



Integrasi kecerdasan buatan dalam evaluasi pendidikan menengah

Author Name(s): Sutrisno Sutrisno, Badeni Badeni, Muhammad Kristiawan, Eko Risdianto

Publication details, including author guidelines

URL: <https://jurnal.konselingindonesia.com/index.php/jkp/about/submissions#authorGuidelines>

Editor: Linda Fitria

Article History

Received: 7 Feb 2025

Revised: 20 Mar 2025

Accepted: 22 Apr 2025

How to cite this article (APA)

Sutrisno, S., Badeni, B., Kristiawan, M., & Risdianto, E. (2025). Integrasi kecerdasan buatan dalam evaluasi pendidikan menengah. Jurnal Konseling dan Pendidikan. 13(1), 503-511. <https://doi.org/10.29210/1139700>

The readers can link to article via <https://doi.org/10.29210/1139700>

SCROLL DOWN TO READ THIS ARTICLE



Indonesian Institute for Counseling, Education and Therapy (as publisher) makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications. However, we make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors and are not the views of or endorsed by Indonesian Institute for Counseling, Education and Therapy. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Indonesian Institute for Counseling, Education and Therapy shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to, or arising out of the use of the content.

Jurnal Konseling dan Pendidikan is published by Indonesian Institute for Counseling, Education and Therapy comply with the [Principles of Transparency and Best Practice in Scholarly Publishing](#) at all stages of the publication process. Jurnal Konseling dan Pendidikan also may contain links to web sites operated by other parties. These links are provided purely for educational purpose.



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

Copyright by Sutrisno, S., Badeni, B., Kristiawan, M., & Risdianto, E. (2025).

The author(s) whose names are listed in this manuscript declared that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript. This statement is signed by all the authors to indicate agreement that the all information in this article is true and correct.

Jurnal Konseling dan Pendidikan

ISSN 2337-6740 (Print) | ISSN 2337-6880 (Electronic)



Article

Integrasi kecerdasan buatan dalam evaluasi pendidikan menengah



Sutrisno Sutrisno^{*)}, Badeni Badeni, Muhammad Kristiawan, Eko Risdianto
Universitas Bengkulu, Indonesia

ABSTRAK

Artikel ini membahas integrasi kecerdasan buatan (AI) dalam evaluasi pembelajaran pada pendidikan menengah, sebagai respons terhadap keterbatasan metode evaluasi konvensional. Penelitian ini menggunakan metode tinjauan sistematis berbasis PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) yang mencakup empat tahap utama: identifikasi, penyaringan, uji kelayakan, dan inklusi. Proses pencarian dilakukan melalui basis data Scopus dan Semantic Scholar, menghasilkan 99 artikel. Setelah melalui tahap penyaringan dan eliminasi berdasarkan kriteria inklusi dan eksklusi, empat artikel yang relevan dianalisis lebih lanjut. Hasil kajian menunjukkan bahwa AI mampu meningkatkan objektivitas, efisiensi, dan personalisasi dalam evaluasi siswa melalui teknologi seperti machine learning dan pemrosesan bahasa alami. AI berperan penting dalam pengembangan rubrik penilaian yang presisi, pemetaan kesenjangan pembelajaran, serta pemberian umpan balik yang adaptif. Namun, tantangan seperti keterbatasan infrastruktur, isu etika data, dan resistensi guru masih menjadi hambatan. Studi ini menekankan bahwa AI sebaiknya digunakan sebagai alat bantu, bukan pengganti, dalam proses evaluasi. Temuan ini diharapkan menjadi dasar dalam perumusan kebijakan pendidikan menengah yang lebih adaptif dan inklusif terhadap perkembangan teknologi.

Keywords:

Kecerdasan buatan
Evaluasi pembelajaran
Sekolah menengah

Corresponding Author:

Sutrisno Sutrisno,
Universitas Bengkulu
Email: sutrisnobkl33@gmail.com

Pendahuluan

Evaluasi pendidikan merupakan aspek fundamental dalam proses pembelajaran karena berfungsi sebagai alat ukur keberhasilan siswa, efektivitas pengajaran, dan pencapaian kurikulum (Zhou, 2023). Seiring perkembangan paradigma pendidikan dari berorientasi pada hasil ke berpusat pada siswa, metode evaluasi pun mengalami evolusi. Pada awalnya, evaluasi hanya mengandalkan tes tulis konvensional yang bersifat sumatif. Model ini bertujuan untuk mengukur seberapa banyak pengetahuan yang dapat diingat atau dihafal siswa, tanpa mempertimbangkan keragaman potensi, gaya belajar, dan latar belakang peserta didik (Gayathri, 2024).

