

TRANSFORMASI EVALUASI PEMBELAJARAN BERBASIS ARTIFICIAL INTELLIGENCE PADA SEKOLAH MENENGAH KEJURUAN

Rulli Widianoro ^{a*)}, Hindarto ^{a)}, Jasno ^{a)}, Sumardi ^{a)}

^{a)}Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

^{*)}e-mail korespondensi: q100250013@student.ums.ac.id

Article history: received 06 Mei 2026; revised 20 Mei 2026; accepted 26 June 2026

DOI : <https://doi.org/10.33751/jmp.v14i02.120>

Abstrak. Transformasi digital di sektor pendidikan telah mendorong adopsi kecerdasan buatan (Artificial Intelligence/AI) sebagai instrumen evaluasi pembelajaran yang semakin relevan. Penelitian ini bertujuan untuk menganalisis secara komprehensif transformasi evaluasi pembelajaran berbasis AI pada Sekolah Menengah Kejuruan (SMK) di Indonesia, mencakup dimensi implementasi, tantangan struktural, dan dampaknya terhadap kualitas penilaian kompetensi siswa. Metode yang digunakan adalah Systematic Literature Review (SLR) dengan protokol PRISMA, menganalisis 45 artikel ilmiah yang diterbitkan pada periode 2022–2026 dari database Scopus, Web of Science, dan Google Scholar. Proses seleksi dilakukan melalui tiga tahap penyaringan yang menghasilkan artikel berkualitas tinggi dengan skor QUALIS minimal B2. Hasil penelitian menunjukkan bahwa implementasi AI dalam evaluasi pembelajaran di SMK memberikan efisiensi penilaian hingga 68%, meningkatkan akurasi identifikasi kompetensi siswa sebesar 54%, memfasilitasi umpan balik yang lebih personal dan real-time, serta mengurangi bias penilaian subjektif sebesar 43%. Analisis tematik mengungkap enam pola utama implementasi, yaitu: auto-grading (78%), analisis pola belajar (65%), rekomendasi materi adaptif (61%), penilaian portofolio digital (57%), deteksi kecurangan akademik (52%), dan prediksi hasil belajar (43%). Tantangan utama meliputi keterbatasan infrastruktur teknologi (72%), kurangnya kompetensi digital guru (65%), isu privasi data siswa (48%), resistensi perubahan budaya sekolah (39%), dan ketiadaan regulasi teknis yang spesifik (34%). Penelitian ini merekomendasikan pengembangan ekosistem AI yang adaptif dan kontekstual, disertai penguatan literasi digital tenaga pendidik dan pembentukan kebijakan tata kelola data siswa yang komprehensif.

Kata Kunci: Artificial Intelligence; Evaluasi Pembelajaran; Sekolah Menengah Kejuruan; Transformasi Digital; Penilaian Kompetensi; Systematic Literature Review

TRANSFORMATION OF ARTIFICIAL INTELLIGENCE-BASED LEARNING EVALUATION IN VOCATIONAL HIGH SCHOOLS

Abstract. The digital transformation in the education sector has driven the adoption of Artificial Intelligence (AI) as an increasingly relevant learning evaluation instrument. This study comprehensively analyzes the transformation of AI-based learning evaluation in Vocational High Schools (SMK) in Indonesia, covering implementation dimensions, structural challenges, and its impact on the quality of student competency assessment. The method used is a Systematic Literature Review (SLR) with PRISMA protocol, analyzing 45 scientific articles published in the period 2022–2026 from the Scopus, Web of Science, and Google Scholar databases. The selection process was carried out through three filtering stages producing high-quality articles with a minimum QUALIS score of B2. The results show that AI implementation in learning evaluation at SMK provides assessment efficiency of up to 68%, improves student competency identification accuracy by 54%, facilitates more personalized and real-time feedback, and reduces subjective assessment bias by 43%. Thematic analysis reveals six main implementation patterns: auto-grading (78%), learning pattern analysis (65%), adaptive material recommendations (61%), digital portfolio assessment (57%), academic dishonesty detection (52%), and learning outcome prediction (43%). Key challenges include limited technology infrastructure (72%), lack of teacher digital competency (65%), student data privacy issues (48%), school cultural resistance to change (39%), and absence of specific technical regulations (34%). This study recommends developing an adaptive and contextual AI ecosystem, accompanied by strengthening digital literacy for educators and establishing comprehensive student data governance policies.

Keywords: Artificial Intelligence; Learning Evaluation; Vocational High School; Digital Transformation; Competency Assessment; Systematic Literature Review

I. PENDAHULUAN

Perkembangan pesat teknologi kecerdasan buatan (Artificial Intelligence/AI) telah membawa perubahan fundamental di hampir seluruh sektor kehidupan, termasuk dunia pendidikan. Pada era revolusi industri 4.0 dan menjelang society 5.0, sistem pendidikan Indonesia dihadapkan pada tuntutan untuk beradaptasi dengan teknologi mutakhir yang mampu mendorong peningkatan kualitas pembelajaran secara menyeluruh dan berkelanjutan. Di antara berbagai jenjang pendidikan, Sekolah Menengah Kejuruan (SMK) memiliki posisi strategis yang unik karena berorientasi pada penguasaan kompetensi vokasional yang terukur, relevan, dan siap diuji di dunia industri.

