

IMPLEMENTASI ARTIFICIAL INTELLIGENCE (AI) DALAM SISTEM EVALUASI PEMBELAJARAN: PELUANG DAN TANTANGAN DI ERA DIGITAL

Maulidina Akmal¹, Ridho Riyadi²

^{1,2} PAI FTIK Universitas Islam Negeri K.H Abdurrahman Wahid Pekalongan

1maulidina.akmal24021@mhs.uingusdur.ac.id, 2ridho.riyadi@uingusdur.ac.id

ABSTRACT

The emergence of Artificial Intelligence [AI] has transformed learning assessment practices, particularly in response to the need for authentic and differentiated evaluation in the digital era. While AI promises greater efficiency, its adoption in schools also raises several ethical and technical concerns. This article argues that AI will be most effective when positioned as a teacher's partner through a Human-in-the-Loop approach, rather than as a fully automated replacement. To examine this issue, the study employs a literature review method with thematic analysis of 30 articles from SINTA and Scopus-indexed journals published between 2020 and 2025. The findings indicate that AI offers three significant opportunities: accelerating the assessment process through Automated Essay Scoring, adjusting item difficulty adaptively, and identifying at-risk students using predictive analytics. However, implementation faces three major challenges, namely cultural bias in algorithms, the risk of student data breaches, and the erosion of teachers' sensitivity in assessment. Therefore, this article recommends clear regulations, strengthening teachers' AI literacy, and a collaborative model between AI and teachers to ensure that assessment remains both accurate and humane.

Keywords: *Artificial Intelligence; Learning Assessment; AI Ethics; Personalized Learning*

ABSTRAK

Kehadiran Artificial Intelligence [AI] telah mengubah cara evaluasi pembelajaran dilakukan, khususnya untuk merespons kebutuhan penilaian autentik dan terdiferensiasi di era digital. Meskipun menjanjikan efisiensi, penggunaan AI di sekolah juga memunculkan sejumlah persoalan etis dan teknis. Artikel ini mengajukan argumen bahwa AI akan optimal jika diposisikan sebagai mitra guru melalui pendekatan Human-in-the-Loop, bukan sebagai pengganti yang bekerja sepenuhnya otomatis. Untuk mengkaji hal tersebut, penelitian ini menggunakan metode studi literatur dengan analisis tematik pada 30 artikel dari jurnal SINTA dan Scopus periode 2020-2025. Hasil kajian menunjukkan bahwa AI menawarkan tiga peluang signifikan: mempercepat proses penilaian melalui Automated Essay Scoring, menyesuaikan tingkat kesulitan soal secara adaptif, serta memetakan siswa yang berisiko melalui analitik prediktif. Namun demikian, implementasinya menghadapi tiga tantangan utama berupa bias budaya dalam algoritma, risiko kebocoran data siswa, dan melemahnya kepekaan guru dalam menilai. Dengan demikian, artikel ini merekomendasikan adanya regulasi yang jelas, penguatan literasi AI bagi guru, serta model kolaborasi antara AI dan guru agar penilaian tetap akurat dan manusiawi.

Kata Kunci: *Artificial Intelligence; Evaluasi Pembelajaran; Etika AI; Pembelajaran Personal*

A. Pendahuluan

Transformasi digital telah menjadi pencerahan dalam dunia pendidikan Indonesia, khususnya setelah pandemi COVID-19. Data APJII (2024) mencatat bahwa tingkat penggunaan internet pada sektor pendidikan telah mencapai 89,02%, dan mayoritas pendidik kini telah memanfaatkan platform digital dalam proses pembelajaran. Bersamaan dengan itu, implementasi Kurikulum Merdeka menuntut dilaksanakannya evaluasi yang bersifat autentik, berkelanjutan, dan terdiferensiasi sesuai kebutuhan peserta didik (Kemendikbudristek, 2022). Akan tetapi, di tengah tuntutan tersebut, guru masih dihadapkan pada beban administratif yang tinggi. Laporan Kemendikbudristek (2022) menegaskan bahwa pekerjaan non-pembelajaran masih menyita porsi waktu guru yang signifikan. Kondisi ini membuka ruang bagi hadirnya *Artificial Intelligence* [AI], terutama melalui teknologi seperti *Automated Essay Scoring* [AES] untuk mempercepat proses penilaian (Arifin & Setyosari, 2021). Meskipun menawarkan efisiensi, Munir dan Purnawarman (2023) mengingatkan bahwa kesiapan guru di Indonesia

dalam mengadopsi AI masih menjadi persoalan. Jika tidak diimbangi dengan pemahaman etis, penggunaan AI justru dapat menimbulkan dilema baru dalam praktik evaluasi (Setiawan & Wibowo, 2022).

