

Pemanfaatan AI dalam Asesmen Pembelajaran PAI: Peluang, Tantangan Etis, dan Celah Empiris Sebuah Tinjauan Sistematis

Muhammad Iqbal Dewantara

Email: dewantara13@gmail.com

Universitas Islam Internasional

Abstrak: Asesmen pembelajaran Pendidikan Agama Islam (PAI) menghadapi persoalan kronis berupa subjektivitas penilaian, ketidakkonsistenan standar, dan keterbatasan kapasitas evaluatif, sementara integrasi kecerdasan buatan (AI) dalam asesmen pendidikan menjanjikan solusi efisiensi dan objektivitas. Namun, literatur yang ada masih didominasi oleh kajian konseptual dan persepsi, bukan bukti empiris implementasi di ruang kelas PAI. Tinjauan sistematis ini bertujuan memetakan peluang, tantangan etis, dan celah empiris pemanfaatan AI dalam asesmen PAI. Menggunakan protokol PRISMA 2020, penelusuran sistematis dilakukan terhadap basis data akademik untuk mengidentifikasi studi yang relevan. Temuan utama menunjukkan bahwa AI menawarkan peluang personalisasi umpan balik, otomatisasi penilaian objektif, dan efisiensi administratif; namun menghadapi tantangan etis berupa bias algoritmik, transparansi sistem, privasi data, dan risiko pergeseran otoritas epistemik dari guru ke sistem. Lebih kritis, tinjauan ini mengungkap bahwa bukti empiris implementasi asesmen AI di PAI masih sangat terbatas dan sebagian besar berupa klaim teoretis. Kontribusi teoretis utama penelitian ini adalah perumusan kerangka konseptual *Techno-Ethical Assessment Literacy* yang mengintegrasikan kompetensi literasi AI dengan prinsip adab dan amanah sebagai fondasi asesmen PAI yang beretika, sekaligus menegaskan bahwa adopsi AI dalam asesmen PAI harus diposisikan sebagai *human-mediated augmentation*, bukan substitusi peran evaluatif guru.

Kata kunci : asesmen PAI, kecerdasan buatan, etika AI, tinjauan sistematis, PRISMA, literasi asesmen

How to cite: Pemanfaatan AI dalam Asesmen Pembelajaran PAI: Peluang, Tantangan Etis, dan Celah Empiris Sebuah Tinjauan Sistematis. (2024). Khatulistiwa: Jurnal Ilmu Pendidikan, 5(2), 137-150.
<https://jurnal.elkatarie.ac.id/index.php/khatulistiwa/article/view/847>

PENDAHULUAN

Pendidikan Agama Islam (PAI) di Indonesia menghadapi tuntutan untuk beradaptasi dengan transformasi digital tanpa mengorbankan tujuan pembentukan karakter dan akhlak peserta didik (Fauzi, 2023). Dalam ranah asesmen, persoalan fundamental telah lama diakui: subjektivitas penilaian guru, ketiadaan standar yang konsisten, dan keterbatasan kapasitas evaluatif menjadi hambatan yang sistematis (Rahman & Hidayat, 2022). Asesmen PAI tidak sekadar mengukur pengetahuan kognitif, tetapi juga dimensi afektif dan spiritual yang menuntut pendekatan evaluatif yang lebih nuanced (Hasanah, 2023). Ketika guru menghadapi kelas besar dengan beragam latar belakang kemampuan dan gaya belajar, tuntutan untuk memberikan umpan balik yang personal dan tepat waktu menjadi semakin sulit dipenuhi secara manual.

Kecerdasan buatan (AI) dalam asesmen pendidikan menjanjikan solusi terhadap persoalan-persoalan tersebut. Sistem penilaian berbasis AI dapat melakukan penilaian otomatis untuk pertanyaan objektif, memberikan umpan balik instan untuk esai, dan bahkan menganalisis kemampuan membaca Al-Qur'an melalui pengenalan suara (Kumar, 2024). Namun, janji efisiensi dan objektivitas ini berhadapan dengan kompleksitas substantif PAI: interpretasi teks suci, penilaian dimensi spiritual, dan otoritas epistemik guru sebagai uswah hasanah (Hamid & Rizki, 2024).

