



## PERAN AI DALAM DUNIA PENDIDIKAN DI SEKOLAH DASAR

Meilan Saiful Suleman<sup>1</sup>, Fitrianti Hasan<sup>2</sup>, Delfi Suharty Kadir<sup>3</sup>, Dian Novian<sup>4</sup>,  
Abd. Azis Bouty<sup>5</sup>

<sup>1,2,3,4,5</sup>Magister Pendidikan Teknologi Pendidikan, Universitas Negeri Gorontalo

### Informasi Artikel

Riwayat Artikel:

Diterima: 5 Desember 2025

Revisi: 30 Januari 2026

Diterima: 9 Februari 2026

Diterbitkan: 30 April 2026

Keywords:

Artificial intelligence; primary school; human-in-the-loops

Kata Kunci:

Kecerdasan buatan, sekolah dasar, human-in-the-loops

DOI :

10.31932/jpdp.v12i1.5800

Surel Korespondensi:

715525009@mahasiswa.ung.ac.id

### Abstract

This study aimed to compare the contributions and limitations of artificial intelligence in elementary schools. The method used was a systematic literature review with light bibliometric analysis of articles published between 2020 and 2025. The review focused on differentiated learning, assessment and formative feedback, explainable AI, human-in-the-loop mechanisms, and data governance. Following a staged screening process based on PRISMA guidelines, 72 articles were analyzed thematically and mapped using keyword co-occurrence. The results indicate that most studies reported improvements in literacy and numeracy when artificial intelligence was used to support instructional strategies, with a moderate average effect size. Automated feedback systems reduced teachers' routine workload by approximately 24 minutes per week and improved the frequency and accuracy of responses. While the detection of misconceptions was considered adequate, the level of agreement between AI and teacher assessment ranged from moderate to high. However, several challenges remain, including uneven fairness audits, limited teacher control features, insufficient system transparency, and low levels of open science practices. In conclusion, artificial intelligence is most effective when it reinforces teachers' pedagogical roles and supports students' intellectual responsibility. These findings have implications for policy development, particularly in child-friendly explainability, strict data governance, and effective human-in-the-loop procedures.

### Abstrak

Tujuan penelitian ini adalah untuk membandingkan kontribusi dan kelemahan kecerdasan artifisial di sekolah dasar. Metode yang digunakan adalah tinjauan literatur sistematis dengan analisis bibliometrik ringan terhadap artikel periode 2020–2025. Fokus tinjauan adalah pembelajaran yang berbeda, asesmen dan umpan balik formatif, kecerdasan buatan yang dapat dijelaskan, mekanisme putaran manusia, dan tata kelola data. Setelah penyaringan bertahap yang dilakukan sesuai dengan pedoman PRISMA, 72 artikel dianalisis secara tematik dan dipetakan menggunakan ko-kemunculan kata kunci. Hasil menunjukkan bahwa sebagian besar penelitian melaporkan peningkatan literasi dan numerasi ketika kecerdasan artifisial digunakan sebagai penguat strategi pendidikan. Dengan dampak rata-rata pada tingkat sedang, hasil menunjukkan peningkatan literasi dan numerasi. Sistem umpan balik otomatis mengurangi beban rutin guru sekitar 24 menit per minggu dan meningkatkan frekuensi dan ketepatan respons. Sementara deteksi miskonsepsi dianggap memadai, tingkat kesepakatan antara kecerdasan artifisial dan penilaian guru dinilai sedang hingga tinggi. Namun demikian, audit keadilan tidak merata, fitur kontrol guru yang terbatas, penjelasan sistem yang kurang, dan keterbukaan sains yang rendah. Penelitian menunjukkan bahwa kecerdasan artifisial paling efektif ketika memperkuat peran pedagogis guru dan menegakkan tanggung jawab intelektual siswa. Ini dapat berdampak pada kebijakan tentang penjelasan yang ramah anak, tata kelola data yang ketat, dan prosedur manusia dalam putaran yang efektif.

*This is an open access article under the CC BY-SA license.*

Copyright © 2026 by Author. Published by STKIP Persada Khatulistiwa



### Pendahuluan

Selama dua tahun terakhir, percakapan tentang Gen AI di pendidikan bergerak di antara antusiasme dan kehati-hatian.

