

Studi Pedagogis dan Etika Teknologi pada Penggunaan Sistem Penilaian Esai Berbasis AI di Pendidikan Tinggi

Rizki Adha¹, Lusianto², Dodi Syaripudin³, Agus Nursikuwagus⁴, Usep Mohamad Ishaq⁵, Andrias Darmayadi⁶

^{1, 2, 3, 4, 5, 6} Doktor Sistem Informasi, Universitas Komputer Indonesia, Indonesia

Email: ¹rizkiadha67@gmail.com, ²maslusianto@gmail.com, ³dodisyaripudin@gmail.com,

⁴agusnursikuwagus@email.unikom.ac.id, ⁵usep.mohamad.ishaq@email.unikom.ac.id,

⁶andrias.darmayadi@email.unikom.ac.id

Abstrak - Penggunaan kecerdasan artifisial (AI) dalam penilaian esai semakin berkembang di pendidikan tinggi karena kemampuannya meningkatkan efisiensi dan konsistensi evaluasi. Namun, penerapannya menimbulkan tantangan pedagogis dan etika terkait keadilan, transparansi, serta legitimasi keputusan akademik. Penelitian ini bertujuan menganalisis implikasi pedagogis dan etika penggunaan sistem penilaian esai berbasis AI dengan mengintegrasikan persepsi mahasiswa dan dosen. Penelitian menggunakan pendekatan mixed-methods exploratory melalui survei kuantitatif terhadap 36 mahasiswa dan 24 dosen, dilengkapi analisis kualitatif. Hasil menunjukkan mahasiswa memiliki persepsi positif terhadap kegunaan, kemudahan, dan nilai pedagogis, namun tetap berhati-hati terhadap aspek kepercayaan dan transparansi. Sementara itu, dosen menunjukkan tingkat penerimaan sangat tinggi selama sistem diterapkan dalam kerangka terkontrol. Integrasi temuan menunjukkan bahwa pendekatan Human-in-the-Loop merupakan faktor kunci menjaga keseimbangan antara efisiensi teknologi, nilai pedagogis, dan legitimasi akademik. Penelitian ini berkontribusi dengan menegaskan bahwa sistem penilaian esai berbasis AI paling tepat diterapkan dalam paradigma AI-assisted assessment, di mana dosen tetap memegang kendali utama dalam pengambilan keputusan akademik, serta memberikan rekomendasi praktis bagi institusi dalam merancang kebijakan penilaian berbasis AI yang bertanggung jawab.

Kata Kunci - kecerdasan artifisial, penilaian esai otomatis, pedagogi, etika teknologi, human-in-the-loop.

Abstract - The use of artificial intelligence (AI) in essay grading is growing in higher education due to its potential to improve evaluation efficiency and consistency. However, its implementation poses pedagogical and ethical challenges related to fairness, transparency, and the legitimacy of academic decisions. This study aims to analyze the pedagogical and ethical implications of using an AI-based essay grading system by integrating student and lecturer perceptions. The study employed a mixed-methods exploratory approach through a quantitative survey of 36 students and 24 lecturers, complemented by qualitative analysis. The results indicate that students have positive perceptions of usability, ease of use, and pedagogical value, but remain cautious about trustworthiness and transparency. Meanwhile, lecturers demonstrated a very high level of acceptance as long as the system is implemented within a controlled framework. The integration of the findings indicates that a Human-in-the-Loop approach is a key factor in maintaining a balance between technological efficiency, pedagogical value, and academic legitimacy. This study contributes by confirming that an AI-based essay grading system

is most appropriately implemented within an AI-assisted assessment paradigm, where lecturers retain primary control over academic decision-making, and provides practical recommendations for institutions in designing responsible AI-based grading policies.

Keywords - artificial intelligence, automated essay grading, pedagogy, technology ethics, human-in-the-loop.