Evaluasi pembelajaran merupakan salah satu komponen terpenting dalam siklus pendidikan. Melalui evaluasi yang valid, reliabel, dan komprehensif, para pemangku kepentingan pendidikan dapat memperoleh gambaran yang akurat mengenai pencapaian kompetensi siswa, efektivitas metode pengajaran, dan arah kebijakan pendidikan yang perlu diambil. Namun, sistem evaluasi di SMK selama ini masih menghadapi berbagai keterbatasan yang bersifat struktural. Penilaian berbasis kertas dan pensil yang mendominasi, ujian praktik yang sarat unsur subjektivitas, serta prosedur penilaian portofolio yang memakan waktu lama menjadi hambatan utama dalam mewujudkan sistem evaluasi yang efisien dan berkeadilan.

Kehadiran teknologi AI membuka cakrawala baru dalam dunia evaluasi pendidikan. Melalui kemampuan komputasi yang masif, AI memungkinkan proses penilaian otomatis (auto-grading), analisis mendalam terhadap pola belajar siswa, pemberian umpan balik yang personal dan real-time, serta prediksi kinerja akademis berbasis data historis. Platform-platform AI pendidikan seperti Gradescope, Turnitin AI, Khan Academy's AI tutor, dan berbagai sistem learning analytics yang terintegrasi dengan LMS telah membuktikan efektivitasnya di berbagai negara. Di Indonesia, adopsi teknologi ini mulai diwujudkan melalui kebijakan Merdeka Belajar yang mendorong inovasi dalam proses pembelajaran dan penilaian.

Meskipun potensinya sangat besar, transformasi evaluasi pembelajaran berbasis AI di SMK bukan tanpa tantangan. Keberagaman kondisi geografis, ketimpangan infrastruktur teknologi antardaerah, keterbatasan kompetensi digital guru, serta isu-isu etis seputar privasi data dan bias algoritma menjadi variabel-variabel kompleks yang harus dipertimbangkan secara serius. Selain itu, karakteristik khas SMK — seperti keberagaman program keahlian, struktur penilaian Uji Kompetensi Keahlian (UKK), dan keterlibatan industri dalam proses sertifikasi — menambah dimensi kompleksitas tersendiri yang berbeda dari sekolah menengah umum.

Penelitian-penelitian terdahulu tentang AI dalam evaluasi pendidikan lebih banyak berfokus pada konteks pendidikan tinggi atau sekolah menengah umum di negara-negara maju. Kajian yang secara spesifik mengeksplorasi transformasi evaluasi berbasis AI pada konteks SMK di Indonesia masih sangat terbatas, sehingga menciptakan kesenjangan pengetahuan (knowledge gap) yang signifikan. Padahal, Indonesia memiliki lebih dari 14.000 SMK dengan lebih dari 5 juta siswa yang tersebar di seluruh wilayah nusantara — sebuah skala yang menjadikan transformasi evaluasi berbasis AI berpotensi berdampak luar biasa jika diimplementasikan dengan tepat.

Berangkat dari latar belakang tersebut, penelitian ini bertujuan untuk: (1) mengidentifikasi pola-pola utama implementasi AI dalam evaluasi pembelajaran di SMK; (2) menganalisis tantangan-tantangan struktural yang menghambat adopsi AI untuk evaluasi di SMK; dan (3) mengevaluasi dampak implementasi AI terhadap kualitas penilaian kompetensi siswa SMK di Indonesia. Dengan menggunakan pendekatan Systematic Literature Review (SLR) yang ketat, penelitian ini diharapkan mampu memberikan kontribusi teoritis dan praktis bagi pengembangan kebijakan evaluasi pendidikan vokasional berbasis AI di Indonesia.

Secara teoretis, penelitian ini berkontribusi pada pengembangan literatur AIED dengan memperkenalkan perspektif pendidikan vokasional yang selama ini kurang terwakili dalam arus utama penelitian AI dan pendidikan. Kontribusi praktis dari penelitian ini mencakup penyediaan peta jalan (roadmap) implementasi AI evaluasi yang dapat diadaptasi oleh sekolah, dinas pendidikan, dan pembuat kebijakan di tingkat nasional. Di sisi lain, temuan penelitian ini diharapkan dapat menjadi referensi bagi pengembang platform AI pendidikan dalam merancang produk yang lebih responsif terhadap kebutuhan kontekstual SMK Indonesia, alih-alih sekadar mengadaptasi platform yang dikembangkan untuk konteks pendidikan umum di negara maju.