Sejumlah pakar lintas negara melihat peluang di desain pembelajaran, pengaturan belajar, umpan balik, sampai asesmen. Namun, mereka juga menegaskan bahwa risiko untuk

timbulnya masalah baru akan muncul apabila adopsi dilakukan secara terburu-buru dengan bukti yang lemah dan desain yang berpusat pada manusia. Ini terutama berlaku karena banyak komentar bergantung pada opini daripada data kelas. Di tingkat makro, kerangka "Education 5.0" menegaskan orientasi human-centric dan keberlanjutan, dan menghubungkan kurikulum, kompetensi kerja masa depan, dan kolaborasi antara manusia dengan AI untuk nilai sosial yang lebih luas (Chigbu & Makapela, 2025). Di tingkat sekolah dasar, kebutuhan mendesak tersebut tampak sebagai kebutuhan diferensiasi belajar sejak dini, umpan balik formatif tepat waktu, serta dukungan untuk guru yang menghadapi beban administrasi dan heterogenitas kemampuan siswa. Tapi prinsip human-in-the-loop yang kuat, tata kelola data yang aman, dan algoritme yang kuat menjadi tiga tema yang muncul sebagai prioritas bagi pakar lintas dunia untuk menggunakan AI pada anak usia sekolah (Ifenthaler et al., 2024).

Di sisi lain, pembicaraan kritis mengingatkan agar integrasi GenAI tidak menurunkan agensi berpikir

siswa, melainkan mengukuhkan budaya belajar yang menumbuhkan tanggung jawab intelektual (Costa & Murphy, 2025). Tinjauan sistematis berskala besar memetakan empat ranah utama : tutoring/adaptif, asesmen dan manajemen cerdas, profiling/prediksi, dan produk baru. Tinjauan ini juga menunjukkan banyak ruang kosong, termasuk kebutuhan dasar teori yang lebih kokoh. Panel Delphi internasional membuat tiga pekerjaan rumah: keadilan, keterpercayaan algoritme, dan privasi. Sarana mereka sederhana tetapi sulit: memastikan bahwa manusia tetap "di dalam loop". Pembicaraan tentang XAI mengingatkan bahwa, dari perspektif keterjelasan sistem, itu relatif dan memerlukan persetujuan lintas pemangku kepentingan, termasuk pelajar.

Perspektif filsafat pendidikan bahkan mengangkat risiko "membungkam" proses berpikir bila GenAI hanya dipakai secara instrumental. Bukti empiris awal memang menunjukkan sebagian siswa meraih nilai lebih baik dan bersikap positif, tetapi ada sinyal penurunan visibilitas serta keterlibatan guru pada

kondisi tertentu. Gagasannya lalu melebar: ada yang mendorong kemitraan manusia-AI sepanjang hayat belajar, ada pula kritik dekolonial yang membongkar jejak Eurocentrisme dalam tata kelola global. Intinya, potensinya besar; tapi untuk konteks sekolah dasar, bukti kelas riil masih tipis. Uji coba yang ketat, AI yang ramah anak, mekanisme human-in-the-loop yang berfungsi, dan evaluasi keadilan dan privasi yang berpijak pada konteks lokal adalah tugas penting ke depan.

Walaupun sangat mungkin, ada perbedaan antara solusi AI yang tersedia dan kebutuhan kelas dasar yang sebenarnya. Pertama, tidak banyak bukti kausal tentang bagaimana AI memengaruhi literasi dan numerasi di sekolah dasar. Akibatnya, penelitian dan kebijakan sebagian besar berfokus pada pendidikan tinggi atau konteks umum (Wang et al., 2024). Kedua, proses penerapan biasanya lebih cepat daripada gagasan pedagogis. Akibatnya, adopsi impulsif tanpa penilaian etika, ekosistem, dan efektivitas yang cukup (Giannakos et al., 2025). Ketiga, AI mungkin menjanjikan, tetapi tidak dapat

menyelesaikan masalah black box, bias, dan akuntabilitas. Transparansi dan penjelasan yang dibutuhkan siswa dan guru sekolah dasar masih belum terpenuhi (Farrow, 2023). Selain itu, penelitian tentang etika menunjukkan bahwa kerangka normatif AI dominan di Eropa, sehingga kurang peka terhadap konteks lokal dan ketimpangan bahasa dan infrastruktur. Ini adalah masalah yang sangat penting bagi sekolah dasar di negara lain (Nemorin, 2024). Sementara itu, temuan Delphi lintas pakar menunjukkan bahwa tiga hal yang paling penting dan paling sulit untuk ditangani adalah keadilan dan equity, privasi, dan keandalan algoritme, merupakan hal-hal terimplikasi jauh lebih kompleks saat menyangkut data anak (Ifenthaler et al., 2024). Dengan demikian, ada implementation gap antara peluang AI dan kesiapan pedagogi, tata kelola, serta kapasitas guru di sekolah dasar.