I. PENDAHULUAN

Penilaian pembelajaran merupakan komponen fundamental dalam pendidikan tinggi karena berfungsi tidak hanya untuk mengukur capaian pembelajaran mahasiswa, tetapi juga untuk membentuk kualitas dan arah proses belajar itu sendiri [1]. Dalam konteks penilaian esai, dosen dituntut untuk menilai kemampuan berpikir kritis, pemahaman konseptual, serta kualitas argumentasi mahasiswa secara komprehensif [2]. Praktik penilaian esai di perguruan tinggi kerap menghadapi berbagai kendala, seperti keterbatasan waktu dosen, tingginya beban kerja akademik, serta tantangan dalam menjaga konsistensi dan objektivitas penilaian [3].

Perkembangan teknologi kecerdasan artifisial (Artificial Intelligence/AI) mendorong munculnya berbagai inovasi dalam bidang pendidikan, termasuk pada proses penilaian pembelajaran [4], [5]. Salah satu bentuk penerapan AI yang semakin banyak digunakan adalah Automated Essay Scoring (AES)[6], yaitu sistem penilaian esai yang memanfaatkan kemampuan komputasi untuk menghasilkan skor dan umpan balik secara otomatis [7]. Sistem ini menawarkan efisiensi dan kecepatan dalam proses evaluasi, sehingga dipandang mampu membantu dosen dalam mengelola penilaian akademik secara lebih efektif.

Tinjauan sistematis terhadap 49 studi menunjukkan bahwa LLM semakin banyak diterapkan untuk penilaian otomatis, khususnya pada penilaian esai dan jawaban singkat, serta memiliki potensi besar untuk mentransformasi asesmen Pendidikan[8]. Model LLM ini memiliki kemampuan pemrosesan bahasa alami yang lebih canggih sehingga mampu memahami konteks teks, struktur argumentasi, serta hubungan antar gagasan secara lebih mendalam. Perkembangan Large Language Models (LLMs) membuka peluang integrasi dengan platform pembelajaran digital (Learning Management System/LMS) untuk

mendukung penilaian otomatis dan pemberian umpan balik kepada mahasiswa [9].

Meskipun menawarkan berbagai manfaat, penerapan sistem penilaian esai berbasis AI juga memunculkan sejumlah tantangan pedagogis dan etika yang perlu mendapat perhatian dalam implementasinya di lingkungan Pendidikan [10]. Dari sisi pedagogis, muncul pertanyaan mengenai sejauh mana AI mampu menilai aspek-aspek kompleks seperti kreativitas, kedalaman analisis, dan konteks pemikiran mahasiswa. Meskipun umpan balik berbasis AI menawarkan keunggulan dalam hal kecepatan, konsistensi, dan kemudahan akses, efektivitasnya dalam mendorong pembelajaran mendalam masih lebih rendah dibandingkan umpan balik dosen karena keterbatasannya dalam memahami konteks, pengembangan isi, dan kebutuhan individual mahasiswa [11]. Kondisi ini menimbulkan kekhawatiran bahwa penggunaan AI secara tidak tepat dapat mereduksi penilaian menjadi sekadar proses mekanis pemberian skor.

Dari sisi etika, penggunaan AI dalam penilaian akademik memunculkan berbagai isu, seperti keadilan (*fairness*), transparansi (*transparency*), dan akuntabilitas (*accountability*), yang perlu diperhatikan agar proses asesmen tetap dapat dipertanggungjawabkan dan dipercaya [12]. Sistem AI berpotensi menghasilkan bias algoritmik yang dapat memengaruhi kelompok mahasiswa atau gaya penulisan tertentu. kurangnya transparansi mengenai dasar pemberian skor dapat menurunkan tingkat kepercayaan mahasiswa terhadap hasil penilaian. Persoalan mengenai tanggung jawab atas kesalahan penilaian AI juga menjadi isu krusial dalam konteks akademik.

Untuk merespons tantangan tersebut, berbagai pendekatan dikembangkan dalam implementasi AI di pendidikan tinggi. Penelitian sebelumnya menunjukkan bahwa integrasi Gemini AI dengan pendekatan Retrieval-Augmented Generation (RAG) dan mekanisme *Human-in-the-Loop* mampu mendukung proses penilaian esai yang lebih efisien, akuntabel, serta tetap mempertahankan peran dosen dalam menentukan nilai akhir [13]. Pendekatan ini sejalan dengan prinsip Responsible AI yang menekankan keterlibatan manusia, transparansi, dan akuntabilitas.