II. METODE PENELITIAN

1. Desain Penelitian

Penelitian ini menggunakan pendekatan Systematic Literature Review (SLR) yang mengacu pada protokol PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Pemilihan metode SLR didasarkan pada keunggulannya dalam mensintesis bukti-bukti ilmiah secara komprehensif, transparan, dan dapat direproduksi (reproducible), sehingga meminimalkan bias seleksi dan konfirmasi yang sering mengganggu tinjauan literatur naratif konvensional. SLR dipandang sebagai metode yang paling tepat untuk menjawab pertanyaan penelitian yang bersifat eksploratif-komprehensif tentang transformasi evaluasi berbasis AI di SMK.

Pertanyaan penelitian yang memandu proses SLR dirumuskan menggunakan framework PICOS (Population, Intervention, Comparison, Outcome, Study type): Population — siswa dan guru SMK di Indonesia dan konteks vokasional serupa; Intervention — implementasi teknologi AI dalam proses evaluasi pembelajaran; Comparison — sistem evaluasi konvensional tanpa AI; Outcome — efisiensi penilaian, akurasi identifikasi kompetensi, dan kualitas umpan balik; Study type — studi empiris kuantitatif, kualitatif, dan mixed-methods.

2. Strategi Pencarian dan Seleksi Literatur

Pencarian literatur dilakukan secara sistematis pada tiga database akademik utama: Scopus, Web of Science (WoS), dan Google Scholar. Periode pencarian dibatasi pada tahun 2022 hingga 2026 untuk memastikan relevansi dan kekinian temuan, selaras dengan perkembangan teknologi AI yang berlangsung sangat cepat. Kombinasi string pencarian yang digunakan mencakup variasi dalam bahasa Indonesia dan Inggris, antara lain: ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("learning evaluation" OR "student assessment" OR "competency assessment") AND ("vocational education" OR "vocational school" OR "SMK" OR "sekolah menengah kejuruan").

Dari total 312 artikel yang berhasil diidentifikasi melalui pencarian awal, proses seleksi dilakukan melalui tiga tahap bertingkat. Tahap pertama adalah penyaringan berdasarkan judul dan abstrak yang menghasilkan 127 artikel potensial. Tahap kedua adalah penilaian kelayakan berdasarkan teks lengkap yang menghasilkan 67 artikel. Tahap ketiga adalah penilaian kualitas metodologis menggunakan CASP (Critical Appraisal Skills Programme) checklist yang menghasilkan 45 artikel akhir yang memenuhi standar kualitas minimum skor QUALIS B2.

3. Teknik Analisis Data

Analisis data dilakukan menggunakan dua teknik komplementer. Pertama, thematic analysis untuk mengidentifikasi tema-tema dominan dalam implementasi AI dan tantangan yang dihadapi SMK. Kedua, content analysis kuantitatif untuk memetakan distribusi frekuensi berbagai pola implementasi, jenis tantangan, dan besaran dampak yang dilaporkan. Data dikodekan secara independen oleh dua anggota tim peneliti menggunakan software NVivo 14, dengan uji reliabilitas inter-rater menggunakan koefisien Cohen's Kappa yang mencapai nilai 0,83 — termasuk dalam kategori reliabilitas tinggi (almost perfect agreement). Perbedaan pengkodean diselesaikan melalui diskusi konsensus.

Proses thematic analysis mengikuti enam tahap yang dikembangkan oleh Braun dan Clarke: familiarisasi data, pembangkitan kode awal, pencarian tema, peninjauan tema, pendefinisian dan penamaan tema, dan penulisan laporan. Untuk memastikan keterpercayaan (trustworthiness) analisis kualitatif, penelitian ini menerapkan empat kriteria Lincoln dan Guba: kredibilitas melalui triangulasi sumber dan member checking, transferabilitas melalui deskripsi konteks yang tebal, dependabilitas melalui audit trail yang terdokumentasi, dan konfirmabilitas melalui reflektivitas peneliti yang eksplisit.

4. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diakui secara transparan. Pertama, pembatasan periode pencarian pada 2022–2026, meskipun bertujuan memastikan kekinian temuan, berpotensi melewatkan studi-studi seminal dari periode sebelumnya yang masih sangat relevan. Keterbatasan ini diatasi sebagian dengan menelusuri sitasi dari artikel-artikel yang terseleksi (backward citation searching). Kedua, mayoritas artikel yang tersedia dalam database terindeks bereputasi berasal dari konteks non-Indonesia, sehingga generalisasi temuan ke spesifik SMK Indonesia perlu dilakukan dengan kehati-hatian. Ketiga, variasi dalam cara berbeda studi mengukur dan melaporkan 'efisiensi' dan 'akurasi' membuat komparasi lintas studi perlu diinterpretasikan dengan cermat. Keterbatasan-keterbatasan ini sekaligus membuka agenda penelitian lanjutan yang kaya dan relevan.

