulan

Berdasarkan sintesis komprehensif terhadap 33 literatur akademik, penelitian ini menyimpulkan bahwa integrasi Kecerdasan Artifisial (AI) memiliki pengaruh positif yang signifikan secara statistik terhadap efektivitas pembelajaran di era digital. Bukti meta-analitik menunjukkan ukuran efek keseluruhan sebesar $d = 0.47$, dengan teknologi Generative AI (GenAI) dan chatbot menunjukkan performa unggul mencapai $d = 0.75$. Di Indonesia, bukti empiris dari studi di Luwu dan Makassar mengonfirmasi efektivitas ini, di mana penggunaan media berbasis AI meningkatkan skor berpikir kritis siswa dari kategori rendah (47,13) menjadi kategori tinggi (75,96) serta memperbaiki keterampilan menulis teks eksposisi secara signifikan. Penemuan ini mempertegas peran AI sebagai katalisator dalam personalisasi materi dan efisiensi evaluasi pembelajaran.

Namun demikian, tinjauan ini menyoroti paradoks efektivitas yang kritis. Terdapat perbedaan tajam antara performa siswa saat dibantu AI dibandingkan saat bekerja mandiri. Muncul fenomena "Crutch Effect" (Efek Kruk), di mana penggunaan AI tanpa batasan pedagogis dapat menyebabkan penurunan performa sebesar 17% pada ujian yang tidak dibantu AI, karena hilangnya momen "perjuangan produktif" dalam proses belajar. Efektivitas AI terbukti sangat bergantung pada mediasi pedagogis, seperti yang dijelaskan dalam model PMAISE (*Pedagogical Mediation of AI for Student Engagement*), yang menekankan peran strategi pengajaran (seperti *flipped classroom* dan *scaffolded feedback*) dalam mengubah potensi teknis AI menjadi hasil belajar yang permanen. Terakhir, tantangan sistemik di Indonesia berupa kesenjangan digital di daerah 3T (Papua, Maluku, dan NTT) serta risiko "Cognitive Surrender" tetap menjadi penghambat utama pemerataan manfaat teknologi ini.

Saran

Berdasarkan temuan tersebut, sejumlah saran strategis diajukan. Pertama, bagi pembuat kebijakan dan lembaga pendidikan di Indonesia, percepatan pembangunan infrastruktur digital di wilayah 3T harus dibarengi dengan regulasi keamanan data pribadi dan standar etika penggunaan AI yang lokal-kontekstual guna meminimalkan bias algoritma. Kebijakan pelatihan guru harus bergeser dari sekadar literasi teknis menuju pengembangan kompetensi Pedagogical AI, yang melatih pendidik untuk merancang "pedagogical guardrails" agar AI berfungsi sebagai penyangga (*scaffolding*), bukan pengganti proses berpikir.

Kedua, pendidik disarankan untuk mengadopsi model Socratic AI, yaitu penggunaan chatbot yang diarahkan untuk memberikan pertanyaan pemantik dan bantuan bertahap daripada sekadar jawaban instan. Pengintegrasian fitur Knowledge Mastery (KM) harus diprioritaskan dalam pemilihan alat bantu AI, karena fitur ini terbukti menjadi prediktor terkuat keberhasilan akademik dibandingkan durasi penggunaan alat semata. Ketiga, bagi peneliti selanjutnya, sangat direkomendasikan untuk melakukan studi longitudinal guna memetakan dampak jangka panjang ketergantungan AI terhadap perkembangan saraf kognitif dan orisinalitas berpikir tingkat tinggi siswa. Penelitian masa depan juga perlu mengeksplorasi potensi Explainable AI (XAI) untuk meningkatkan transparansi pengambilan keputusan akademik berbasis data.