Kajian terdahulu menunjukkan bahwa diskursus AI dalam pendidikan di Indonesia bergerak pada dua kutub. Kutub pertama bersifat optimistis dan menekankan efisiensi, sebagaimana tergambar dalam penelitian Arifin dan Setyosari (2021) mengenai pengembangan AES untuk esai Bahasa Indonesia. Kutub kedua bersifat kritis dan menyoroti aspek etika, seperti yang dikemukakan Setiawan dan Wibowo (2022) terkait risiko privasi dan bias algoritma. Namun demikian, kedua kutub tersebut cenderung memosisikan AI dan guru dalam relasi yang dikotomis, yakni AI sebagai alat otomatisasi atau sebagai ancaman terhadap peran pendidik. Padahal, Yaniawati dan Kariadinata (2023) telah mengindikasikan pentingnya redefinisi peran guru di era AI, meskipun belum merumuskan model implementasi yang konkret. Kekosongan inilah yang menjadi research gap sekaligus kebaruan dari

penelitian ini. Berbeda dengan kajian sebelumnya, tulisan ini mengajukan model “*Human-in-the-Loop AI Evaluation*”, yaitu kerangka kolaboratif yang menempatkan guru sebagai penentu akhir dalam sistem evaluasi berbasis AI yang beretika dan kontekstual.

Bertolak dari kesenjangan tersebut, penelitian ini berupaya mengkaji implementasi AI dalam sistem evaluasi pembelajaran secara kritis dan komprehensif. Secara rinci, penelitian ini memiliki tiga tujuan. Pertama, mengidentifikasi berbagai peluang yang ditawarkan AI dalam evaluasi pembelajaran. Kedua, menganalisis tantangan etis dan teknis yang menyertai implementasi AI. Ketiga, merumuskan kerangka “*Human-in-the-Loop AI Evaluation*” sebagai model yang selaras dengan arah kebijakan pendidikan nasional (Kemendikbudristek, 2022). Dengan pencapaian ketiga tujuan tersebut, diharapkan artikel ini dapat memberikan kontribusi konseptual sekaligus praktis bagi pengembangan evaluasi pembelajaran yang lebih efisien, adil, dan manusiawi di era digital.

Penelitian ini berpijak pada argumentasi bahwa AI tidak

seharusnya diposisikan sebagai pengganti guru dalam kegiatan evaluasi. Sebaliknya, AI akan memberikan manfaat optimal apabila difungsikan sebagai mitra kognitif guru melalui skema *Human-in-the-Loop* (Yaniawati & Kariadinata, 2023). Premis ini berangkat dari pandangan Munir dan Purnawarman (2023) bahwa teknologi pendidikan harus tetap berorientasi pada manusia. Oleh karena itu, hipotesis yang diajukan dalam penelitian ini adalah bahwa tanpa adanya regulasi yang memadai, penguatan literasi AI bagi guru, dan desain sistem yang kolaboratif, implementasi AI dalam evaluasi berisiko menimbulkan bias, mengurangi kepekaan pedagogis guru, dan mengikis nilai kemanusiaan dalam pendidikan.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan kualitatif dengan metode studi kepustakaan (*library research*). Pendekatan ini dipilih karena penelitian bertujuan untuk mengkaji secara mendalam konsep, perkembangan, peluang, tantangan, serta implikasi penerapan kecerdasan buatan (*Artificial Intelligence/AI*) dalam evaluasi

pendidikan berdasarkan berbagai sumber ilmiah yang relevan. Data penelitian diperoleh melalui penelusuran literatur berupa artikel jurnal terindeks, buku ilmiah, prosiding, dan dokumen resmi yang membahas evaluasi pendidikan, transformasi digital, serta implementasi AI dalam pembelajaran. Metode studi kepustakaan dipilih karena mampu memberikan landasan teoritis yang komprehensif serta menghasilkan sintesis pengetahuan yang relevan terhadap isu-isu kontemporer dalam dunia pendidikan (Creswell & Creswell, 2018).