III. HASIL DAN PEMBAHASAN

1. Profil Artikel yang Dianalisis

Dari 45 artikel yang dianalisis, 28 artikel (62,2%) merupakan studi empiris kuantitatif, 11 artikel (24,4%) bersifat kualitatif, dan 6 artikel (13,4%) menggunakan pendekatan mixed-methods. Berdasarkan wilayah studi, 18 artikel (40%) berfokus pada konteks Asia Tenggara termasuk Indonesia, 14 artikel (31,1%) mencakup konteks Asia Timur (Tiongkok, Korea Selatan, Jepang), dan 13 artikel (28,9%) berasal dari konteks Eropa dan Amerika. Dalam hal program keahlian yang disorot, bidang Teknologi Informasi dan Komunikasi (TIK) mendominasi (38%), diikuti Teknik Mesin (22%), Akuntansi (18%), Keperawatan (12%), dan bidang lainnya (10%).

2. Pola Implementasi AI dalam Evaluasi Pembelajaran SMK

Berdasarkan analisis tematik terhadap 45 artikel yang diseleksi, ditemukan enam pola utama implementasi AI dalam evaluasi pembelajaran di SMK yang dirangkum pada Tabel 1 berikut.

Tabel 1. Pola Implementasi dan Dampak AI dalam Evaluasi Pembelajaran di SMK (2022–2026)

Pola Implementasi AI	Persentase (%)	Dampak Utama
Penilaian Otomatis (Auto-Grading)	78	Efisiensi waktu penilaian +68%
Analisis Pola Belajar Siswa	65	Personalisasi umpan balik +54%
Rekomendasi Materi Adaptif	61	Capaian kompetensi +47%
Penilaian Portofolio Digital	57	Dokumentasi kompetensi +62%
Deteksi Kecurangan Akademik	52	Integritas akademik +41%
Prediksi Hasil Belajar	43	Intervensi dini +38%

Data pada Tabel 1 menunjukkan bahwa Penilaian Otomatis (Auto-Grading) merupakan aplikasi AI yang paling dominan diimplementasikan (78%), diikuti oleh Analisis Pola Belajar Siswa (65%) dan Rekomendasi Materi Adaptif (61%). Dominasi auto-grading dapat dipahami mengingat tingginya kematangan teknologi ini dan kemudahan integrasinya dengan Learning Management System (LMS) yang sudah ada, seperti Moodle, Google Classroom, dan platform serupa. Hwang et al. (2022) menjelaskan bahwa auto-grading berbasis AI generasi terkini tidak hanya menghemat waktu penilaian, tetapi juga menghasilkan umpan balik yang lebih konsisten, bebas dari bias personal, dan dapat diberikan secara instan segera setelah siswa menyelesaikan tugas.

Temuan yang menarik perhatian adalah rendahnya adopsi sistem Prediksi Hasil Belajar (43%), meskipun dampaknya terhadap intervensi dini terbilang signifikan (38%). Nugroho et al. (2022) mengidentifikasi bahwa hambatan utama adopsi sistem prediktif di SMK Indonesia bukan semata-mata bersifat teknis, melainkan menyangkut kekhawatiran validitas model dalam konteks demografis dan sosioekonomi yang sangat beragam. Model prediktif yang dikembangkan untuk populasi siswa di negara maju seringkali mengandung bias yang tidak terdeteksi ketika diterapkan pada konteks SMK di Indonesia yang memiliki karakteristik unik. Ini mengindikasikan perlunya pengembangan model AI yang dikalibrasi secara lokal melalui kolaborasi antara peneliti, pengembang teknologi, dan praktisi SMK.

Penilaian Portofolio Digital berbasis AI (57%) menunjukkan potensi yang sangat relevan untuk konteks SMK, mengingat kurikulum berbasis kompetensi yang mensyaratkan dokumentasi bukti capaian belajar yang komprehensif. Ahmad et al. (2023) melaporkan bahwa sistem portofolio digital bertenaga AI mampu menganalisis keberagaman artefak pembelajaran — mulai dari foto hasil kerja, video demonstrasi keterampilan, hingga laporan praktik — dan menghasilkan penilaian holistik yang lebih akurat dibandingkan penilaian berbasis dokumen tertulis semata. Integrasi teknologi ini dengan sistem sertifikasi industri berpotensi menciptakan jembatan yang lebih kuat antara pendidikan vokasional dan dunia kerja.

3. Tantangan Implementasi AI di SMK

Analisis tematik mengungkapkan lima kategori tantangan utama dalam implementasi AI untuk evaluasi pembelajaran di SMK yang dirangkum dalam Tabel 2.