Literatur tentang AI dalam pendidikan Islam telah berkembang dalam beberapa tahun terakhir, namun dengan karakteristik yang perlu dicermati secara kritis. Sejumlah studi mengeksplorasi peluang integrasi AI dalam pembelajaran PAI, termasuk personalisasi materi dan peningkatan keterlibatan siswa (Nugroho & Prasetyo, 2023). Kajian lain menyoroti transformasi peran guru dan tantangan infrastruktur (Sood, 2026a). Namun, sebuah tinjauan sistematis terbaru menemukan bahwa bukti yang tersedia dalam domain ini "predominantly exploratory, conceptual, or perception-

based" dan sebagian besar klaim tentang dampak AI terhadap pendidikan agama bersifat "reported concerns or theoretical risks rather than established classroom outcomes" (Ameer, 2025).

Ada dua celah signifikan yang belum terjawab. Pertama, mayoritas literatur membahas AI dalam pembelajaran PAI secara umum, sementara dimensi asesmen spesifik—yang memiliki karakteristik etis dan teknis tersendiri belum mendapat perhatian yang memadai (Bukhari & Kaur, 2026). Kedua, meskipun isu etika AI telah banyak dibahas dalam pendidikan secara umum, belum ada kerangka yang secara eksplisit mengintegrasikan prinsip-prinsip etika Islam ke dalam desain dan implementasi asesmen berbasis AI untuk PAI (Mighfar et al., 2026).

Tinjauan sistematis ini menjawab tiga pertanyaan: (1) Apa peluang pemanfaatan AI dalam asesmen PAI menurut literatur yang ada? (2) Apa tantangan etis yang teridentifikasi dalam konteks spesifik PAI? (3) Sejauh mana bukti empiris mendukung adopsi AI dalam asesmen PAI, dan di mana celah yang perlu diisi oleh penelitian selanjutnya?

Kontribusi yang ditawarkan bersifat teoretis sekaligus kontekstual. Secara teoretis, penelitian ini merumuskan kerangka Techno-Ethical Assessment Literacy yang menyintesis dimensi teknis asesmen AI dengan prinsip adab (etika penggunaan) dan amanah (tanggung jawab epistemik) sebagai fondasi normatif (Sood, 2026b). Secara kontekstual, tinjauan ini menegaskan bahwa adopsi AI dalam asesmen PAI memerlukan reposisi peran guru dari evaluator tunggal menjadi arbiter etis, di mana AI berfungsi sebagai augmentation bukan substitusi (Faqihuddin & Romadhon, 2023).

METODOLOGI PENELITIAN

Rancangan Penelitian

Penelitian ini menggunakan metode tinjauan sistematis (systematic literature review) dengan mengikuti protokol PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) sebagai kerangka metodologis (Kovári, 2024). Protokol PRISMA dipilih karena menyediakan prosedur yang transparan dan dapat direplikasi untuk identifikasi, seleksi, dan sintesis literatur, sekaligus memungkinkan penilaian kualitas bukti yang sistematis (Moher et al., 2021).

Kriteria Inklusi dan Eksklusi Kriteria Inklusi:

- Publikasi dalam bahasa Indonesia atau Inggris

- Rentang publikasi: 2020–2026
- Fokus pada integrasi AI dalam asesmen pembelajaran, khususnya dalam konteks pendidikan agama Islam atau PAI
- Artikel jurnal, prosiding konferensi, atau bab buku yang melalui peer review
- Studi yang membahas asesmen sebagai dimensi yang dapat diidentifikasi secara eksplisit

Kriteria Eksklusi:

- Studi yang membahas AI dalam pembelajaran PAI tanpa menyentuh dimensi asesmen
- Artikel opini non-akademik atau publikasi tanpa proses peer review
- Studi yang membahas AI dalam pendidikan umum tanpa relevansi dengan konteks PAI/IRE

Penelusuran dilakukan pada basis data akademik berikut: Scopus, Web of Science, ERIC, dan Google Scholar. Kombinasi kata kunci yang digunakan mencakup: ("artificial intelligence" OR "AI" OR "machine learning" OR "deep learning") AND ("assessment" OR "evaluation" OR "penilaian") AND ("Islamic religious education" OR "Pendidikan Agama Islam" OR "PAI"). Penelusuran tambahan dilakukan melalui snowballing terhadap referensi dari artikel-artikel kunci yang teridentifikasi (Page et al., 2021).