Beragam telaah terbaru menempatkan GenAI sebagai peluang sekaligus peringatan untuk pendidikan. Komentar pakar lintas negara menyoroti potensi pada desain pembelajaran, pengaturan belajar, umpan balik, dan asesmen, tetapi

mengingatkan bahaya adopsi tergesa-gesa tanpa bukti kuat dan tanpa desain berpusat manusia. Empat rumpun aplikasi—tutoring/adaptif, asesmen dan manajemen cerdas, profiling/prediksi, dan produk baru—diidentifikasi dalam penelitian sistematis-bibliometrik skala besar. Selain itu, penelitian tersebut menunjukkan bahwa masih ada area yang kurang diperhatikan yang memerlukan landasan teori yang lebih kokoh.

Selain itu, penelitian Delphi internasional mengidentifikasi tiga masalah utama: privasi, keterpercayaan algoritme, dan keadilan. Rekomendasi prinsip kebijakannya adalah “human-in-the-loop Untuk menjelaskan model, ulasan penting tentang XAI menekankan bahwa penjelasan harus relatif dan dimediasi oleh berbagai pemangku kepentingan, termasuk pelajar. Analisis teoritis dalam filsafat pendidikan memperingatkan tentang bahaya “membungkam” proses berpikir ketika GenAI digunakan secara instrumental semata. Bukti empiris awal menunjukkan penurunan visibilitas dan keterlibatan guru dalam kondisi tertentu, tetapi

juga menunjukkan peningkatan capaian belajar siswa dan sikap positif siswa di beberapa mata pelajaran saat menggunakan AI. Kritik dekolonial mengungkapkan nilai-nilai Eurocentrik dalam tata kelola global AI pendidikan, sementara perspektif profesi kesehatan menawarkan kerangka “pasangan belajar manusia-AI” sepanjang hayat untuk kerja sama dan diskusi nyata.

Singkatnya, literatur baru-baru ini telah membahas potensi dan bahaya penggunaan AI dalam pendidikan, tetapi ada beberapa kekosongan penting mengenai konteks sekolah dasar. Pertama, ada sedikit bukti berskala dan in situ yang tersedia tentang pengaruh kecerdasan buatan terhadap capaian literasi, numerasi, dan pembelajaran self-regulated pada anak usia sekolah dasar (Wang et al., 2024; Giannakos et al., 2025). Kedua, desain XAI yang ramah anak, yang dapat dipahami siswa dan membantu guru menjelaskan proses algoritmik, hampir tidak diuji di kelas dasar. Ketiga, efisiensi privasi dan kepercayaan digariskan sebagai prioritas utama dalam agenda kebijakan.

Namun, belum ada bukti empiris yang menjelaskan bagaimana hal ini dapat diterapkan ke praktik kelas dasar, seperti minimisasi data, consent yang signifikan bagi anak dan orang tua, dan audit bias pada tugas kurikuler (Ifenthaler et al., 2024). Keempat, bias Eurocentric dalam etika AI pendidikan menimbulkan kesenjangan konteks—bahasa, budaya, dan infrastruktur—yang dapat memengaruhi keberterimaan serta keadilan manfaat AI bagi siswa sekolah dasar di luar pusat produksi pengetahuan (Nemorin, 2024).

Penelitian ini bisa mengisi celah pengetahuan dan menempatkan sekolah dasar sebagai panggung utama. Fokusnya adalah menguji bagaimana AI dapat memberi umpan balik formatif dan latihan yang dipersonalisasi untuk literasi dan numerasi, langsung di dalam dinamika kelas. Supaya keputusan sistem tidak terasa “kotak hitam”, rancangan yang dipakai menyertakan XAI ramah anak sehingga siswa dan guru bisa menangkap alasan di balik setiap rekomendasi—selaras dengan gagasan *assessment as learning* (Farrow, 2023). Dampak intervensi tidak dinilai sepihak: peneliti

memadukan pendekatan kuasi-eksperimental, *classroom analytics*, dan etnografi kelas untuk melihat efek pada hasil belajar, beban kerja guru, serta indikator keadilan (Ifenthaler et al., 2024; Giannakos et al., 2025). Di saat yang sama, desain kurikuler dan antarmuka ditinjau melalui lensa dekolonial guna meminimalkan bias bahasa dan budaya yang kerap luput dalam penerapan AI di kelas dasar (Nemorin, 2024). Sejauh yang kami ketahui, kombinasi fokus pada jenjang sekolah dasar, penggunaan XAI yang ramah anak, pengukuran keadilan, dan evaluasi di banyak situs seperti ini belum terdokumentasi dengan memadai dalam literatur AI pendidikan yang masih didominasi studi di pendidikan tinggi (Wang et al., 2024).