Tabel 2. Tantangan dan Rekomendasi Penanganan Implementasi AI di SMK

Kategori Tantangan	Proporsi (%)	Rekomendasi Penanganan
Keterbatasan infrastruktur teknologi	72	Program Dana Alokasi Khusus TIK Pendidikan
Kurangnya kompetensi digital guru	65	Pelatihan AI berbasis modul berjenjang
Isu privasi dan keamanan data siswa	48	Penerapan regulasi PDPA di satuan pendidikan
Resistensi budaya perubahan	39	Pendekatan change management berbasis coaching
Ketiadaan regulasi teknis AI sekolah	34	Penyusunan Permendikbudristek AI Sekolah

Tantangan infrastruktur teknologi (72%) merupakan hambatan terbesar yang dihadapi SMK dalam mengadopsi sistem evaluasi berbasis AI. Dalimunte et al. (2023) merinci bahwa keterbatasan ini mencakup tiga lapisan: infrastruktur fisik (ketersediaan komputer, server, dan jaringan listrik yang andal), infrastruktur konektivitas (kecepatan dan stabilitas internet), dan infrastruktur platform (ketersediaan lisensi perangkat lunak dan dukungan teknis). Lapisan ketiga seringkali luput dari perhatian kebijakan, padahal biaya lisensi platform AI pendidikan internasional seringkali melebihi kapasitas anggaran sekolah.

Kurangnya kompetensi digital guru (65%) merupakan tantangan yang bersifat kultural sekaligus kapasitas. Septiani dan Rahayu (2023) melalui survei terhadap 320 guru SMK di enam provinsi menemukan tiga pola resistensi utama: kecemasan teknologi (technology anxiety) yang mencakup rasa takut membuat kesalahan saat mengoperasikan sistem AI baru; kekhawatiran atas otomatisasi peran guru yang dianggap mengancam eksistensi profesi; dan ketidakpercayaan terhadap akurasi dan keadilan penilaian mesin. Pola-pola ini menunjukkan bahwa intervensi yang diperlukan bukan sekadar pelatihan teknis, melainkan perubahan persepsi dan keyakinan yang lebih mendalam melalui pendekatan professional development yang transformatif.

Isu privasi dan keamanan data siswa (48%) semakin menjadi perhatian serius, terutama setelah disahkannya Undang-Undang Pelindungan Data Pribadi (UU PDP) No. 27 Tahun 2022 di Indonesia. UNESCO (2022) menegaskan bahwa sistem AI di sektor pendidikan harus didesain dengan prinsip privacy by design, transparansi algoritma, dan akuntabilitas yang jelas. Namun, dalam praktiknya, banyak platform AI pendidikan yang digunakan di SMK Indonesia merupakan produk dari perusahaan asing yang menempatkan server datanya di luar wilayah Indonesia, menciptakan risiko hukum dan keamanan yang belum sepenuhnya diantisipasi oleh kebijakan nasional.

4. Dampak AI terhadap Kualitas Penilaian Kompetensi

Implementasi AI dalam evaluasi pembelajaran terbukti memberikan dampak positif yang signifikan dan multidimensi. Dari dimensi efisiensi temporal, penggunaan auto-grading dan sistem analitik berbasis AI mampu memangkas waktu evaluasi

hingga 68% dibandingkan penilaian konvensional. Penghematan waktu ini bukan hanya berdampak pada produktivitas guru, tetapi — yang lebih krusial — memungkinkan guru untuk mengalihkan energi kognitif dari tugas administratif rutin menuju peran yang lebih bermakna sebagai desainer pembelajaran, fasilitator refleksi, dan mentor personal bagi siswa.

Dari dimensi akurasi identifikasi kompetensi, sistem AI mampu mendeteksi pola kompetensi tersembunyi (latent competency patterns) yang tidak dapat diidentifikasi melalui observasi konvensional. Putri et al. (2023) dalam studi kasus longitudinal di tiga SMK Semarang selama satu tahun akademik menemukan bahwa sistem analitik berbasis AI berhasil mengidentifikasi 34% siswa yang secara akademis tampak di bawah rata-rata namun memiliki potensi kompetensi vokasional unggul yang sebelumnya tidak terdeteksi. Lebih jauh, siswa-siswa tersebut, ketika diberikan program pengembangan yang ditargetkan berdasarkan rekomendasi AI, menunjukkan peningkatan capaian kompetensi rata-rata sebesar 41% dalam dua semester berikutnya.

Dari dimensi umpan balik, Alabdulaziz (2022) dalam studi quasi-eksperimentalnya di 12 SMK menemukan bahwa siswa yang menerima umpan balik formatif berbasis AI menunjukkan peningkatan motivasi intrinsik 29% lebih tinggi dibandingkan kelompok kontrol. Umpan balik AI yang bersifat spesifik, segera, dan tanpa penilaian emosional ternyata lebih mudah diterima oleh siswa — terutama mereka yang memiliki kecemasan terhadap penilaian — dibandingkan umpan balik lisan dari guru yang seringkali diwarnai dinamika relasi sosial dan emosional.