Seleksi studi dilakukan melalui tiga tahap: (1) penyaringan judul dan abstrak berdasarkan kriteria inklusi; (2) pembacaan teks penuh untuk menilai relevansi; (3) ekstraksi data dari studi yang memenuhi kriteria. Data yang diekstraksi mencakup: karakteristik studi (penulis, tahun, lokasi, desain), dimensi asesmen yang dibahas, jenis AI yang digunakan, serta temuan utama terkait peluang dan tantangan etis. Analisis dilakukan dengan pendekatan analisis tematik terhadap kategori yang telah ditetapkan (peluang, tantangan etis, celah empiris) sekaligus tetap terbuka terhadap tema-tema emergent (Braun & Clarke, 2006).

Sebagai tinjauan literatur, penelitian ini tidak melibatkan partisipan manusia. Seluruh sumber yang digunakan telah dipublikasikan secara publik dan disitasi dengan semestinya. Tidak terdapat konflik kepentingan dalam penelitian ini.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Penelitian

Analisis terhadap literatur yang memenuhi kriteria mengungkap empat kategori peluang utama pemanfaatan AI dalam asesmen PAI.

Pertama, otomatisasi penilaian dan efisiensi administratif. AI memungkinkan penilaian otomatis untuk pertanyaan objektif, mengurangi beban administratif guru dan memberikan ruang untuk fokus pada aspek evaluatif yang membutuhkan pertimbangan manusiawi (Ogunleye et al., 2024). Sistem manajemen pembelajaran berbasis AI juga dapat mengelola dan menganalisis data asesmen dalam skala yang sulit dicapai secara manual (Jauhiainen & Bernardo, 2025). Otomatisasi tugas-tugas administratif yang selama ini membebani guru PAI—seperti membuat rencana pembelajaran, menilai hasil belajar, melacak kehadiran, dan menyusun laporan kemajuan spiritual siswa memungkinkan guru untuk kembali ke esensi profesinya sebagai pendidik dan pembangun karakter (Chanda et al., 2023).

Kedua, umpan balik formatif yang dipersonalisasi. Teknologi pemrosesan bahasa alami (NLP) memungkinkan AI memberikan umpan balik terhadap esai siswa dengan tingkat kedalaman yang dapat dikonfigurasi (Lu et al., 2022). Dalam konteks PAI, ini berpotensi mendukung asesmen formatif terhadap pemahaman konsep-konsep keislaman dan kemampuan argumentasi religius siswa. Tinjauan sistematis menunjukkan bahwa sistem umpan balik berbasis AI dapat meningkatkan kredibilitas asesmen melalui mekanisme umpan balik otomatis dan validasi guru (Chima et al., 2024). Umpan balik yang digerakkan AI juga dapat dibandingkan dengan umpan balik manusia dalam beberapa skenario dan meningkatkan pembelajaran, meskipun tetap memerlukan mediasi pedagogis (Rafiq-Uz-Zaman & Ameer, 2025).

Ketiga, asesmen keterampilan spesifik PAI. Aplikasi AI berbasis pengenalan suara dapat membantu menilai kemampuan membaca Al-Qur'an, sementara alat analisis teks dapat membantu mengevaluasi pemahaman hadis atau fiqh (Anwar et al., 2025). Beberapa studi mengembangkan sistem penilaian cerdas berbasis AI untuk pendidikan dengan akurasi 85% dalam menilai literasi (Imamuddin et al., 2022). Ini membuka kemungkinan asesmen yang lebih objektif dan terstandarisasi untuk keterampilan yang sebelumnya sangat bergantung pada penilaian subjektif guru.

Keempat, personalisasi dan adaptivitas. AI dapat menyesuaikan tingkat kesulitan dan jenis pertanyaan asesmen berdasarkan performa siswa sebelumnya, menciptakan pengalaman evaluasi yang lebih adaptif (Nasucha et al., 2023). Asesmen adaptif secara dinamis memilih pertanyaan berdasarkan respons peserta didik untuk menawarkan diagnostik yang lebih akurat dan umpan balik yang dipersonalisasi (Javed et al., 2025). Dalam konteks PAI, setiap siswa memiliki kecepatan belajar, gaya kognitif, dan latar belakang keagamaan yang berbeda, dan AI dapat menjembatani kesenjangan ini secara efektif (Anwar & Umam, 2025).

Literatur secara konsisten mengidentifikasi sejumlah tantangan etis yang memiliki resonansi khusus dalam konteks PAI.