Seiring waktu, muncul kritik terhadap pendekatan konvensional karena dianggap tidak mencerminkan kompetensi siswa secara menyeluruh, khususnya dalam konteks pembelajaran abad ke-21 yang menekankan keterampilan berpikir kritis, kolaboratif, dan kreatif. Oleh karena itu, muncul evaluasi formatif, autentik, dan berbasis portofolio yang lebih menekankan proses belajar, namun implementasinya sering terkendala waktu, sumber daya, dan kemampuan pendidik dalam melakukan penilaian yang adil dan objektif (Abdurahim-Salain, 2024). Kemajuan teknologi informasi membawa perubahan bagi dunia pendidikan, termasuk dalam hal evaluasi. Dalam dua dekade terakhir, mulai digunakan sistem evaluasi berbasis komputer (CBT) yang meningkatkan efisiensi dan akurasi penilaian. Namun, pendekatan ini mulai dipertanyakan efektivitasnya karena cenderung

bersifat satu dimensi, tidak mampu menangkap potensi individual, dan rentan terhadap bias (Myshchyn, 2022).

Metode evaluasi konvensional menghadirkan sejumlah permasalahan signifikan. Pertama, penilaian berbasis ujian standar kadang mengabaikan keragaman gaya belajar siswa dan tidak memberikan ruang bagi siswa dengan kecerdasan majemuk (Dilekli & Tezci, 2022). Kemudian format soal pilihan ganda atau esai terbatas tidak dapat menilai keterampilan abad 21 seperti kolaborasi, berpikir kritis, dan kreativitas secara menyeluruh (Rusmin et al., 2024). Terakhir evaluasi ini juga memunculkan bias yang berdampak pada keadilan penilaian (Tarun & Krueger, 2016). Studi Bandyopadhyay dan Chugh (2020) mengungkapkan bahwa siswa dari latar belakang sosio-ekonomi rendah cenderung memiliki performa lebih rendah dalam tes standar karena faktor lingkungan belajar, bukan kemampuan akademik semata. Hal ini menunjukkan bahwa evaluasi konvensional dapat memperlebar ketimpangan pendidikan.

Dalam konteks inilah, kecerdasan buatan (*Artificial Intelligence/AI*) mulai dipandang sebagai solusi potensial untuk meningkatkan kualitas dan akurasi evaluasi pembelajaran (Liu, 2024). Secara teknis, AI mengatasi kelemahan dalam evaluasi pendidikan melalui algoritma adaptif dan pembelajaran mesin yang memungkinkan penilaian *real-time*, personalisasi umpan balik, dan analisis data siswa secara longitudinal (Tambat, 2024). Teknologi seperti natural *language processing* dapat menilai respons terbuka siswa secara otomatis, sementara *machine learning* dapat memprediksi kebutuhan belajar berdasarkan pola interaksi siswa sebelumnya (Ansari, 2024). Lewis et al. (2024) menemukan bahwa AI mampu merancang *asesmen* berbasis simulasi interaktif yang lebih kontekstual dan mengukur kompetensi holistik siswa. Studi Khlaif (2024) didapat bahwa AI lebih mengarah pada evaluasi yang lebih objektif karena mengurangi faktor bias manusia yang mungkin muncul dalam evaluasi tradisional.

Beberapa negara telah mulai menerapkan AI dalam pendidikan dengan hasil yang menjanjikan. Di Singapura, misalnya, pemerintah telah mengembangkan sistem pembelajaran berbasis AI yang memungkinkan analisis data siswa secara *real-time* dan memberikan umpan balik yang dapat membantu siswa meningkatkan hasil belajar mereka (OswalLine et al., 2024). Menurut Karmakar dan Das (2024), penerapan AI dalam pendidikan dapat meningkatkan keterlibatan siswa serta hasil akademik mereka secara signifikan. Penelitian oleh Mertala et al. (2022) menunjukkan bahwa penggunaan AI dalam sistem pendidikan Finlandia telah berhasil menghasilkan metode evaluasi yang lebih inklusif dan menyeluruh. Studi Ubah et al. (2022) menunjukkan bahwa penggunaan teknologi AI dalam pendidikan dapat membantu mengidentifikasi siswa yang memerlukan perhatian lebih, serta memberikan pelatihan dan materi pembelajaran yang lebih sesuai dengan kebutuhan individu mereka.

Penerapan AI di lingkungan sekolah menengah tidak lepas dari tantangan. Studi Aletras (2023) dibanyak sekolah menengah ditemukan bahwa AI tidak didukung oleh kurikulum. Kurikulum AI sering belum tersedia atau belum terintegrasi dengan baik dalam silabus sekolah menengah. Tilton et al. (2023) mengungkapkan bahwa sekolah di daerah tertinggal atau pedesaan biasanya kekurangan perangkat keras dan koneksi internet yang mendukung AI. Pengadaan perangkat dan platform AI memerlukan dana besar yang tidak semua sekolah mampu penuhi. Ubah et al. (2022) menemukan kurangnya literasi AI kalangan guru. Guru belum siap secara pengetahuan dan keterampilan, akibatnya sulit untuk memperkenalkan pembelajaran berbasis AI secara mendalam. Beberapa guru merasa AI bisa mengurangi peran mereka, dimana menurut Bond et al. (2020) masih ada skeptisisme tentang efektivitas AI dalam pembelajaran. Menurut Luckin dan Cukurova (2019), pembelajaran AI kadang hanya jadi bagian dari proyek atau ekstrakurikuler, bukan bagian inti pendidikan.