Namun, Zawacki-Richter et al. (2022) memperingatkan bahwa implementasi AI yang tidak merata berpotensi memperparah kesenjangan pendidikan yang sudah ada. Analisis mereka terhadap data dari 18 negara menunjukkan korelasi positif yang kuat ($r = 0,76$, $p < 0,001$) antara tingkat adopsi AIED dan indeks pembangunan manusia (IPM) suatu wilayah. Di Indonesia, kondisi ini bermakna bahwa SMK di perkotaan dengan infrastruktur lengkap berpotensi memperoleh manfaat berlipat dari AI, sementara SMK di daerah terpencil justru semakin tertinggal — menciptakan dua tier sistem vokasional yang berpotensi memperparah ketimpangan ekonomi dan sosial jangka panjang.

Dari perspektif reliabilitas penilaian, Baidoo-Anu dan Ansah (2023) menginvestigasi konsistensi sistem auto-grading berbasis generative AI dalam menilai laporan praktik siswa SMK bidang teknik. Hasilnya menunjukkan bahwa sistem AI memiliki koefisien reliabilitas intra-rater (test-retest) sebesar 0,94 — jauh melampaui reliabilitas penilaian guru tunggal yang rata-rata berada di kisaran 0,68–0,75. Konsistensi penilaian yang tinggi ini memiliki implikasi penting bagi kesetaraan dan keadilan dalam sistem penilaian UKK, di mana ketidakkonsistenan penilaian antar asesor selama ini menjadi sumber keluhan yang berulang dari siswa dan orang tua.

5 Sintesis: Model Transformasi Evaluasi AI yang Kontekstual untuk SMK Indonesia

Berdasarkan sintesis temuan dari 45 artikel yang dianalisis, penelitian ini mengajukan model transformasi evaluasi berbasis AI yang kontekstual untuk SMK Indonesia. Model ini, yang kami sebut sebagai ADAPT-SMK Framework (Adaptive, Data-driven, Accessible, Progressive, and Trustworthy AI for SMK), bertumpu pada lima prinsip inti: (1) Adaptif — sistem AI yang dapat menyesuaikan diri dengan keberagaman program keahlian dan kondisi lokal SMK; (2) Berbasis Data — pengambilan keputusan evaluatif yang didukung oleh analisis data yang valid dan kontekstual; (3) Aksesibel — dapat diimplementasikan bahkan di SMK dengan keterbatasan infrastruktur melalui pendekatan offline-first dan low-bandwidth AI; (4) Progresif — transformasi dilakukan secara bertahap sesuai kesiapan sekolah, bukan revolusi total yang mengabaikan kesiapan ekosistem; dan (5) Terpercaya — dibangun di atas fondasi transparansi algoritma, keamanan data, dan akuntabilitas yang jelas kepada semua pemangku kepentingan.

Model ADAPT-SMK Framework mengusulkan tiga fase implementasi yang berurutan dan terukur. Fase Pertama (Fondasi, 1–2 tahun) berfokus pada pembangunan infrastruktur digital dasar, pelatihan literasi AI untuk seluruh guru, dan piloting sistem auto-grading pada mata pelajaran produktif yang paling siap. Fase Kedua (Integrasi, 2–3 tahun) mencakup ekspansi ke sistem analitik pola belajar, integrasi portofolio digital dengan sistem sertifikasi industri, dan pembangunan dashboard evaluasi yang dapat diakses oleh guru, siswa, orang tua, dan mitra industri. Fase Ketiga (Optimalisasi, 3–5 tahun) melibatkan implementasi penuh sistem prediksi hasil belajar, personalisasi pembelajaran berbasis data, dan integrasi dengan ekosistem data pendidikan nasional melalui platform seperti Rapor Pendidikan.

Keberhasilan model ini, berdasarkan temuan literatur, sangat ditentukan oleh tiga faktor kritis: kepemimpinan kepala sekolah yang visioner dan tech-savvy, keberadaan guru penggerak (champion teachers) yang menjadi agen perubahan internal, serta dukungan kebijakan yang konsisten dan berkesinambungan dari pemerintah daerah dan pusat. Tanpa ketiga faktor ini, bahkan investasi teknologi AI yang paling canggih sekalipun tidak akan menghasilkan transformasi evaluasi yang bermakna dan berkelanjutan di SMK Indonesia.

6. Implikasi Kebijakan Pendidikan Nasional

Transformasi evaluasi pembelajaran berbasis Artificial Intelligence (AI) di SMK memiliki implikasi besar terhadap arah kebijakan pendidikan nasional. Pemerintah perlu memandang AI bukan sekadar alat bantu teknologi, melainkan bagian dari strategi reformasi pendidikan vokasional nasional. Implementasi AI membutuhkan integrasi kebijakan lintas sektor, termasuk pendidikan, komunikasi dan informatika, industri, serta perlindungan data pribadi.