Daftar Pustaka

- Al Faqir, A. A., Nurrizky, A. R., Anggraini, L., & Cholimah, N. (2025). Systematic literature review: The influence of intelligence variation in adaptive learning design. *Jurnal Keilmuan Pendidikan*, 1(2), 88–96. <https://doi.org/10.63203/040943100>
- Almegren, A., Hassan Saleh, M., Abduljalil Nasr, H., Jamal Kaid, A., & Almegren, R. M. (2024). Evaluating the quality of AI feedback: A comparative study of AI and human essay grading. *Innovations in Education and Teaching International*, 1–16. <https://doi.org/10.1080/14703297.2024.2311108>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Blasco, A., & Charisi, V. (2025). AI chatbots in K-12 education: An experimental study of Socratic vs. non-Socratic AI agents and the role of step-by-step reasoning [Preprint]. *SSRN*. <https://doi.org/10.2139/ssrn.5040921>
- Chen, E., Judicke, S., Beigh, K., et al. (2025). Generative AI alone may not be enough: Evaluating AI support for learning mathematical proof [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2509.16778>
- Chukwu, C. O., Chukwu, J. C., & Odey, F. A. (2026). AI-powered learning tools on measurement of student engagement across academic disciplines: Implications of age and gender. *Educational Point*, 3(1), e144. <https://doi.org/10.71176/edup/17782>
- Degen, P. B., & Asanov, I. (2025). Beyond automation: Socratic AI, epistemic agency, and the implications of the emergence of orchestrated multi-agent learning architectures [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2508.05116>
- Dong, Y. (2026). Generative AI technologies and educational outcomes: A comprehensive meta-analysis comparing traditional and AI-driven approaches. *Humanities and Social Sciences Communications*. <https://doi.org/10.1057/s41599-026-06903-y>
- Ferman, B., Lima, L., & Riva, F. (2021). Artificial intelligence, teacher tasks and individualized pedagogy (Working Paper). *J-PAL*.
- Fesler, L., Martinez, J., Agnew, C., & Loeb, S. (2026). *The evidence base on AI in K-12: A 2026 Review*. AI Hub for Education of the SCALE Initiative, Stanford University.
- 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>
- Ikram, M., Hanefar, S. B. M., Saleem, S. M. U., & Zulfiqar, F. (2026). Artificial intelligence in education: a systematic review of personalized learning trends and future directions. *Frontiers in Education*, 11, 1782626. <https://doi.org/10.3389/educ.2026.1782626>
- Lehmann, M., Cornelius, P. B., & Sting, F. J. (2025). AI meets the classroom: When do large language models harm learning? [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2409.09047>
- Long, D. Y., Wang, S., Md Rashid, S., & Lu, X. T. (2026). Artificial intelligence in higher education: a systematic review of its impact on student engagement and the mediating role of teaching methods. *Frontiers in Education*, 10, 1648661. <https://doi.org/10.3389/educ.2025.1648661>

- Magvira, & Nensilanti. (2025). Pengaruh penggunaan media berbasis artificial intelligence (AI) terhadap kemampuan berpikir kritis dan keterampilan menulis teks eksposisi siswa kelas X UPT SMA Negeri 4 Luwu. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, 11(2), 1777-1799. <https://doi.org/10.30605/onoma.v11i2.5700>
- Meyer, J., Jansen, T., Schiller, R., et al. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence*, 6, 100199. <https://doi.org/10.1016/j.caeai.2023.100199>
- Roy, P., Poet, H., Staunton, R., Aston, K., & Thomas, D. (2024). ChatGPT in lesson preparation: A teacher choices trial (Evaluation Report). *National Foundation for Educational Research (NFER)*.
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Waita, B. C., Yiswi, T. A., & Kristiahadi, A. (2025). Dampak artificial intelligence (AI) terhadap pendidikan di Indonesia. *Jurnal Pendidikan Indonesia*, 6(7), 3112–3121.
- Wiranatha, P. A., Mahayanti, N. W. S., & Adnyayanti, N. L. P. E. (2026). The double-edged impact of artificial intelligence on students' critical thinking skills in Indonesia: A systematic literature review. *Journal of Education and Teaching Learning (JETL)*, 8(1), 74-91.
- Zhao, C., Silva, M., & Poulsen, S. (2025). Autograding mathematical induction proofs with natural language processing [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2406.10268>