Penelitian ini bertujuan untuk menganalisis secara komprehensif bagaimana AI diintegrasikan dalam evaluasi pendidikan menengah, dengan menelaah manfaat, tantangan, serta efektivitasnya dalam meningkatkan kualitas asesmen. Secara khusus, penelitian ini akan mengevaluasi aspek-aspek seperti personalisasi penilaian, objektivitas, efisiensi, serta pengaruhnya terhadap hasil belajar siswa. Asumsi awal yang dibawa adalah bahwa penggunaan AI dapat memperbaiki kelemahan evaluasi tradisional dan menciptakan sistem penilaian yang lebih adil dan inklusif.

Dalam menjawab pertanyaan penelitian ini, pendekatan sistematis melalui metode PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) digunakan untuk menyeleksi dan meninjau artikel secara transparan, sistematis, dan replikatif. PRISMA dipilih karena mampu meminimalkan bias pemilihan literatur dan memberikan kerangka kerja yang kuat untuk membandingkan berbagai temuan penelitian. Berbeda dengan metode tinjauan naratif atau *scoping review*, PRISMA memberikan kejelasan dalam proses penyaringan dan kriteria inklusi-eksklusi, sehingga meningkatkan validitas hasil kajian.

Metode

Kajian ini bertujuan untuk memberikan tinjauan komprehensif mengenai peran dan integrasi kecerdasan buatan (AI) dalam evaluasi pendidikan pada tingkat pendidikan menengah. Penelitian ini menggunakan metode PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) untuk memastikan ketelitian dan transparansi dalam proses seleksi serta evaluasi artikel yang digunakan. Metode PRISMA telah banyak digunakan dalam penelitian literatur sistematis untuk memperoleh hasil yang objektif dan dapat diandalkan (Reis et al., 2023). Menurut Tassinari et al. (2023), PRISMA menyediakan daftar periksa dan diagram alir untuk meningkatkan pelaporan tinjauan sistematis dan meta-analisis, memastikan bahwa semua studi yang relevan.

Penelitian ini secara khusus menggunakan dua basis data utama, yaitu Scopus dan Semantic Scholar, dalam proses pengumpulan literatur. Pemilihan kedua basis data ini didasarkan pada beberapa pertimbangan. Scopus merupakan salah satu pangkalan data ilmiah terbesar dan paling bereputasi di dunia, yang menyediakan akses ke jurnal-jurnal *peer-reviewed* berkualitas tinggi di berbagai bidang ilmu. Sementara itu, Semantic Scholar dipilih karena memiliki keunggulan dalam penelusuran artikel berbasis kecerdasan buatan, yang mampu menampilkan artikel-artikel relevan dengan efisiensi tinggi, terutama dalam topik-topik yang bersifat multidisipliner seperti pendidikan dan teknologi. Penggunaan dua basis data ini dianggap sudah memadai untuk tujuan kajian sistematis karena keduanya saling melengkapi. Scopus menjamin kualitas dan kredibilitas sumber, sedangkan Semantic Scholar memperluas cakupan pencarian artikel yang relevan secara tematis dan terkini. Meskipun terdapat banyak database lain yang tersedia, penelitian ini membatasi pada dua sumber tersebut untuk menjaga konsistensi dan fokus dalam proses penyaringan serta analisis artikel.

Reis et al. (2023) mengatakan bahwa metode PRISMA mendorong pencarian database yang lengkap untuk meminimalkan bias dan memastikan bahwa semua studi yang relevan disertakan. Terkait dengan kajian ini, pencarian ini menggunakan kata kunci yang relevan dengan topik penelitian, yaitu *Artificial Intelligence*, *Educational Evaluation*, dan *Secondary Education*. Kata kunci ini dipilih dengan tujuan untuk menyaring artikel-artikel yang berkaitan dengan penggunaan AI dalam konteks evaluasi pendidikan pada tingkat pendidikan menengah. Pemilihan artikel dilakukan tanpa membatasi jenis pendekatan metodologisnya. Oleh karena itu, baik studi kualitatif maupun kuantitatif disertakan dalam seleksi, selama artikel tersebut memenuhi kriteria inklusi: (1) Artikel yang dipublikasikan antara tahun 2019 hingga 2024; (2) Artikel yang berbahasa Inggris untuk memastikan keberagaman literatur internasional; (3) Artikel yang diterbitkan dalam jurnal *peer-reviewed* untuk memastikan kualitas ilmiah dan validitas hasil penelitian; (4) Artikel yang dapat diakses lengkap