Meskipun kajian mengenai AES dan penerapan AI dalam pendidikan terus berkembang, sebagian besar penelitian masih berfokus pada aspek teknis dan akurasi sistem. Kajian empiris yang menelaah persepsi mahasiswa dan dosen terhadap implikasi pedagogis dan etika penggunaan sistem penilaian esai berbasis AI, khususnya yang terintegrasi dengan LMS dan menerapkan pendekatan RAG serta HIL, masih relatif terbatas, terutama dalam konteks pendidikan tinggi di Indonesia. Persepsi pengguna merupakan faktor kunci yang memengaruhi penerimaan, adopsi, dan keberlanjutan penggunaan teknologi dalam lingkungan pendidikan, karena menentukan kemauan pengguna untuk menerima dan memanfaatkan teknologi dalam proses pembelajaran [14].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menganalisis persepsi mahasiswa dan dosen terhadap penggunaan sistem penilaian esai berbasis AI yang

terintegrasi dengan LMS Moodle serta didukung oleh pendekatan RAG dan Human-in-the-Loop. Fokus analisis mencakup aspek kegunaan, kemudahan penggunaan, kepercayaan, transparansi, nilai pedagogis, serta penerimaan etis terhadap sistem tersebut. Penelitian ini diharapkan dapat memberikan kontribusi teoritis dalam memperluas kajian penilaian berbasis AI dari perspektif pedagogis dan etika, sekaligus memberikan rekomendasi praktis bagi perguruan tinggi dalam menerapkan sistem penilaian AI yang berpusat pada manusia dan bertanggung jawab.

II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan mixed-methods eksploratif untuk mengkaji implikasi pedagogis dan etika penggunaan sistem penilaian esai berbasis kecerdasan artifisial di pendidikan tinggi [15]. Pendekatan ini dipilih karena fenomena yang diteliti tidak hanya berkaitan dengan efektivitas sistem secara teknis, tetapi juga mencakup persepsi, pengalaman, dan pertimbangan akademik para pengguna dalam konteks pembelajaran, khususnya mahasiswa dan dosen.

Penggunaan metode campuran memungkinkan peneliti memperoleh pemahaman yang lebih komprehensif mengenai penerimaan dan tantangan implementasi sistem penilaian esai berbasis AI. Data kuantitatif digunakan untuk memetakan kecenderungan persepsi pengguna secara sistematis, sedangkan data kualitatif digunakan untuk menggali pengalaman dan refleksi pedagogis serta etika yang tidak sepenuhnya dapat dijelaskan melalui data numerik. Dengan demikian, pendekatan ini dinilai tepat untuk menangkap kompleksitas penggunaan AI dalam penilaian akademik.

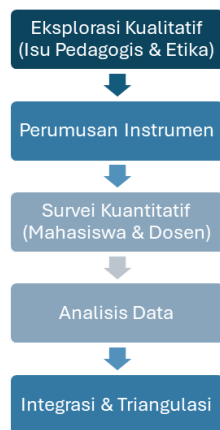
A. Desain Penelitian

Desain penelitian yang digunakan adalah exploratory mixed-methods, dimana pengumpulan dan analisis data dilakukan melalui kombinasi pendekatan kualitatif dan kuantitatif secara bertahap [15]. Tahap awal penelitian bersifat eksploratif dan bertujuan untuk mengidentifikasi isu-isu utama terkait penggunaan sistem penilaian esai berbasis AI, khususnya yang berkaitan dengan aspek pedagogis dan etika, seperti kepercayaan terhadap sistem, transparansi penilaian, serta peran dosen dalam pengambilan keputusan akademik.

Berdasarkan hasil eksplorasi tersebut, penelitian dilanjutkan dengan pengumpulan data kuantitatif menggunakan kuesioner untuk mengukur persepsi mahasiswa dan dosen secara lebih luas dan terstruktur. Data kuantitatif ini digunakan untuk memetakan kecenderungan umum persepsi pengguna terhadap sistem penilaian esai berbasis AI serta hubungannya dengan niat penggunaan sistem.