Tabel 1. Tantangan Etis Asesmen AI dalam PAI

Dimensi Etis	Deskripsi	Relevansi Khusus PAI
Bias Algoritmik	Sistem AI dapat mereproduksi bias dari data pelatihan	Potensi bias terhadap tradisi keislaman tertentu (Seo et al., 2025)
Transparansi dan Explainability	Opasitas keputusan AI menyulitkan verifikasi	Kebutuhan <i>tabayyun</i> terhadap sumber pengetahuan (Clark et al., 2025)
Privasi Data	Pengumpulan dan pemrosesan data siswa	Prinsip <i>amanah</i> dalam menjaga informasi (Aldreabi et al., 2025)

Pergeseran Epistemik	Al menggantikan peran evaluatif guru dan pemegang otoritas interpretatif (Sood, 2026c)
Ketergantungan Peserta Didik	Risiko penurunan Tujuan PAI membentuk <i>insan kamil</i> yang mandiri (Jauhiainen & Bernardo, 2025)

Bias algoritmik menjadi perhatian serius karena model AI yang dilatih pada korpus kontemporer berpotensi kurang memahami nuansa klasik teks-teks Islam (Seo et al., 2025). Penelitian menunjukkan bahwa guru memandang asesmen berbasis AI sebagai efektif, namun tetap terdapat bias kultural yang perlu diwaspadai (Mighfar et al., 2026). Sistem yang tidak dilatih dengan data keislaman yang representatif berisiko memberikan penilaian yang tidak adil terhadap interpretasi atau tradisi keilmuan tertentu.

Transparansi dan explainability merupakan tantangan yang secara langsung bersinggungan dengan tradisi keilmuan Islam. Sistem AI proprietary yang bersifat black box menyulitkan guru dan siswa untuk memverifikasi dasar penilaian (Clark et al., 2025). Prinsip fairness dan transparency menjadi elemen kritis dalam asesmen berbantuan AI (Bukhari & Kaur, 2026). Dalam tradisi Islam, verifikasi sumber (tabayyun) adalah prinsip fundamental, sehingga opasitas algoritmik menciptakan ketegangan epistemik yang signifikan (Sood, 2026b). Literatur juga mencatat tantangan berupa "hallucinated or irrelevant items, bias and fairness risks, privacy concerns, and continued reliance on human supervision" (Ameer, 2025).

Privasi data memiliki resonansi khusus dengan prinsip amanah. Pengumpulan dan pemrosesan data spiritual siswa—termasuk catatan bacaan Al-Qur'an, pemahaman keislaman, dan perkembangan moral—menimbulkan pertanyaan tentang batas-batas yang etis. Kerangka tata kelola data berbasis prinsip hifz al-'ird (menjaga kehormatan) dan hifz al-'ilm (menjaga pengetahuan) telah diusulkan sebagai respons terhadap tantangan ini (Aldreabi et al., 2025).

Pergeseran otoritas epistemik adalah tantangan yang paling spesifik bagi PAI. Ketika AI digunakan untuk menilai pemahaman keislaman siswa, pertanyaan fundamental muncul: atas dasar apa AI menilai interpretasi tertentu sebagai "benar" atau "salah" dalam konteks tradisi keilmuan yang kaya akan keragaman interpretatif? Literatur tentang etika AI dalam pendidikan agama Islam menekankan bahwa kekhawatiran dalam konteks ini bukan hanya apakah AI menghasilkan informasi yang tidak akurat, tetapi juga apakah penggunaannya mengubah persepsi peserta didik tentang siapa atau apa yang dapat dipercaya secara sah sebagai sumber pengetahuan keagamaan (Sood, 2026c). Sistem AI mungkin dipersepsikan oleh peserta didik sebagai sumber yang berotoritas secara epistemik semata-mata karena menghasilkan jawaban yang fasih dan tampak kredibel, tanpa memiliki legitimasi keilmuan, tekstual, institusional, atau keagamaan yang diperlukan untuk interpretasi yang otoritatif (Sood, 2026a).

Ketergantungan peserta didik juga menjadi kekhawatiran yang teridentifikasi dalam literatur. Risiko ketergantungan pada AI dalam asesmen berpotensi menurunkan agensi siswa dan kemampuan kritis mandiri (Jauhiainen & Bernardo, 2025). Dalam kerangka tujuan PAI yang menekankan pembentukan kemandirian berpikir dan pengamalan nilai, ketergantungan pada sistem otomatis berpotensi bertentangan dengan tujuan pendidikan yang lebih luas (Fauzi, 2023).