Sejalan dengan kebutuhan tersebut, penelitian ini diarahkan untuk memotret secara menyeluruh bagaimana AI bekerja di kelas dasar. Pertama, penelitian ini menilai pengaruh pembelajaran terpersonalisasi dan umpan balik formatif berbantuan AI terhadap capaian literasi serta numerasi, lalu membandingkannya dengan praktik pengajaran biasa. Kedua, penelitian ini

menguji sejauh mana solusi ini dapat diterima dan membantu kerja sehari-hari: dari sudut pandang guru (kualitas umpan balik, kemudahan dan efisiensi asesmen) maupun siswa (pengalaman belajar dan kemandirian). Ketiga, penelitian ini mengevaluasi dimensi XAI—apakah penjelasan sistem benar-benar dipahami oleh siswa dan guru serta mampu memperkuat *assessment as learning*. Keempat, penelitian ini menganalisis keadilan dampak: apakah manfaat AI tersebar merata di berbagai kelompok, misalnya berdasarkan kemampuan awal, status sosial ekonomi, atau bahasa yang digunakan di rumah. Kelima, temuan-temuan tersebut menjadi dasar untuk menyusun pedoman tata kelola yang menekankan privasi, keterpercayaan, dan keadilan bagi konteks sekolah dasar.

### **Metode**

Penelitian ini menggunakan Systematic Literature Review (SLR) yang diperkaya analisis bibliometrik ringan. Metode SLR dipilih karena mampu memberikan kerangka penelaahan yang jelas, penting, dan replikabel untuk mengidentifikasi,

menilai, dan menyintesis bukti lintas disiplin. Ini sesuai dengan kecepatan berkembangnya kecerdasan artifisial dalam bidang pendidikan. Menurut rekomendasi studi AIED terbaru, elemen bibliometrik ditambahkan untuk menggambarkan konsep dan tren terbaru. Ini terutama mencakup *co-occurrence* kata kunci dan telaah *meta-data* dasar. Selain itu, fokus pada *explainability*, *humans-in-the-loop*, dan tata kelola menjadi justifikasi metodologis agar kualitas bukti serta implikasi etis dapat dinilai secara terstruktur.

Penelitian ini tidak melibatkan partisipan manusia secara langsung, oleh karena itu, unit analisis penelitian ini adalah dokumen ilmiah: artikel jurnal bereputasi (*peer-reviewed*) dan makalah konferensi. Rentang waktu 2020–2025 dipilih untuk menangkap pergeseran dari AIED pra-GenAI ke era LLM. Penelusuran dilakukan pada Scopus, Web of Science, ScienceDirect (Elsevier), SpringerLink, Wiley Online Library, IEEE Xplore, Taylor & Francis Online, dan SAGE Journals. Kriteria inklusi meliputi publikasi *peer-reviewed* yang berfokus pada AI dalam konteks sekolah

dasar/elementary/primary (intervensi, adopsi, asesmen, XAI, analitik belajar, keadilan) dalam bahasa Inggris atau Indonesia pada rentang tahun tersebut. Grey literature, kajian murni teknis tanpa konteks pendidikan, serta studi jenjang menengah/tinggi tanpa implikasi ke sekolah dasar dikeluarkan. Purposeful sampling digunakan dengan pertimbangan relevansi topik dan kualitas outlet (Q1/Q2), dilengkapi strategi forward-backward citation chasing. Ukuran sampel akhir ditetapkan melalui penyaringan bertahap—judul, abstrak, hingga teks penuh—sampai tercapai saturasi tematik. Tiga orang penelaah independen melakukan seleksi dan ekstraksi untuk meminimalkan bias.

Pada penelitian ini memanfaatkan penelusuran sistematis yang memanfaatkan kombinasi Boolean keywords. Contoh inti: “(artificial intelligence OR AI OR machine learning OR generative AI OR large language model\*) AND (primary OR elementary) AND (education OR school) AND (assessment OR feedback OR tutoring OR explainable AI OR XAI OR equity OR privacy)”. Perangkat lunak tertentu digunakan sebagai

referensi deduplikasi untuk mengelola semua sitasi. Metadata (penulis, tahun, jurnal, DOI), konteks (negara, jenjang), desain/metode, jenis AI, luaran (literasi/numerasi, SRL, beban kerja guru), komponen XAI/etika, dan keterbatasan studi adalah semua bagian dari ekstraksi data terstruktur. Keandalan proses dijaga melalui pilot screening 30–50 artikel, double-screening oleh tiga orang peneliti, penghitungan Cohen’s  $\kappa \geq 0,70$  untuk seleksi dan pengodean, serta audit trail keputusan inklusi/eksklusi.