Analisis data dilakukan dengan menggunakan teknik analisis isi (*content analysis*), yaitu mengidentifikasi, mengelompokkan, membandingkan, dan menginterpretasikan informasi dari berbagai sumber untuk menemukan pola, kesamaan, maupun perbedaan pandangan terkait pemanfaatan AI dalam evaluasi pendidikan. Selanjutnya, hasil analisis disusun secara deskriptif-analitis sehingga mampu memberikan gambaran yang sistematis mengenai peluang, tantangan, serta implikasi penggunaan AI terhadap peningkatan kualitas pembelajaran. Teknik ini

dipilih karena memungkinkan peneliti memperoleh pemahaman yang lebih mendalam dan objektif terhadap fenomena yang dikaji berdasarkan bukti-bukti ilmiah yang telah dipublikasikan (Krippendorff, 2019).

C. Hasil Penelitian dan Pembahasan

Pada bagian ini ditemukan beberapa data dari studi literatur terhadap beberapa jurnal, artikel ataupun penelitian terdahulu yaitu peluang dan tantangan implementasi AI dalam evaluasi pembelajaran. Pertama menunjukkan bahwa AI memiliki potensi besar dalam meningkatkan efisiensi proses evaluasi. Berdasarkan kajian Jumadi et al. (2023), penggunaan AI dalam penilaian sains dapat mengurangi waktu yang dibutuhkan guru untuk mengoreksi hingga 60%. Hal ini terjadi karena sistem dapat secara otomatis memeriksa jawaban objektif dan memberikan rekapitulasi nilai secara langsung.

Selanjutnya, yang kedua dan ketiga berkaitan dengan personalisasi serta analisis data. Wahyudi dan Utami (2022) menemukan bahwa sistem pembelajaran adaptif berbasis AI mampu menyesuaikan tingkat kesulitan soal sesuai dengan

kemampuan siswa setelah melakukan evaluasi awal. Sementara itu, Sutarto et al. (2021) menyatakan bahwa sebagian besar data analisis membantu guru memetakan kelemahan kelas secara komprehensif dalam waktu singkat. Dari hal tersebut, AI tidak hanya cepat, tetapi juga dapat memberikan gambaran yang lebih detail. Di sisi lain, ditemukan pula tantangan signifikan. Putri dan Rahman (2023) mengidentifikasi adanya bias algoritmik dalam sistem penilaian AI. Penelitian mereka menunjukkan bahwa siswa dengan gaya bahasa tertentu cenderung mendapatkan skor lebih rendah karena data pelatihan didominasi oleh satu kelompok. Selain itu, Kurniawan dan Sari (2022) menyoroti persoalan etika dan privasi. Mereka mencatat bahwa 70% platform AI pendidikan belum memiliki kebijakan perlindungan data siswa yang jelas di Indonesia.

Tantangan keempat dan kelima berkaitan dengan sumber daya manusia. Maharani dan Fadhillah (2023) menemukan kekhawatiran guru bahwa penggunaan AI secara berlebihan dapat mengurangi aspek humanis dalam evaluasi, seperti penilaian sikap dan kreativitas.

Senada dengan itu, Syah dan Saputra (2023) menyatakan bahwa literasi digital guru masih menjadi kendala utama. Survei mereka pada 120 guru menunjukkan hanya 38% yang merasa mampu mengoperasikan dan mengevaluasi hasil AI secara kritis. Dua kajian terdahulu yang terakhir melengkapi gambaran di lapangan. Abidin (2022) menyimpulkan bahwa tantangan terbesar evaluasi di era digital adalah perubahan paradigma dari penilaian produk ke penilaian proses. Fitria (2023) menambahkan bahwa persepsi pendidik terhadap AI dalam asesmen masih terbagi. Sebagian melihatnya sebagai alat bantu, sebagian lagi meragukan validitasnya. Terakhir, Hidayat (2021) menekankan bahwa kebijakan sekolah terkait teknologi evaluasi belum merata, sehingga implementasi di lapangan menjadi tidak konsisten.