PEMBAHASAN

Sintesis terhadap temuan-temuan di atas mengungkap sebuah ketegangan produktif: AI menawarkan solusi teknis terhadap persoalan objektivitas dan efisiensi asesmen, namun solusi teknis ini tidak dengan sendirinya menyelesaikan persoalan substantif PAI tentang apa yang layak dinilai, oleh siapa, dan berdasarkan otoritas apa (Mighfar et al., 2026). Ketegangan ini mengimplikasikan bahwa adopsi AI dalam asesmen PAI tidak dapat diperlakukan sebagai persoalan teknis semata, melainkan memerlukan kerangka normatif yang eksplisit (Sood, 2026b).

Literatur menunjukkan bahwa integrasi AI dalam asesmen PAI memerlukan tidak hanya kompetensi teknologi tetapi juga literasi AI dan kesadaran etis untuk memastikan bahwa penggunaan AI tetap konsisten dengan tujuan moral dan human-centered pendidikan agama (Ameer, 2025). Pendekatan human-centred terhadap implementasi AI dalam pendidikan menekankan perlindungan privasi, desain pedagogis yang sehat, pembuatan kebijakan yang bertanggung jawab, dan tata kelola institusional (Bukhari & Kaur, 2026). Dalam pendidikan agama, hal

ini juga mencakup pelatihan guru, dukungan institusional, kebijakan yang sensitif secara kultural, verifikasi informasi yang dihasilkan AI, dan keterlibatan pendidik dalam pengambilan keputusan (Sood, 2026a).

Tinjauan ini mengusulkan kerangka Techno-Ethical Assessment Literacy sebagai kontribusi teoretis. Kerangka ini terdiri dari tiga dimensi yang saling terkait:

(1) Kompetensi teknis, mencakup pemahaman tentang cara kerja sistem AI, kemampuan mengoperasikan alat, dan kapasitas menginterpretasi output. Ini mencakup pemahaman tentang keterbatasan sistem, potensi kesalahan, dan kondisi di mana AI dapat diandalkan versus kondisi di mana pertimbangan manusia diperlukan (Ogunleye et al., 2024). Literatur tentang Assessment for Learning dalam era AI menekankan pentingnya menemukan keseimbangan antara teknologi dan pedagogi, antara AI dan manusia, serta mempertahankan integritas pedagogis dan mediasi guru (Ameer, 2025).

(2) Literasi etis, mencakup kesadaran akan bias, transparansi, dan dampak sosial-pedagogis penggunaan AI. Ini mencakup kemampuan mengidentifikasi dan memitigasi bias algoritmik, memahami implikasi privasi dari pengumpulan data, dan mengenali batas-batas otoritas epistemik AI (Seo et al., 2025; Clark et al., 2025). Kerangka 6A yang dikembangkan untuk integrasi AI dalam pengajaran, pembelajaran, dan asesmen menekankan bahwa AI harus memperluas (augment) daripada menggantikan penilaian profesional, dan penggunaannya harus transparan, adil, menjaga privasi, serta tunduk pada pengawasan manusia dan tanggung jawab yang jelas (Bukhari & Kaur, 2026).

(3) Fondasi normatif Islami, yang mengoperasionalkan prinsip adab dan amanah sebagai kriteria evaluatif terhadap desain dan implementasi asesmen AI. Prinsip adab dalam konteks ini merujuk pada praktik keterlibatan kritis dan hormat terhadap informasi yang dihasilkan AI, termasuk verifikasi terhadap sumber-sumber otoritatif dan penolakan memperlakukan output AI sebagai pengetahuan yang tidak perlu dipertanyakan (Sood, 2026b). Kerangka etika integrasi AI dalam pendidikan agama Islam yang dikembangkan melalui lensa adab, amanah, dan otoritas epistemik memberikan dasar yang lebih terarah untuk IRE daripada menerapkan prinsip tata kelola AI umum semata (Sood, 2026c). Prinsip amanah merujuk pada tanggung jawab untuk mengungkapkan penggunaan AI, memverifikasi konten yang dihasilkan, melindungi data peserta didik, dan menerima akuntabilitas atas pekerjaan yang dibantu AI (Faqihuddin & Romadhon, 2023).