Dimulainya penelitian diawali dengan perumusan pertanyaan riset tentang peran AI dalam pembelajaran dan asesmen di sekolah dasar serta operasionalisasi XAI, privacy-trust-equity, dan humans-in-the-loop. Agar dapat direplikasi, strategi pencarian kemudian dirancang untuk lintas basis data besar. Screening berlangsung bertingkat: pertama pada judul-abstrak untuk mengonfirmasi cakupan sekolah dasar, disusul telaah teks penuh untuk menilai kesesuaian fokus dan intervensi AI, lalu penilaian kualitas terkait kekuatan metodologi, kejelasan luaran, dan transparansi data. Selanjutnya, data disintesiskan melalui analisis tematik untuk isi dan

substansi, serta analisis bibliometrik ringan untuk peta topik dan keterkaitan konsep. Ini dilakukan dengan menggunakan coding sheet terstandar. Pada tahap akhir, validasi dilakukan melalui debriefing internal antar rekan dan pengujian sensitivitas atas keputusan penting yang berkaitan dengan inklusi. Proses ringkasnya: definisi RQ → strategi pencarian → penelusuran multibasis data → deduplikasi → screening judul-abstrak → screening teks penuh → kaji kualitas → ekstraksi → analisis tematik + bibliometrik → sintesis & pelaporan → validasi dengan lampiran alur PRISMA.

Analisis isi tematik. Mengikuti Braun & Clarke (2006), analisis ini mengekstrak pola lintas studi pada: domain aplikasi (tutoring/adaptif, asesmen cerdas, profil/prediksi, produk baru), hasil belajar (literasi/numerasi, SRL), praktik guru/beban kerja, serta XAI & tata kelola (privasi, keadilan, keandalan). Pengodean dilakukan oleh tiga peneliti dan perbedaan diselesaikan melalui konsensus.

Analisis bibliometrik ringan. Dilakukan co-occurrence mapping atas kata kunci penulis/Keywords

Plus untuk memetakan lanskap tema dan irisan topik (seperti feedback, pembelajaran khusus, kesetaraan, XAI). Ini dilakukan dengan mengacu pada praktik AIED terbaru. Implementasi memanfaatkan bibliometrix/biblioshiny (R) dan VOSviewer untuk jaringan ko-kata dan peta densitas. Hasil bibliometrik berfungsi sebagai scaffold yang menguatkan interpretasi tematik—bukan analisis kuantitatif utama.

Dikarenakan tidak ada partisipan manusia dalam penelitian ini, izin etik formal tidak diperlukan. Meski demikian, penelitian mematuhi etika akademik: sitasi yang akurat, penghormatan hak cipta, dan pemanfaatan sumber berlisensi. Saat menafsirkan isu privasi dan keadilan dalam studi primer yang ditelaah, peneliti menerapkan kehati-hatian dan konteks, sejalan dengan rekomendasi literatur tentang XAI dan tata kelola AI di pendidikan.

### **Hasil dan Pembahasan**

Gambaran Awal: Dari tahun 2020 hingga 2025, penelusuran menghasilkan 1.742 rekaman awal dari delapan basis data utama. Dari 1.231 artikel yang tersisa setelah

deduplikasi, 214 dikirim untuk peninjauan full-text, dan 78 studi memenuhi kriteria inklusi untuk sintesis (Gambar 1. PRISMA flow). Tahun publikasi (n=78): median 2022; IQR 2021–2024; bidang/mata pelajaran: literasi bahasa (34%), matematika/numerasi (29%), STEM campuran (18%), dan lintas mata pelajaran (19%). 4. Tipe studi (n=78): eksperimental/kuasi-eksperimental (26%), survei (21%), desain-berbasis riset/DBR (14%), studi kasus (12%), analitik pembelajaran/EDM (11%), campuran (9%), ulasan/posisi (7%). 5. Wilayah (n=78): seluruh sekolah dasar Pelaporan efek terstandar (Hedges  $g$ /Cohen  $d$ ) tersedia pada  $k=22$  (28%); saran untuk artikel Q1 adalah 62%; saran untuk artikel Q2 adalah 26%; dan saran untuk artikel lain adalah 12%.

Hasil utama disajikan sesuai dengan tujuan penelitian dan pertanyaan penelitian: deskripsi disajikan tanpa penafsiran mendalam sesuai urutan RQ. RQM1. Implikasi kecerdasan buatan pada hasil belajar dan praktik kelas di sekolah dasar. 2.1 Hasil belajar (literasi dan numerasi) : Proporsi penelitian yang melaporkan peningkatan signifikan pada  $\geq 1$

indikator hasil belajar adalah 56/78 (72%). Metrik efek terstandar (subset  $k=22$ ): median  $g=0,32$  (IQR 0,20–0,45; rentang 0,05–0,71). Untuk mapel, median  $g=0,30$  untuk literasi ( $k=12$ ) dan median  $g=0,30$  untuk numerasi ( $k=8$ ).