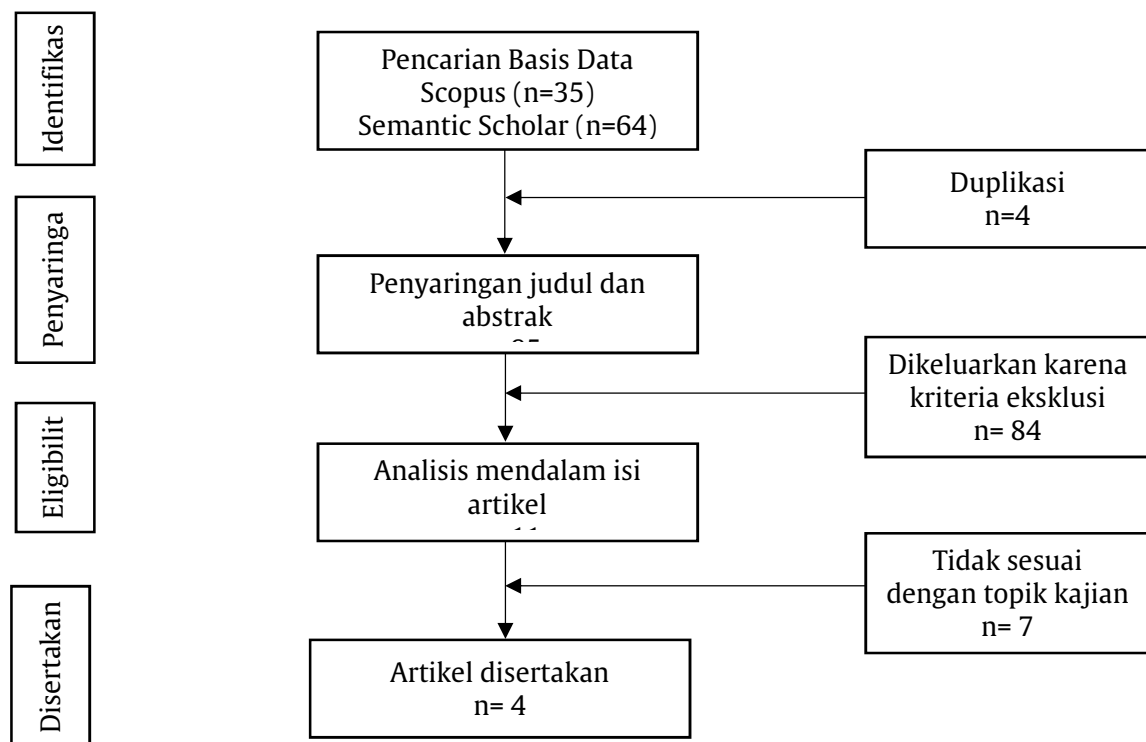
Selain itu, terdapat pula kriteria eksklusi yang mencakup: (1) Artikel publikasi di bawah tahun 2019; (2) Artikel yang tidak relevan dengan penggunaan kecerdasan buatan dalam evaluasi pendidikan di tingkat pendidikan menengah; (3) Artikel yang tidak memenuhi kriteria kualitas ilmiah atau tidak dapat diakses melalui platform yang digunakan; (4) Penelitian yang tidak diterbitkan dalam jurnal *peer-reviewed* atau tidak memiliki metodologi yang jelas dan dapat diuji; (5) Artikel tidak dapat diakses

Proses penyaringan literatur dalam studi ini mengikuti alur kerja yang sistematis dan berlapis sebagaimana yang diatur dalam panduan PRISMA. Tahap awal dimulai dengan identifikasi sumber

literatur melalui dua pangkalan data ilmiah utama, yakni Scopus dan Semantic Scholar. Dari proses pencarian dengan kata kunci yang telah ditentukan secara tematik, diperoleh total 99 artikel: 35 artikel dari Scopus dan 64 artikel dari Semantic Scholar. Langkah berikutnya adalah penyaringan duplikasi untuk menjamin bahwa setiap artikel yang dianalisis bersifat unik dan tidak redundan. Dalam tahap ini ditemukan 4 artikel duplikat, yang kemudian dihapus dari daftar untuk menghindari bias dalam analisis dan perhitungan temuan.

Setelah penyaringan duplikasi, dilakukan tahap seleksi awal dengan menelaah judul dan abstrak dari 95 artikel yang tersisa. Proses ini bertujuan untuk menilai kecocokan awal terhadap kriteria inklusi, yaitu artikel yang relevan secara tematik dengan fokus kajian (AI dalam evaluasi pendidikan menengah), dipublikasikan antara tahun 2019–2024, berbahasa Inggris, dan terbit dalam jurnal ilmiah peer-reviewed. Pada tahap ini, 84 artikel dieliminasi karena tidak memenuhi kriteria tersebut, baik karena konteks yang tidak sesuai, tidak fokus pada pendidikan menengah, atau metodologi yang kurang dapat dipertanggungjawabkan secara akademik.

Selanjutnya, dilakukan evaluasi kelayakan secara menyeluruh terhadap 11 artikel yang lolos seleksi awal. Evaluasi ini mencakup pemeriksaan terhadap isi penuh artikel, fokus topik, kualitas penyajian data, validitas metodologis, dan kontribusi teoritis terhadap permasalahan penelitian. Dalam tahap ini, 7 artikel kembali dieliminasi karena tidak secara langsung membahas integrasi kecerdasan buatan dalam evaluasi pendidikan menengah atau tidak memenuhi standar kualitas ilmiah yang ditetapkan. Akhirnya, hanya 4 artikel yang dinyatakan memenuhi seluruh kriteria seleksi dan dianggap layak untuk dimasukkan dalam tinjauan akhir. Artikel-artikel ini mewakili kontribusi empiris dan teoretis yang relevan, serta memberikan beragam perspektif mengenai penerapan AI dalam sistem evaluasi pendidikan menengah, baik dari segi manfaat, tantangan, maupun potensi pengembangannya di masa depan.