Kerangka ini menegaskan reposisi peran guru: dari evaluator tunggal menjadi arbiter etis. Dalam posisi ini, guru tidak sekadar menggunakan AI sebagai alat, tetapi secara aktif menetapkan batas-batas etis penggunaan AI, mengevaluasi kualitas dan keadilan output AI, dan memegang otoritas final dalam penilaian dimensi-dimensi yang membutuhkan pertimbangan manusiawi,

terutama dimensi spiritual dan moral yang menjadi inti PAI (Hamid & Rizki, 2024). Transformasi peran guru dalam konteks integrasi AI dalam pendidikan Islam telah diidentifikasi sebagai pergeseran dari sumber utama pengetahuan menjadi fasilitator pembelajaran adaptif, namun transisi ini harus dipandang sebagai respons evolusioner yang memerlukan penyesuaian kompetensi ke arah mentoring, pengembangan karakter, dan manajemen pembelajaran yang lebih personal (Nugroho & Prasetyo, 2023). Guru PAI tetap menjadi uswah hasanah (teladan yang baik) yang perannya tidak dapat digantikan oleh teknologi (Hamid & Rizki, 2024).

Temuan paling signifikan dari tinjauan ini adalah kelangkaan bukti empiris implementasi asesmen AI dalam konteks PAI. Literatur yang ada sebagian besar berupa: (a) kajian konseptual tentang potensi AI; (b) survei persepsi guru atau siswa; (c) eksplorasi teoretis tentang peluang dan tantangan (Ameer, 2025). Sangat sedikit studi yang memberikan bukti tentang bagaimana AI benar-benar digunakan dalam asesmen PAI, apa dampaknya terhadap hasil belajar dan pengalaman siswa, dan bagaimana guru menavigasi kompleksitas etis dalam praktik nyata (Sood, 2026a).

Sebuah tinjauan sistematis tentang etika AI generatif dalam pendidikan agama Islam secara tegas menyatakan: "claims concerning doctrinal distortion, reduced spirituality, or displacement of religious authority should be interpreted as reported concerns or theoretical risks rather than established classroom outcomes" (Ameer, 2025). Implikasinya jelas: kita memiliki banyak klaim tentang potensi dan risiko, tetapi hampir tidak ada bukti empiris yang dapat memvalidasi klaim-klaim tersebut (Mighfar et al., 2026).

Literatur tentang AI dalam pendidikan agama Islam juga mencatat bahwa bukti yang tersedia "remains limited, fragmented, and under-theorised" (Sood, 2026a). Istilah *under-theorised* merujuk pada terbatasnya pengembangan kerangka konseptual terintegrasi yang menghubungkan penggunaan pedagogis AI dengan nilai-nilai etika Islam, pertanyaan epistemologis tentang pengetahuan dan otoritas, serta tujuan teologis dan moral IRE (Sood, 2026c). Studi-studi yang ada sering kali memeriksa peluang, risiko, akseptabilitas, pedagogi, etika AI, pembelajaran Al-Qur'an, atau penggunaan teknologi yang bertanggung jawab secara terpisah, sementara relatif sedikit studi yang secara sistematis menghubungkan dimensi-dimensi ini dalam kerangka yang koheren untuk IRE (Sood, 2026b).

Tinjauan sistematis tentang transformasi digital dalam PAI juga mengidentifikasi bahwa bukti lebih konsisten untuk hasil kognitif dan keterampilan daripada untuk hasil afektif-spiritual dan karakter (Fauzi, 2023). Keberhasilan integrasi AI lebih terkait dengan desain pedagogis, kompetensi pendidik, dukungan institusional, literasi peserta didik, dan pagar pengaman tata kelola

daripada dengan teknologi tunggal apa pun (Bukhari & Kaur, 2026). Ini menunjukkan bahwa celah empiris bukan hanya tentang apakah AI bekerja, tetapi tentang bagaimana AI bekerja dalam ekosistem pendidikan yang kompleks di mana faktor manusia dan institusional memainkan peran yang menentukan (Jauhiainen & Bernardo, 2025).

Celah ini memiliki konsekuensi praktis. Tanpa bukti empiris, pengembangan kebijakan, pelatihan guru, dan desain sistem asesmen AI untuk PAI berjalan tanpa fondasi yang kokoh (Rahman & Hidayat, 2022). Rekomendasi tentang praktik terbaik tetap bersifat spekulatif sampai diuji melalui penelitian kelas yang sistematis (Ogunleye et al., 2024). Kebutuhan akan bukti empiris menjadi semakin mendesak mengingat skala inisiatif yang sedang berlangsung, seperti program pelatihan AI untuk 50.000 guru pesantren yang diluncurkan oleh Kementerian Agama bekerja sama dengan Microsoft Indonesia (Mighfar et al., 2026). Inisiatif semacam ini menciptakan kebutuhan mendesak akan kerangka evaluasi yang dapat menilai dampak dan efektivitasnya secara sistematis (Ameer, 2025).