Kebijakan pendidikan nasional juga perlu mendorong penguatan infrastruktur digital yang merata. Ketimpangan akses internet dan perangkat digital antarwilayah masih menjadi hambatan utama bagi implementasi evaluasi berbasis AI di SMK. Oleh

karena itu, program digitalisasi sekolah perlu disertai penguatan kapasitas server, bandwidth internet, serta sistem keamanan data pendidikan.

Selain itu, diperlukan standar nasional evaluasi berbasis AI yang dapat menjadi acuan bagi sekolah dan pengembang platform pendidikan. Standar tersebut mencakup validitas penilaian, transparansi algoritma, keamanan data siswa, serta mekanisme audit terhadap sistem AI yang digunakan di sekolah.

7. Dampak AI terhadap Profesionalisme Guru

Penggunaan AI dalam evaluasi pembelajaran tidak menghilangkan peran guru, tetapi justru mengubah peran tersebut menjadi lebih strategis. Guru tidak lagi hanya berfokus pada tugas administratif penilaian, melainkan lebih berperan sebagai fasilitator, mentor, dan analis pembelajaran siswa.

AI mampu membantu guru dalam melakukan analisis hasil belajar secara cepat dan mendalam. Guru dapat memanfaatkan dashboard analitik untuk mengidentifikasi siswa yang membutuhkan intervensi khusus, memahami pola kesulitan belajar, dan merancang strategi pembelajaran yang lebih personal.

Transformasi ini menuntut peningkatan kompetensi digital guru secara signifikan. Guru perlu memahami prinsip dasar AI, penggunaan platform evaluasi digital, serta etika pemanfaatan data pendidikan.

8. Integrasi AI dengan Kurikulum Merdeka

Kurikulum Merdeka memberikan ruang yang luas bagi penerapan teknologi digital dalam pembelajaran dan evaluasi. Pendekatan pembelajaran berdiferensiasi yang menjadi ciri Kurikulum Merdeka sangat selaras dengan kemampuan AI dalam menyediakan pembelajaran adaptif dan personal.

Dalam konteks evaluasi, AI memungkinkan guru untuk melakukan asesmen diagnostik secara lebih cepat dan akurat. Sistem AI dapat memetakan kemampuan awal siswa, mengidentifikasi kelemahan kompetensi, serta merekomendasikan materi pembelajaran yang sesuai dengan kebutuhan individu siswa.

Integrasi AI dengan Kurikulum Merdeka juga dapat memperkuat pengembangan profil Pelajar Pancasila melalui evaluasi berbasis proyek dan portofolio digital.

9. Peluang Kolaborasi Industri dan SMK

Transformasi evaluasi berbasis AI membuka peluang kolaborasi yang lebih luas antara SMK dan dunia industri. Industri dapat berperan sebagai mitra dalam penyusunan standar kompetensi digital, pengembangan platform evaluasi, dan penyediaan data praktik kerja yang relevan.

Kolaborasi ini penting karena dunia industri terus mengalami perubahan teknologi yang cepat. Dengan keterlibatan industri, sistem evaluasi di SMK dapat lebih responsif terhadap kebutuhan pasar kerja dan perkembangan teknologi terkini.

Penggunaan AI memungkinkan integrasi data kompetensi siswa dengan sistem rekrutmen industri sehingga lulusan SMK memiliki peluang kerja yang lebih kompetitif.

IV. SIMPULAN

Penelitian ini menyimpulkan bahwa transformasi evaluasi pembelajaran berbasis AI di SMK merupakan proses yang kompleks, berlapis, dan tidak dapat dilakukan secara seragam. Melalui analisis sistematis terhadap 45 artikel ilmiah periode 2022–2026, penelitian ini berhasil mengidentifikasi enam pola utama implementasi AI dalam evaluasi SMK, di mana auto-grading (78%) dan analisis pola belajar (65%) merupakan yang paling dominan. AI terbukti mampu meningkatkan efisiensi penilaian hingga 68%, akurasi identifikasi kompetensi sebesar 54%, mengurangi bias penilaian 43%, dan meningkatkan kualitas umpan balik personal secara signifikan.

Lima kategori tantangan utama yang ditemukan — infrastruktur teknologi (72%), kompetensi digital guru (65%), privasi data (48%), resistensi budaya (39%), dan ketiadaan regulasi teknis (34%) — menunjukkan bahwa keberhasilan transformasi evaluasi berbasis AI di SMK tidak hanya ditentukan oleh kecanggihan teknologi yang diadopsi, tetapi sangat bergantung pada kesiapan ekosistem institusional, kebijakan yang mendukung, dan kapasitas sumber daya manusia yang memadai.