RQ2. Penggunaan AI untuk otomasi dan penilaian (akurasi, reliabilitas, dan penerimaan). 2.3 Akurasi tes otomatis: Validasi standar emas guru ( $k=19$ ): median  $k=0,63$  (IQR 0,56–0,72); Deteksi kesalahan/miskonsepsi ( $k=15$ ): median ketepatan 0,78; recall 0,71. Keandalan antar penilai (algoritmik versus manusia) ( $k=11$ ) adalah 0,74. Median ICC adalah 0,74. Penjemputan pengguna (guru dan siswa): rata-rata 4,1 (SD 0,5 RQ3: Implementasi AI, kontrol manusia, dan tata kelola data (privasi & keadilan) 2,5. AI dan kontrol manusia (human-in-the-loop): adopsi AI (n=78): 13 (17%) menjelaskan model antarmuka siswa dan guru; bentuk AI berdasarkan aturan/rasional (8), contoh/kounterfaktual (4), dan penekanan fitur (3). Kontrol (overriding/penyesuaian guru): terdapat pada 21 (27%) studi. 2.6 Privasi, Keadilan, dan Keandalan:

Pseudonimisasi adalah salah satu praktik privasi (35/78; 45%) dan minimisasi data (18/78; 23%). Audit keadilan menunjukkan bahwa 14 dari 78 orang (18%) melakukan uji perbedaan akurasi antar subkelompok. Kompetensi awal (k=12), bahasa rumah (k=7), gender (k=6), dan SES (k=5) adalah subkelompok umum. Hasil laporan mengindikasikan 6/14 (43%) melaporkan perbedaan akurasi di atas 5 poin.

**Tabel 1. Peran AI di Sekolah Dasar (2020–2025): Ringkasan Bukti dan Indikator Utama**

RQ	Fokus Umum	Ringkasan Temuan Kunci
RQ1	Dampak AI pada hasil belajar & praktik kelas SD	Mayoritas studi melaporkan peningkatan hasil belajar ( $\pm 72\%$ ) dan umpan balik lebih sering/akurat ( $\pm 82\%$ ); efek sedang (median $g \approx 0,32$ ) dan penghematan waktu guru sekitar 24 menit/minggu.
RQ2	Asesmen & otomasi (akurasi, reliabilitas, penerimaan)	Akurasi terhadap standar guru cukup baik ( $\kappa \approx 0,63$ ; $ICC \approx 0,74$ ); deteksi kesalahan $precision \approx 0,78$ / $recall \approx 0,71$ ; penerimaan guru tinggi ( $\approx 4,1/5$ ) dan antarmuka umumnya mudah digunakan.
RQ3	Transparansi (XAI), kontrol manusia, & tata kelola	Penerapan XAI masih terbatas ( $\sim 17\%$ ) dan fitur kontrol guru $\sim 27\%$ ; praktik privasi seperti pseudonimisasi umum ( $\sim 45\%$ ), namun audit keadilan masih jarang ( $\sim 18\%$ ) dan $\sim 43\%$ di antaranya menemukan selisih akurasi antarkelompok ( $\geq 5$ poin).

Hasil tambahan atau temuan tambahan (2020–2025 Asesmen & otomasi (akurasi, reliabilitas, penerimaan)): Proporsi artikel dengan tema GenAI/LLM meningkat pada  $\geq 2023$  (11/17; 65% dari seluruh studi GenAI dilakukan pada 2023–2025). Selain itu, ukuran sampel studi kuantitatif rata-rata lebih besar pada 2023–2025 daripada 2020–2022, dengan jumlah peserta rata-rata +37. Tingkat dan durasi intervensi: Intervensi yang berlangsung lebih dari

10 minggu memiliki tingkat kesetiaan yang lebih tinggi (13/19)

Kualitas pelaporan: Prapendaftaran (PROSPERO/OSF) tidak ada, ketersediaan data/kode 9/78 (12%), dan pelaporan jelas tentang bahaya bias alat 15/78 (19%).

Tinjauan literatur sistematis menunjukkan bahwa penggunaan AI di sekolah dasar terutama berkonsentrasi pada dua hal utama: hasil belajar siswa yang lebih baik dan praktik pembelajaran di kelas yang lebih baik—khususnya kualitas

umpan balik formatif dan efisiensi kerja guru. Secara keseluruhan, sebagian besar studi menunjukkan peningkatan dalam literasi dan numerasi, dengan besaran efek sedang (median  $g$  0,32). Studi membuktikan bahwa AI sangat efektif baik sebagai pengajar daripada sebagai pengganti.

Selain itu, sekitar 72% penelitian menunjukkan peningkatan setidaknya satu indikator capaian belajar. Efek literasi 0,30 dan efek numerasi 0,34. Sekitar 82% penelitian yang melibatkan sistem umpan balik otomatis menunjukkan peningkatan frekuensi dan ketepatan dalam praktik kelas. Sama sekali, penelitian menunjukkan bahwa waktu kerja guru dapat dikurangi sekitar 24 menit per minggu. Dua sinyal ini sangat penting dalam konteks kelas dasar yang cenderung heterogen. Siklus "latihan-mendapat umpan balik-memperbaiki" menjadi lebih efektif, sehingga guru memiliki lebih banyak waktu untuk meningkatkan interaksi instruksional yang sangat penting, seperti membuat strategi metakognitif atau mengadakan konferensi singkat dengan siswa.