Gambar 1 Gambar Metode PRISMA

Setelah melalui proses seleksi berdasarkan metode PRISMA, artikel-artikel yang lolos tahap akhir kemudian dianalisis secara tematik. Analisis tematik dilakukan untuk mengidentifikasi pola-pola utama, kategori, serta tema-tema yang relevan terkait penggunaan kecerdasan buatan dalam evaluasi

pendidikan menengah. Proses ini bertujuan untuk menggali secara mendalam berbagai pendekatan, manfaat, tantangan, serta dampak dari penerapan AI dalam konteks evaluasi pembelajaran. Melalui pendekatan tematik, setiap artikel ditelaah untuk menemukan aspek-aspek kunci seperti: (1) bentuk penerapan teknologi AI dalam evaluasi, (2) jenis data atau indikator yang digunakan dalam penilaian berbasis AI, (3) dampak AI terhadap efektivitas dan objektivitas evaluasi, serta (4) tantangan teknis, etis, maupun kebijakan yang muncul dalam implementasinya di sekolah menengah. Tema-tema tersebut kemudian dibandingkan untuk mendapatkan sintesis pemahaman yang komprehensif.

Meskipun metode PRISMA merupakan pendekatan yang kuat dan banyak digunakan untuk meninjau literatur secara sistematis dan transparan, terdapat beberapa keterbatasan signifikan ketika diterapkan dalam konteks evaluasi pendidikan, khususnya pada topik yang bersifat multidisipliner seperti penerapan kecerdasan buatan (AI). Pertama, PRISMA sangat bergantung pada kualitas dan keterjangkauan publikasi ilmiah yang tersedia di basis data tertentu. Dalam konteks evaluasi pendidikan, banyak praktik inovatif atau studi penerapan teknologi yang dilaporkan melalui laporan institusional, studi kasus di konferensi, atau artikel non-peer-reviewed yang tidak terindeks dalam Scopus atau Semantic Scholar. Akibatnya, penerapan PRISMA dapat mengecualikan sumber informasi yang sebenarnya relevan dan kaya akan konteks praktis, terutama dalam studi pendidikan di negara berkembang. Kedua, PRISMA dirancang untuk studi yang memiliki variabel dan temuan yang relatif terstandarisasi, seperti dalam bidang kesehatan. Sementara dalam bidang pendidikan, terutama evaluasi pembelajaran, pendekatannya sangat beragam, baik dari segi paradigma (kuantitatif, kualitatif, campuran), desain, maupun indikator keberhasilan. Hal ini dapat menyulitkan proses sintesis temuan, karena studi-studi tersebut tidak selalu dapat dibandingkan secara langsung dalam kerangka meta-analisis yang ketat. Ketiga, PRISMA tidak secara eksplisit menangani analisis konteks sosial, budaya, dan kebijakan pendidikan, padahal faktor-faktor tersebut sangat berpengaruh dalam penerapan AI di lingkungan sekolah. Tanpa integrasi pendekatan kontekstual atau framework interpretatif, penggunaan PRISMA berisiko menghasilkan pemahaman yang terlalu teknis dan terlepas dari realitas pendidikan di lapangan. Keempat, meskipun PRISMA meningkatkan transparansi dalam proses seleksi artikel, ia tidak memberikan panduan spesifik untuk analisis tematik atau naratif yang lebih cocok digunakan dalam kajian pendidikan berbasis eksploratif. Dengan demikian, peneliti perlu mengombinasikan PRISMA dengan metode analisis kualitatif seperti analisis tematik untuk menangkap dimensi konseptual yang tidak dapat diungkap hanya melalui kuantifikasi data literatur.

Hasil dan Pembahasan

Melalui proses PRISMA, ditemukan empat artikel yang benar-benar relevan dan memenuhi standar yang dipilih untuk dianalisis lebih lanjut. Berikut pada Tabel 1 adalah empat artikel yang disertakan.

Tabel 1. Artikel yang disertakan

Penulis (tahun)	Judul	Metode	Temuan
(Fernández-Sánchez et al., 2025)	Navigating the Future of Pedagogy: The Integration of AI Tools in Developing Educational Assessment Rubrics	Pendekatan metodologis aksi-penelitian yang digunakan dalam penelitian ini. Alat AI seperti ChatGPT digunakan untuk pengembangan rubrik.	AI meningkatkan presisi dan efisiensi rubrik. Mendukung pengalaman belajar yang dipersonalisasi dan strategi penilaian inovatif.
(Pusporini & Nurdianto, 2024)	Utilization of artificial intelligence in Outcome-Based Curriculum	Teknologi AI: pembelajaran mesin, pemrosesan bahasa alami. Menganalisis	AI mengidentifikasi kesenjangan pembelajaran, memprediksi hasil, merekomendasikan jalur yang dipersonalisasi. Alat