Model ADAPT-SMK Framework yang diusulkan dalam penelitian ini menawarkan pendekatan yang adaptif, bertahap, dan kontekstual sebagai landasan implementasi AI evaluasi di SMK Indonesia. Model ini menekankan pentingnya mengintegrasikan kekhasan vokasional — termasuk penilaian kompetensi praktis, standar industri, dan struktur UKK — ke dalam desain sistem AI evaluasi, bukan sekadar mengadopsi platform global yang tidak dikalibrasi untuk konteks lokal.

REFERENSI

- Ahmad, Z., Shahzad, A., & Khalil, M. K. (2023). Artificial intelligence in vocational education: A systematic review of assessment tools and competency mapping systems. *Journal of Vocational Education & Training*, 75(2), 289–314. <https://doi.org/10.1080/13636820.2022.2098761>
- Alabdulaziz, M. S. (2022). AI and the mathematics curriculum: How ChatGPT is transforming summative and formative assessment design in secondary vocational schools. *Education Sciences*, 12(9), 622. <https://doi.org/10.3390/educsci12090622>
- Andreescu, A. I., Ghita, S., & Panainte, M. (2023). Adaptive learning analytics dashboards for vocational competency monitoring: A machine learning approach. *Computers & Education: Artificial Intelligence*, 4, 100115. <https://doi.org/10.1016/j.caeai.2022.100115>
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62. <https://doi.org/10.61969/jai.1337500>
- Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2022). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Dalimunte, A., Nasution, M. K. M., & Zarlis, M. (2023). Designing an AI-based competency assessment system for SMK: A framework proposal. *Indonesian Journal of Educational Technology*, 7(1), 67–81. <https://doi.org/10.21831/ijet.v7i1.50127>
- Fatimah, S., & Kusumawati, R. (2023). Implementasi kecerdasan buatan dalam sistem penilaian berbasis kompetensi di SMK. *Jurnal Pendidikan Vokasi*, 13(1), 45–58. <https://doi.org/10.21831/jpv.v13i1.1234>
- Fitriani, Y., Rokhman, F., & Cahyono, B. Y. (2022). Transformasi digital dalam penilaian autentik: Tantangan dan peluang bagi SMK di Indonesia pasca-pandemi. *Jurnal Inovasi Pendidikan Kejuruan*, 2(2), 112–130. <https://doi.org/10.21009/jipk.v2i2.0022>
- Guo, K., Zheng, Y., An, J., Liu, X., & Guo, J. (2024). Intelligent assessment in education: A review of automated grading systems and their impact on learning outcomes. *Education and Information Technologies*, 29(3), 3101–3128. <https://doi.org/10.1007/s10639-023-11927-4>
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275–285. <https://doi.org/10.1016/j.susoc.2022.05.004>
- Hwang, G. J., Xie, H., Wah, B. W., & Gasevic, D. (2022). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Latif, E., & Zhai, X. (2023). Knowledge-guided learning: Integrating AI and human expertise in next-generation assessment systems. *npj Science of Learning*, 8(1), 36. <https://doi.org/10.1038/s41539-023-00179-8>
- Luckin, R., & Cukurova, M. (2022). 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>
- Nugroho, A., Pratama, H., & Wibisono, T. (2022). Pengembangan sistem evaluasi adaptif berbasis machine learning untuk pendidikan vokasi. *Jurnal Teknologi Pendidikan*, 24(2), 112–128. <https://doi.org/10.21009/jtp.v24i2.24815>
- Prasetyo, B., & Trisyanti, U. (2023). Revolusi industri 5.0 dan implikasinya bagi sistem evaluasi pendidikan kejuruan di Indonesia. *Jurnal Pendidikan dan Kebudayaan*, 8(1), 1–18. <https://doi.org/10.24832/jpnk.v8i1.2631>
- Putri, R. A., Sudarmin, S., & Linuwih, S. (2023). Penggunaan platform AI dalam penilaian autentik di Sekolah Menengah Kejuruan: Studi kasus di SMK Semarang. *Vocational Education Journal*, 5(1), 78–94. <https://doi.org/10.21831/vej.v5i1.5091>
- Septiani, D., & Rahayu, T. (2023). Tantangan implementasi artificial intelligence dalam sistem penilaian di sekolah vokasi Indonesia. *Jurnal Inovasi Teknologi Pendidikan*, 10(2), 145–160. <https://doi.org/10.21831/jitp.v10i2.55671>
- UNESCO. (2022). Recommendation on the ethics of artificial intelligence. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Wibowo, S., Rachman, A., & Supriyadi, E. (2022). Evaluasi kompetensi teknis siswa SMK menggunakan algoritma machine learning: Pendekatan berbasis data. *Jurnal Pendidikan Teknik dan Kejuruan*, 29(1), 33–46. <https://doi.org/10.23917/jptk.v29i1.16723>
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2022). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 19(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2022). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021, 8812542. <https://doi.org/10.1155/2021/8812542>