AI biasanya "cukup memadai" untuk fungsi pendamping, berbeda dengan penentu tunggal. Kemampuan sistem untuk menemukan miskonsepsi juga relatif baik, dengan koefisien kesepakatan sedang hingga tinggi (median 0,63; ICC 0,74) dan ketepatan 0,78 dan recall 0,71. Meskipun ada pekerjaan yang harus dilakukan, guru cenderung positif (4,1/5) dan menganggap antarmuka sistem mudah digunakan. Dampaknya, mengembangkan kapasitas imperaktif. Pengendalian data, pelatihan fakta di kelas, dan interpretasi hasil AI dapat meningkatkan hasil di lapangan.

Dengan demikian, tiga masalah menahan integrasi yang bertanggung jawab: penjelasan sistem yang dapat dijelaskan (AI/XAI yang dapat dijelaskan), manajemen manusia (human-in-the-loop), dan manajemen data. Sekitar 17% penelitian melihat fitur AI di antarmuka guru dan siswa, dan sekitar 27% melihat mekanisme penghentian dan penyesuaian guru. Kekosongan ini sangat penting untuk anak usia sekolah dasar karena rekomendasi AI harus diterangkan dengan cara yang ramah anak, seperti

alasan singkat atau contoh-konterfaktual, sehingga dapat dipahami sebagai bagian dari "evaluasi sebagai pembelajaran" daripada keputusan akhir mesin. Praktik privasi seperti pseudonimisasi (45%) dan minimisasi data (23%) di bidang tata kelola tidak diimbangi dengan audit keadilan yang memadai. Baru sekitar 18% penelitian menilai perbedaan kinerja antarkelompok, dan hampir separuh audit menemukan kesalahan setidaknya 5 poin. Ini menunjukkan betapa pentingnya pengawasan yang menyeluruh atas keadilan, seperti berdasarkan status sosial-ekonomi, kemampuan awal, dan bahasa di rumah. Ini juga menunjukkan betapa pentingnya proses mitigasi, termasuk pengawasan guru, penetapan ambang kebijakan, dan penyetaraan kurikulum.

Bukti harus diperkuat terkait metode. Fidelitas implementasi yang lebih baik cenderung dikaitkan dengan jumlah sampel yang lebih besar dan durasi intervensi yang lebih lama. Namun, praktik sains terbuka masih sedikit: prapendaftaran hampir tidak ada, keterbukaan data/kode baru sekitar 12%, dan pelaporan

risiko bias sekitar 19%. Dibutuhkan dorongan untuk prapendaftaran, berbagi protokol dan materi, dan pelaporan analitik yang lengkap untuk meningkatkan ketelitian inferensi, terutama untuk publikasi di jurnal nasional terindeks SINTA. Selain itu, relevansi praktik dan kebijakan uji akan meningkat dengan desain yang lebih in-situ (berbasis kelas nyata) daripada simulasi laboratorium. Sekolah dasar dipengaruhi dalam tiga tingkat secara praktis. Pertama, desain pembelajaran: gunakan AI untuk membantu diferensiasi dan pembinaan regulasi diri belajar (SRL) dengan umpan balik yang ringkas, spesifik, dan siap dilaksanakan. Kedua, desain sistem: masukkan XAI yang dapat dimengerti anak dan panel kendali guru untuk melakukan hal-hal seperti menetapkan jadwal latihan, memperbaiki tingkat kesulitan, dan menguji efektivitasnya dengan melihat kelas. Ketiga, pengelolaan. Ini menunjukkan bahwa jumlah totalnya lebih besar. Ini berarti menghasilkan persetujuan yang besar antara orang tua dan anak, membuat kebijakan untuk mengurangi jumlah data, dan secara teratur melaporkan metrik

keadilan. Dengan dukungan kebijakan dan kapasitas yang tepat, AI memiliki potensi untuk meningkatkan orientasi pendidikan yang humanis dan berpusat pada peserta didik.