Mengenai efisiensi AI, menurut hasil penelitian Jumadi et al. (2023) yang membuktikan AI dapat memangkas waktu koreksi secara signifikan. Akan tetapi, penelitian ini mengkritisi jika efisiensi dijadikan satu-satunya tujuan. Berbeda halnya dengan pandangan yang ingin

menggantikan guru sepenuhnya, tulisan ini berargumen bahwa waktu yang dihemat justru harus digunakan guru untuk validasi dan interaksi yang lebih berkualitas dengan siswa. Mengenai personalisasi dan analisis data mendukung atau memperkuat studi Wahyudi dan Utami (2022) serta Sutarto et al. (2021). AI memang unggul dalam memproses data individual dan kelas. Namun, temuan Syah dan Saputra (2023) menjadi catatan penting. Karena hanya 38% guru yang siap, maka potensi AI tidak akan maksimal. Dengan kata lain, teknologi tanpa kapasitas guru sama saja dengan alat yang tidak terpakai. Isu bias dan etika tidak jauh berbeda dengan temuan Putri dan Rahman (2023) serta Kurniawan dan Sari (2022). Keduanya sama-sama memperingatkan risiko jika AI diterapkan tanpa regulasi. Pembahasan ini mengkritisi praktik implementasi yang terburu-buru di banyak sekolah. Seharusnya, audit algoritma dan kebijakan privasi menjadi syarat utama sebelum AI digunakan untuk evaluasi sumatif.

Kekhawatiran tentang pergeseran peran guru senada dengan Maharani dan Fadhillah (2023). Begitu juga dengan persepsi

guru yang terbelah seperti yang dilaporkan Fitria (2023). Dari sini muncul rekomendasi utama penelitian ini, yaitu model “Human-in-the-Loop”. Dalam model ini, AI berfungsi sebagai asisten analitik, sementara guru memegang otoritas penuh dalam pengambilan keputusan evaluasi. Dengan begitu, aspek humanis yang ditekankan Abidin (2022) tetap terjaga. Yang terakhir, menurut Hidayat (2021) mengenai ketimpangan kebijakan memperkuat argumen bahwa implementasi AI tidak bisa diserahkan sepenuhnya ke sekolah. Diperlukan kebijakan dari pemerintah atau yayasan yang jelas terkait etika, pelatihan, dan standar penggunaan AI dalam evaluasi. Tanpa payung kebijakan, kesenjangan digital antar sekolah akan semakin lebar.

D. Kesimpulan

Pemanfaatan Artificial Intelligence (AI) dalam evaluasi pembelajaran memiliki potensi yang besar untuk meningkatkan kualitas proses penilaian. AI mampu meningkatkan efisiensi kerja guru melalui otomatisasi penilaian, mendukung pembelajaran yang lebih personal, serta membantu

menganalisis hasil belajar peserta didik secara lebih cepat dan komprehensif. Meskipun demikian, implementasi AI juga menghadapi berbagai tantangan, seperti potensi bias algoritma, persoalan etika dan perlindungan data, rendahnya literasi digital guru, perbedaan persepsi terhadap penggunaan AI, serta belum meratanya kebijakan sekolah dalam mendukung pemanfaatan teknologi evaluasi.

Dari pembahasan menunjukkan bahwa AI tidak seharusnya dipandang sebagai pengganti peran guru, melainkan sebagai teknologi pendukung yang dapat membantu proses pengambilan keputusan dalam evaluasi pembelajaran. Oleh karena itu, penerapan model Human-in-the-Loop menjadi pendekatan yang paling relevan karena mengombinasikan kemampuan AI dalam mengolah data dengan pertimbangan pedagogis, etis, dan profesional yang dimiliki guru. Selain itu, diperlukan dukungan kebijakan yang lebih jelas, peningkatan kompetensi literasi digital guru, serta penyusunan standar etika penggunaan AI agar implementasinya dapat berlangsung secara optimal, adil, dan tetap

berorientasi pada nilai-nilai pendidikan.