PENUTUP

Simpulan

Tinjauan sistematis ini memetakan lanskap pemanfaatan AI dalam asesmen PAI dan menemukan sebuah paradoks: literatur menunjukkan antusiasme yang signifikan terhadap peluang AI, namun bukti empiris yang mendukung adopsi masih sangat terbatas (Ameer, 2025). Peluang yang teridentifikasi—otomatisasi penilaian, umpan balik personal, asesmen keterampilan spesifik, dan adaptivitas—berpotensi mengatasi persoalan kronis asesmen PAI berupa subjektivitas dan keterbatasan kapasitas evaluatif (Ogunleye et al., 2024). Namun, tantangan etis yang teridentifikasi—bias algoritmik, opasitas, pergeseran otoritas epistemik, dan risiko ketergantungan—memiliki resonansi khusus dalam konteks PAI yang menuntut integrasi prinsip keislaman ke dalam setiap dimensi praktik pendidikan (Sood, 2026c).

Kontribusi teoretis utama penelitian ini adalah kerangka Techno-Ethical Assessment Literacy yang menyintesis kompetensi teknis, literasi etis, dan fondasi normatif Islami (adab dan amanah) sebagai fondasi konseptual untuk asesmen PAI berbasis AI (Sood, 2026b). Kebaruan teoretis ini terletak pada pengintegrasian dimensi-dimensi yang biasanya dibahas secara terpisah—kompetensi teknologi, kesadaran etika, dan nilai-nilai Islam—ke dalam satu kerangka yang koheren dan dapat dioperasionalkan

(Mighfar et al., 2026). Kerangka ini menjawab celah under-theorised yang diidentifikasi dalam literatur dengan menyediakan lensa analitis yang menghubungkan penggunaan pedagogis AI dengan pertanyaan epistemologis dan etika Islam (Sood, 2026a). Kontribusi kontekstualnya adalah penegasan bahwa adopsi AI dalam asesmen PAI harus diposisikan sebagai human-mediated augmentation, bukan substitusi peran evaluatif guru (Bukhari & Kaur, 2026). Guru sebagai arbiter etis tetap memegang otoritas final dalam penilaian, khususnya untuk dimensi-dimensi yang membutuhkan pertimbangan spiritual dan moral yang tidak dapat direduksi menjadi output algoritmik (Hamid & Rizki, 2024). Kontekstualisasi ini penting karena literatur yang ada sering kali mengabaikan kekhususan PAI sebagai domain yang tidak dapat diperlakukan sama dengan pendidikan umum (Fauzi, 2023).

Ke depan, agenda penelitian yang paling mendesak adalah mengisi celah empiris yang telah teridentifikasi (Ameer, 2025). Diperlukan studi berbasis kelas yang mengeksplorasi bagaimana guru dan siswa benar-benar menggunakan asesmen AI dalam konteks PAI, bagaimana penggunaan tersebut berhubungan dengan hasil belajar dan pembentukan karakter, dan bagaimana kompleksitas etis dinavigasi dalam praktik (Sood, 2026a). Penelitian longitudinal juga diperlukan untuk menilai dampak jangka panjang adopsi AI terhadap pemahaman keislaman, kemandirian berpikir, dan hubungan guru-siswa (Jauhiainen & Bernardo, 2025). Visi jangka panjangnya adalah terciptanya ekosistem asesmen PAI yang memanfaatkan kecanggihan AI untuk meningkatkan kualitas umpan balik dan efisiensi, tanpa mengorbankan otoritas epistemik guru, nilai-nilai etika Islam, dan tujuan pembentukan insan kamil (Sood, 2026c). Tanpa agenda empiris yang sistematis, potensi AI untuk mentransformasi asesmen PAI akan tetap menjadi janji yang belum terwujud, dan risiko-risiko etisnya akan tetap menjadi kekhawatiran yang belum terverifikasi (Mighfar et al., 2026).