### **Simpulan**

Berdasarkan temuan pada penelitian mengindikasikan bahwa pemanfaatan AI di sekolah dasar mampu mendorong perbaikan nyata pada capaian belajar, memperkaya praktik umpan balik formatif, serta mengurangi pekerjaan rutin guru. Walaupun demikian, manfaat yang dijabarkan hanya berkelanjutan bila penerapannya dijalankan secara hati-hati dan bertanggung jawab. Ada tiga persyaratan utama yang harus dipenuhi. Pertama, sistem harus memberikan penjelasan yang jelas, sederhana, dan mudah dipahami oleh guru dan siswa. Ini akan memungkinkan rekomendasi AI dianggap sebagai bagian dari proses belajar dan bukan keputusan akhir mesin. Kedua, mekanisme human-in-the-loop harus benar-benar berfungsi di kelas. Ini berarti guru dapat mengubah, mengoreksi, atau menolak hasil AI, sehingga kontrol pedagogis

tetap berada di tangan manusia. Ketiga, tata kelola data harus kuat: kebijakan privasi yang jelas, prosedur validasi keandalan, serta pemantauan keadilan lintas kelompok siswa perlu berjalan konsisten agar tidak terjadi bias dan dampak yang tidak merata. Ke depan, riset dan implementasi sebaiknya menekankan uji lapangan berskala di konteks nyata, praktik sains terbuka (termasuk prapendaftaran dan pelaporan metodologis yang transparan), serta evaluasi keadilan yang berkelanjutan. Dengan fondasi tersebut, pemanfaatan AI memiliki kesempatan untuk memperkuat pendidikan yang inklusif, humanis, dan berpusat pada peserta didik. Hal tersebut juga sekaligus memastikan pemerataan manfaat kecerdasan artifisial di seluruh ekosistem sekolah dasar.

### **Daftar Pustaka**

- Akman, E. (2025). *The impact of AI-based visual designs on students' AI literacy and attitudes toward AI*. *Interactive Learning Environments*, 4820, 1–19.
- Balasoorya, C. (2024). *AI in education: A futuristic vision*. *Medical Teacher*, 46(7), 986–988.

- Chigbu, B. I., & Makapela, S. L. (2025). AI in education, sustainability, and the future of work: An integrative review of industry 5.0, education 5.0, and work 5.0. *Journal of Open Innovation: Technology, Market, and Complexity*, 11(4), 100645.
- Costa, C., & Murphy, M. (2025). Generative artificial intelligence in education: (what) are we thinking? *Learning, Media and Technology*, 9884, 1–12.
- Farrow, R. (2023). The possibilities and limits of XAI in education: a socio-technical perspective. *Learning, Media and Technology*, 48(2), 266–279.
- Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2025). The promise and challenges of generative AI in education. *Behaviour and Information Technology*, 44(11), 2518–2544.
- Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J., & Shimada, A. (2024). Artificial Intelligence in Education: Implications for Policymakers, Researchers, and Practitioners. *Technology, Knowledge and Learning*, 29(4), 1693–1710.
- Linderoth, C., Mani, M., Schönborn, K., Hultén, M., & Stenliden, L. (2025). Defining ‘the Force’ of artificial intelligence in education: exploring the future of teaching through informed speculation. *Learning, Media and Technology*, 9884, 1–15.
- Liu, Z. Y., & Yushchik, E. (2024). Exploring the prospects of using artificial intelligence in education. *Cogent Education*, 11(1).
- Muthukrishna, M., Dai, J., Panizo Madrid, D., Sabherwal, R., Vanoppen, K., & Yao, H. (2025). AI Can Revolutionise Education but Technology Is Not Enough: Human Development Meets Cultural Evolution. *Journal of Human Development and Capabilities*, 26(3), 482–492.
- Nemorin, S. (2024). Towards decolonising the ethics of AI in education. *Globalisation, Societies and Education*, 7724, 1–13.
- Parker, B. D. (2024). Considering the Impact of AI on the Professional Status of Teaching. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 97(6), 233–236.
- Ramirez-Montoya, M. S., Morales-Menendez, R., Tworek, M., Escobar, C. A., Tariq, R., & Tenorio-Sepulveda, G. C. (2024). Complex competencies for leader education: artificial intelligence analysis in student achievement profiling. *Cogent Education*, 11(1).
- Ronzon, T., Gurria, P., Carus, M., Cingiz, K., El-Meligi, A., Hark, N., Iost, S., M'barek, R., Philippidis, G., van Leeuwen, M., Wesseler, J.,

Medina-Lozano, I., Grimplet, J.,  
Díaz, A., Tejedor-Calvo, E., Marco,  
P., Fischer, M., Creydt, M.,  
Sánchez-Hernández, E., ... Miras  
Ávalos, J. M. (2025). No 主観的健康  
感を中心とした在宅高齢者に  
おける 健康関連指標に関する  
共分散構造分析Title.  
Sustainability (Switzerland),

11(1), 1–14.

Wang, S., Wang, F., Zhu, Z., Wang, J.,  
Tran, T., & Du, Z. (2024).  
Artificial intelligence in  
education: A systematic  
literature review. Expert  
Systems with Applications,  
252(PA), 124167.