Penulis (tahun)	Judul	Metode	Temuan
	Evaluation and Development	data pendidikan untuk wawasan tentang kurikulum dan kinerja siswa	berbasis AI merampingkan evaluasi, memfasilitasi peningkatan kurikulum berkelanjutan
(Duraes et al., 2024)	AI-Driven Educational Transformation in Secondary Schools: Leveraging Data Insights for Inclusive Learning Environments	Teknik pembelajaran mesin dan teknik fusi data	AI meningkatkan pembelajaran yang dipersonalisasi, mengatasi perbedaan, menumbuhkan lingkungan yang inklusif. Kemudian teknik fusi data meningkatkan kinerja dengan menggabungkan prediksi model dasar.
(Ojeda & Castro Arias, 2025)	Changes in Academic Assessment Due to the Use of Artificial Intelligence	Survei terhadap siswa dan guru untuk mengumpulkan informasi tentang pengaruh kecerdasan buatan, penggunaan etisnya, dan penerapannya dalam berbagai kegiatan akademis	kecerdasan buatan digunakan dalam berbagai kegiatan pendidikan, memungkinkan optimalisasi pekerjaan guru, dan mengevaluasi, meningkatkan, dan mempersonalisasi proses belajar mengajar siswa. Khususnya, siswa yang menggunakan perangkat kecerdasan buatan lebih suka dinilai oleh guru manusia

Rekomendasi

Penerapan kecerdasan buatan (AI) dalam evaluasi pendidikan menengah menunjukkan hasil yang sangat menjanjikan. Namun, untuk memastikan bahwa inovasi ini dapat diimplementasikan secara efektif dan berkelanjutan, diperlukan langkah-langkah strategis yang bersifat operasional dan kontekstual. Implementasi AI di lingkungan sekolah menengah perlu dilakukan secara bertahap melalui skema pilot project atau percontohan pada mata pelajaran tertentu, khususnya yang memiliki karakter evaluasi formatif yang kuat. Pendekatan bertahap ini penting untuk meminimalkan resistensi dan memberikan ruang bagi sekolah dan guru untuk beradaptasi. Penelitian oleh Simon dan Zeng (2024) menekankan pentingnya proses adaptif dalam adopsi teknologi pendidikan, di mana keterlibatan guru sebagai pengguna utama harus diprioritaskan sejak tahap awal pengembangan.

Sekolah dapat mulai mengintegrasikan AI ke dalam sistem pembelajaran yang sudah tersedia, seperti *Learning Management System* (LMS) atau aplikasi berbasis digital yang mendukung *assessment for learning* (Alotaibi, 2024). Studi Pusporini & Nurdyanto (2024) menunjukkan bahwa pemanfaatan pembelajaran mesin dan pemrosesan bahasa alami dalam LMS memungkinkan sistem memberikan evaluasi personal dan prediksi performa siswa yang lebih akurat. Hal ini dapat memperkuat upaya perbaikan kurikulum dan pembelajaran berbasis data. Dalam proses implementasi, pelatihan guru menjadi faktor krusial. Guru perlu tidak hanya memahami cara kerja teknologi, tetapi juga bagaimana memanfaatkan hasil evaluasi AI untuk memperkaya proses pembelajaran. Sa et al. (2024) mencatat bahwa persepsi guru terhadap AI sangat menentukan keberhasilan integrasi teknologi ini. Oleh karena itu, pelatihan guru sebaiknya tidak hanya bersifat teknis, tetapi juga pedagogis menyediakan kerangka berpikir baru dalam menafsirkan data hasil penilaian yang dihasilkan oleh AI.

Rekomendasi selanjutnya adalah pengembangan sistem penilaian berbasis AI yang lebih adaptif terhadap keragaman siswa. Alat evaluasi yang dikembangkan perlu mempertimbangkan dinamika kelas yang beragam, termasuk perbedaan kemampuan, latar belakang, dan gaya belajar. Vetrivel et al. (2024) menunjukkan bahwa AI mampu menyusun rubrik penilaian otomatis yang disesuaikan

dengan konteks belajar siswa secara individual. Meski demikian, González-Calatayud et al. (2021) menekankan bahwa keterlibatan guru tetap penting agar aspek emosional, kontekstual, dan sosial dari pembelajaran tidak diabaikan. AI harus berperan sebagai asisten evaluatif, bukan sebagai satu-satunya penilai.

Selain itu, masih dibutuhkan penelitian jangka panjang untuk menilai efektivitas AI terhadap hasil belajar siswa. Meskipun beberapa studi, seperti Duraes et al. (2024), menunjukkan bahwa AI berkontribusi terhadap peningkatan performa akademik dan pembelajaran yang inklusif, pertanyaan tetap terbuka: apakah peningkatan tersebut berkelanjutan dan relevan dengan pengembangan keterampilan abad 21, seperti berpikir kritis dan reflektif? Penelitian masa depan perlu mengevaluasi dampak jangka panjang ini secara empiris dan lintas konteks. Kemudian dari sisi kebijakan, dukungan institusional dan infrastruktur sangat penting. Pemerintah dan lembaga pendidikan harus berperan aktif dalam menyediakan sarana digital, pelatihan, serta regulasi yang mendukung integrasi AI di sekolah. Ismail (2024) menyatakan bahwa ekosistem teknologi pendidikan hanya akan berkembang apabila didukung oleh kebijakan publik yang memadai dan partisipasi aktif dari semua pemangku kepentingan.