DAFTAR PUSTAKA

- APJII. (2024). Survei internet APJII 2024. Asosiasi Penyelenggara Jasa Internet Indonesia. <https://apjii.or.id>
- Abidin, Z. (2022). Tantangan evaluasi pembelajaran di era digital. *Jurnal Evaluasi Pendidikan*, 9(1), 55-67. <https://sinta.kemdikbud.go.id/journals/detail?id=3387>
- Arifin, Z., & Setyosari, P. (2021). Pengembangan model Automated Essay Scoring untuk penilaian esai Bahasa Indonesia. *Jurnal Teknologi Pendidikan*, 9(1), 78-91. <https://sinta.kemdikbud.go.id>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Fitria, T. (2023). Persepsi pendidik terhadap kecerdasan buatan dalam asesmen. *Jurnal Kependidikan*, 12(2), 99-111. <https://sinta.kemdikbud.go.id/journals/detail?id=6501>
- Hidayat, R. (2021). Kebijakan teknologi pendidikan dan kesiapan sekolah. *Jurnal Kebijakan Pendidikan*, 4(3), 145-159. <https://sinta.kemdikbud.go.id/journals/detail?id=2804>
- Jumadi, J., Yulianto, A., & Prasetyo, Z. (2023). Efektivitas penggunaan AI dalam

- penilaian pembelajaran sains. *Jurnal Inovasi Pendidikan IPA*, 9(1), 45-58. <https://sinta.kemdikbud.go.id/journals/detail?id=2089>
- Kemendikbudristek. (2022). Laporan hasil survei digitalisasi pendidikan. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. <https://kemdikbud.go.id>
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.
- Kurniawan, D., & Sari, L. (2022). Isu etika dan privasi dalam penggunaan kecerdasan buatan di pendidikan. *Jurnal Etika dan Hukum Pendidikan*, 4(2), 112-125. <https://sinta.kemdikbud.go.id/journals/detail?id=5123>
- Maharani, S., & Fadhillah, N. (2023). Dampak kecerdasan buatan terhadap profesionalitas guru. *Jurnal Profesi Guru Indonesia*, 7(3), 201-215. <https://sinta.kemdikbud.go.id/journals/detail?id=4456>
- Munir, M., & Purnawarman, P. (2023). Pemanfaatan Artificial Intelligence dalam pembelajaran di Indonesia: Peluang dan tantangan. *Jurnal Pendidikan dan Kebudayaan*, 8(2), 123-138. <https://sinta.kemdikbud.go.id>
- Putri, A., & Rahman, M. (2023). Bias algoritmik dalam sistem penilaian berbasis AI. *Jurnal Teknologi Informasi dan Pendidikan*, 16(1), 33-47. <https://sinta.kemdikbud.go.id/journals/detail?id=6012>
- Setiawan, R., & Wibowo, A. (2022). Etika kecerdasan buatan dalam sistem pendidikan Indonesia. *Jurnal Filsafat Indonesia*, 5(1), 45-56. <https://sinta.kemdikbud.go.id>
- Sutarto, S., Sari, D., & Fathurrohman, M. (2021). Pemanfaatan big data analytics untuk evaluasi hasil belajar. *Jurnal Pendidikan Matematika dan Sains*, 9(2), 88-101. <https://sinta.kemdikbud.go.id/journals/detail?id=3011>
- Syah, R., & Saputra, E. (2023). Literasi digital guru dalam implementasi teknologi evaluasi. *Jurnal Pendidikan Guru*, 11(4), 155-168. <https://sinta.kemdikbud.go.id/journals/detail?id=7221>
- UNESCO. (2021). *AI and education: Guidance for policy-makers*. UNESCO Publishing.
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.
- Wahyudi, A., & Utami, P. (2022). Pembelajaran adaptif berbasis AI untuk personalisasi pendidikan. *Jurnal Inovasi Teknologi Pendidikan*, 9(3), 178-190. <https://sinta.kemdikbud.go.id/journals/detail?id=4190>
- Yaniawati, P., & Kariadinata, R. (2023). Peran guru di era Artificial Intelligence pada pembelajaran abad 21. *Jurnal Inovasi Pendidikan*, 14(2), 89-102. <https://sinta.kemdikbud.go.id>