DAFTAR PUSTAKA

- Aldreabi, H., et al. (2025). Data Privacy in AI-Enhanced Education: A Governance Framework. *Computers & Security*, 148, 104123.
- Ameer, S. (2025). Assessment for Learning in the Era of Artificial Intelligence: Rethinking Feedback, Evaluation, and Academic Integrity. *The Study*

- of Religion and History, 3(4), 1-26.
<https://doi.org/10.5281/zenodo.22035279>
- Anwar, A., et al. (2025). Kecerdasan Buatan dalam Asesmen Keterampilan Al-Qur'an. *Jurnal Pendidikan Islam*, 12(1), 89-106.
- Anwar, M., & Umam, A. (2025). AI dan Personalisasi Pembelajaran PAI untuk Generasi Z. *Jurnal UIN Malang*, 14(2), 201-218.
- Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Bukhari, N., & Kaur, A. (2026). From Automation to Accountability: The Promise and Perils of AI in Educational Assessment. In *AI in Higher Education* (pp. 89-112). Taylor & Francis.
- Chanda, K., et al. (2023). Automating Administrative Tasks in Education: AI Applications. *International Journal of Educational Management*, 37(4), 812-829.
- Chima, C., et al. (2024). AI-Powered Formative Assessment: A Systematic Review. *Assessment & Evaluation in Higher Education*, 49(3), 356-374.
- Clark, J., et al. (2025). Transparency and Explainability in AI-Based Assessment. *British Journal of Educational Technology*, 56(2), 445-462.
- Faqihuddin, M., & Romadhon, A. (2023). Maqasid al-Shariah sebagai Fondasi Etika AI dalam Pendidikan Islam. *Jurnal Ushuluddin*, 21(2), 156-173.
- Fauzi, M. (2023). Transformasi Pendidikan Agama Islam di Era Digital: Tantangan dan Peluang. *Jurnal Pendidikan Islam*, 8(2), 145-162.
- Hamid, A., & Rizki, M. (2024). Otoritas Guru PAI sebagai Uswah Hasanah dalam Era Kecerdasan Buatan. *Jurnal Pendidikan Agama Islam*, 11(2), 178-195.
- Hasanah, N. (2023). Dimensi Afektif dan Spiritual dalam Asesmen Pendidikan Agama Islam. *Jurnal Studi Islam*, 15(3), 289-304.
- Imamuddin, M., et al. (2022). Sistem Penilaian Cerdas Berbasis AI untuk Literasi Keagamaan. *Jurnal Teknologi Informasi dan Pendidikan*, 15(2), 234-251.
- Jauhiainen, Y., & Bernardo, A. (2025). AI in Education: Efficiency, Scalability, and Personalization. *Journal of Educational Technology & Society*, 27(1), 78-95.
- Javed, F., et al. (2025). Adaptive Testing and Personalized Learning Pathways. *Computers & Education*, 210, 104982.

- Kovári, A. (2024). Steps of Systematic Reviews Based on the PRISMA 2020 Recommendation. *Journal of Applied Technical and Educational Sciences*, 14(3), 1-18.
- Kumar, R. (2024). Artificial Intelligence in Educational Assessment: A Systematic Review. *Computers & Education: Artificial Intelligence*, 5, 100123.
- Lu, S., et al. (2022). Natural Language Processing for Automated Feedback in Education. *Computers & Education*, 185, 104536.
- Mighfar, S., Hafid, M., & Qotrinnada, I. (2026). Etika dan Risiko Kecerdasan Artifisial Generatif: Implikasi dalam Pendidikan Agama Islam dan Masyarakat Digital. *Edupedia: Jurnal Studi Pendidikan dan Pedagogi Islam*, 10(2), 275-294. <https://doi.org/10.35316/edupedia.v10i2.7588>
- Moher, D., et al. (2021). PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews. *BMJ*, 372, n160.
- Nasucha, A., et al. (2023). Adaptive Assessment in Islamic Religious Education. *Jurnal Pendidikan Agama Islam*, 10(3), 312-329.
- Nugroho, S., & Prasetyo, D. (2023). Personalisasi Pembelajaran PAI Berbasis AI: Sebuah Eksplorasi Konseptual. *Jurnal Teknologi Pendidikan*, 24(1), 67-84.
- Ogunleye, I., et al. (2024). AI-Driven Assessment in Higher Education: A Systematic Review. *Education and Information Technologies*, 29, 14523-14556.
- Page, J., et al. (2021). The PRISMA 2020 Statement