Dengan demikian, agar AI benar-benar menjadi instrumen yang memperkuat sistem evaluasi pendidikan menengah, pendekatan yang digunakan harus bersifat kolaboratif, kontekstual, dan berorientasi pada pembelajaran holistik. AI harus dipandang bukan sebagai pengganti guru, melainkan sebagai alat bantu strategis untuk memperluas kapasitas pendidik dalam memahami dan mendampingi perkembangan peserta didik secara lebih akurat dan manusiawi.

Simpulan

Integrasi kecerdasan buatan (AI) dalam evaluasi pendidikan menengah menawarkan potensi dalam meningkatkan kualitas asesmen yang selama ini didominasi oleh metode konvensional yang kurang adaptif, tidak personal, dan rentan bias. Melalui analisis sistematis terhadap studi terpilih, ditemukan bahwa AI mampu menghadirkan sistem penilaian yang lebih personal, objektif, efisien, dan inklusif, serta mampu beradaptasi dengan kebutuhan belajar siswa secara individual. AI memungkinkan pengembangan rubrik penilaian yang lebih presisi, memberikan umpan balik real-time, memprediksi performa siswa, dan menyesuaikan kurikulum berdasarkan data belajar yang aktual. Temuan ini memperkuat asumsi bahwa AI dapat memperbaiki kelemahan mendasar dari sistem evaluasi tradisional, dan mendukung transformasi evaluasi menjadi proses yang lebih menyeluruh, berorientasi proses, dan adil. Namun demikian, keberhasilan implementasi AI sangat dipengaruhi oleh kesiapan infrastruktur, kompetensi pendidik, serta penerimaan sosial terhadap teknologi ini. Tantangan seperti keterbatasan akses teknologi, isu etika data, dan resistensi terhadap otomatisasi harus diantisipasi agar integrasi AI tidak justru menciptakan kesenjangan baru dalam pendidikan. Dengan demikian, AI seharusnya diposisikan sebagai mitra strategis dalam pendidikan, bukan pengganti peran guru, guna menciptakan ekosistem evaluasi yang lebih bermakna, kontekstual, dan transformatif bagi seluruh peserta didik.

Referensi

- Abdurahim-Salain, H. (2024). Analysis of portfolio for critical thinking and creativity development among education students. *Environment and Social Psychology*, 9(11), 1–10. <https://doi.org/10.59429/esp.v9i11.3199>
- Aletras, K. (2023). Impacts on Education and Teaching from the Advance of Artificial Intelligence in Secondary Schools. *European Journal of Engineering and Technology Research*, 12(4), 16–22. <https://doi.org/10.24018/ejeng.2023.1.CIE.3133>
- Alotaibi, N. S. (2024). The Impact of AI and LMS Integration on the Future of Higher Education: Opportunities, Challenges, and Strategies for Transformation. *Sustainability*, 16(23), 101–115. <https://doi.org/10.3390/su162310357>

- Ansari, M. T. A. (2024). Role of Artificial Intelligence to Enhance Learning Competencies at Secondary School Students. *International Journal of Information Technology and Management*, 17(1), 102–106. <https://doi.org/10.29070/vh029415>
- Bandyopadhyay, M., & Chugh, S. (2020). Status of Secondary Education in India: A Review of Status, Challenges and Policy Issues. In *Universal Secondary Education in India* (pp. 17–49). Springer Singapore. https://doi.org/10.1007/978-981-15-5366-0_2
- Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: a systematic evidence map. *International Journal of Educational Technology in Higher Education*, 17(1), 2. <https://doi.org/10.1186/s41239-019-0176-8>
- Dilekli, Y., & Tezci, E. (2022). *Measuring, Assessing and Evaluating Thinking Skills in Educational Settings: A Necessity for Twenty-First Century* (pp. 395–415). https://doi.org/10.1007/978-3-031-15963-3_22
- Duraes, D., Bezerra, R., & Novais, P. (2024). AI-Driven Educational Transformation in Secondary Schools: Leveraging Data Insights for Inclusive Learning Environments. *2024 IEEE Global Engineering Education Conference (EDUCON)*, 1–9. <https://doi.org/10.1109/EDUCON60312.2024.10578910>
- Fernández Sánchez, A., Lorenzo Castiñeiras, J. J., & Sánchez Bello, A. (2025). Navigating the Future of Pedagogy: The Integration of AI Tools in Developing Educational Assessment Rubrics. *European Journal of Education*, 60(1), 1–10. <https://doi.org/10.1111/ejed.12826>
- Gayathri, J. (2024). Paradigm Shift in Teaching. *Indian Journal of Computer Science*, 9(1), 41–49. <https://doi.org/10.17010/ijcs/2024/v9/i1/173696>
- González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial Intelligence for Student Assessment: A Systematic Review. *Applied Sciences*, 11(12), 5467–5475. <https://doi.org/10.3390/app11125467>
- Ismail, O. S. (2024). *Empowering the Modern Learning Ecosystem* (pp. 145–167). <https://doi.org/10.4018/979-8-3693-4103-2.ch008>
- Karmakar, S., & Das, T. (2024). Effect of artificial intelligence on education. In *Optimization and Computing using Intelligent Data-Driven Approaches for Decision-Making* (pp. 198–211). CRC Press. <https://doi.org/10.1201/9781003536796-8>
- Khlaif, Z. (2024). *Rethinking Educational Assessment in the Age of Artificial Intelligence* (pp. 131–144). <https://doi.org/10.4018/979-8-3693-7255-5.ch005>
- Lewis, H. J., Johnson, S. M., & Phillips, E. (2024). *Redefining Student Achievement* (pp. 17–36). <https://doi.org/10.4018/979-8-3693-1001-4.ch002>
- Liu, H. H. B. (2024). Learning Evaluation Method Based on Artificial Intelligence Technology and Its Application in Education. *Journal of Electrical Systems*, 20(3s), 1833–1842. <https://doi.org/10.52783/jes.1722>
- Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences driven approach. *British Journal of Educational Technology*, 50(6), 2824–2838. <https://doi.org/10.1111/bjet.12861>
- Mertala, P., Fagerlund, J., & Calderon, O. (2022). Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence (AI) and their implications for AI literacy education. *Computers and Education: Artificial Intelligence*, 3(2), 101–117. <https://doi.org/10.1016/j.caeai.2022.100095>
- Myshchysyn, I. (2022). European experience of evaluating the quality of secondary education. *Visnyk of the Lviv University. Series Pedagogics*, 113(37), 133–142. <https://doi.org/10.30970/vpe.2022.37.11651>
- Ojeda, I., & Castro Arias, S. (2025). Changes in Academic Assessment Due to the Use of Artificial Intelligence. *Communications in Computer and Information Science*, 2273 CCIS, 364–380. https://doi.org/10.1007/978-3-031-75431-9_24

- OswalLine, N., Ateeq, K., & Mathew, S. (2024). Enhancing Educational Feedback Systems using Deep Queue-Dependent Networking: An Automatic StudentAssessment with AI. *2024 10th International Conference on Optimization and Applications (ICOA)*, 1–6. <https://doi.org/10.1109/ICOA62581.2024.10753808>
- Pusporini, W., & Nurdianto, H. (2024). Utilization of artificial intelligence in Outcome-Based Curriculum Evaluation and Development. *Journal of Research in Social Science and Humanities*, 4(2), 132–134. <https://doi.org/10.47679/jrssh.v4i1.109>
- Reis, I. W., Galvão, I., & Ulbricht, V. R. (2023). *Metaverse, Education, and Inclusion: Integrative Review Using the PRISMA Method* (pp. 587–598). https://doi.org/10.1007/978-981-99-5414-8_54
- Rusmin, L., Misrahayu, Y., Pongpalilu, F., Radiansyah, R., & Dwiyanto, D. (2024). Critical Thinking and Problem-Solving Skills in the 21st Century. *Join: Journal of Social Science*, 1(5), 144–162. <https://doi.org/10.59613/svhy3576>
- Sa n, F. G., Özkaya, A. B., Tengiz, F., Geyik, Ö. G., & Geyik, C. (2024). Current evaluation and recommendations for the use of artificial intelligence tools in education. *Turkish Journal of Biochemistry*, 48(6), 620–625. <https://doi.org/10.1515/tjb-2023-0254>
- Simon, P. D., & Zeng, L. M. (2024). Behind the Scenes of Adaptive Learning: A Scoping Review of Teachers' Perspectives on the Use of Adaptive Learning Technologies. *Education Sciences*, 14(12), 1413–1420. <https://doi.org/10.3390/educsci14121413>
- Tambat, P. (2024). Harnessing Artificial Intelligence to Transform Secondary Education in India: Innovations, Challenges, and Future Directions. *International Journal For Multidisciplinary Research*, 6(6), 1–10. <https://doi.org/10.36948/ijfmr.2024.v06i06.30145>
- Tarun, P., & Krueger, D. (2016). A Perspective on Student Evaluations, Teaching Techniques, and Critical Thinking. *Journal of Learning in Higher Education*, 12(1), 1–13.
- Tassinari, L. F. de M., Araújo, T. C. M. de, & Barbosa, J. A. (2023). Seismic Oceanography Method: a Prisma Based Sistematic Review. *Journal of Engineering Research*, 3(25), 2–26. <https://doi.org/10.22533/at.ed.3173252324071>
- Tilton, Z., LaVelle, J. M., Ford, T., & Montenegro, M. (2023). Artificial intelligence and the future of evaluation education: Possibilities and prototypes. *New Directions for Evaluation*, 2023(178–179), 97–109. <https://doi.org/10.1002/ev.20564>
- Ubah, A. E., Onakpojeruo, E. P., Ajamu, J., Mangai, T. R., Isa, A. M., Ayansina, N. B., & Al-Turjman, F. (2022). A Review of Artificial Intelligence in Education. *2022 International Conference on Artificial Intelligence of Things and Crowdsensing (AloTCs)*, 38–45. <https://doi.org/10.1109/AloTCs58181.2022.00104>
- Vetrivel, Vidhyapriya, P., & Arun. (2024). *The Role of AI in Transforming Assessment Practices in Education* (pp. 43–70). <https://doi.org/10.4018/979-8-3693-5443-8.ch003>
- Zhou, M. (2023). Significance of Assessment in Learning: The Role of Educational Assessment Tools. *Science Insights Education Frontiers*, 18(2), 2881–2883. <https://doi.org/10.15354/sief.23.co215>