Selanjutnya, hasil analisis kualitatif dan kuantitatif diintegrasikan melalui proses triangulasi. Integrasi ini bertujuan untuk memastikan bahwa temuan penelitian tidak hanya bersifat deskriptif, tetapi juga kontekstual dan interpretatif, sehingga mampu memberikan gambaran yang lebih utuh mengenai implementasi sistem penilaian esai berbasis AI di pendidikan tinggi. Alur desain penelitian

mixed-methods yang digunakan dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Desain Penelitian Mixed-Methods Eksploratif

1. Eksplorasi Kualitatif (Isu Pedagogis & Etika) Identifikasi awal isu pedagogis dan etika penggunaan AI dalam penilaian esai.
2. Perumusan Instrumen Penyusunan kuesioner berdasarkan temuan eksplorasi dan kajian literatur.
3. Survei Kuantitatif (Mahasiswa & Dosen) Pengumpulan data persepsi mahasiswa dan dosen secara terstruktur.
4. Analisis Data Analisis statistik dan tematik untuk mengidentifikasi pola persepsi.
5. Integrasi & Triangulasi Penggabungan temuan kualitatif dan kuantitatif untuk validasi hasil.

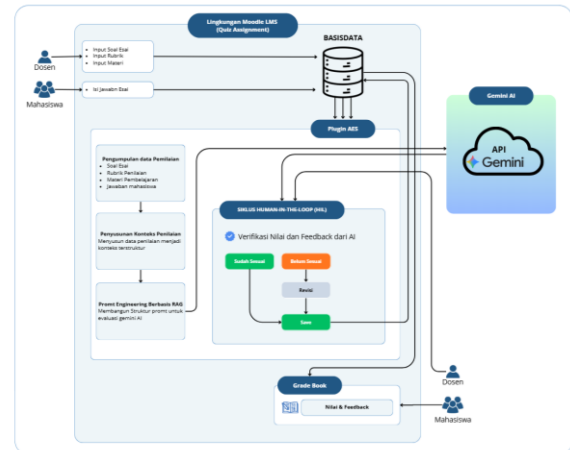
B. Konteks dan Arsitektur Sistem

Sistem yang digunakan dalam penelitian ini merupakan sistem penilaian esai berbasis kecerdasan artifisial yang terintegrasi dengan Learning Management System (LMS) Moodle. Integrasi ini memungkinkan proses pengumpulan jawaban esai, penilaian berbasis AI, serta validasi dosen dilakukan dalam satu ekosistem pembelajaran yang sama, sehingga mendukung alur evaluasi yang efisien dan terkontrol.

Sistem penilaian memanfaatkan Large Language Model sebagai alat bantu analisis jawaban esai. Untuk meningkatkan relevansi konteks penilaian, sistem menerapkan pendekatan Retrieval-Augmented Generation (RAG), yaitu mekanisme yang memungkinkan AI mengakses informasi kontekstual seperti rubrik penilaian, deskripsi tugas, dan materi pembelajaran sebelum menghasilkan skor dan umpan balik. Dengan pendekatan ini, penilaian tidak dilakukan semata-mata berdasarkan teks jawaban mahasiswa, tetapi juga mempertimbangkan konteks akademik yang telah ditentukan oleh dosen.

Sistem dirancang dengan pendekatan *Human-in-the-Loop* (HIL), di mana hasil penilaian yang dihasilkan oleh AI bersifat rekomendatif dan belum menjadi keputusan akhir.

Dosen memiliki kewenangan penuh untuk meninjau, mengoreksi, atau menyesuaikan skor dan umpan balik yang diberikan AI sebelum nilai akhir ditetapkan. Pendekatan ini menempatkan AI sebagai alat bantu (AI-assisted assessment), sementara tanggung jawab akademik dan otoritas penilaian tetap berada pada dosen [13]. Arsitektur sistem penilaian esai berbasis AI yang digunakan dalam penelitian ini ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur Pengembangan sistem AES [13]

Dengan konfigurasi tersebut, sistem penilaian esai berbasis AI dalam penelitian ini tidak dimaksudkan untuk menggantikan peran dosen, melainkan untuk mendukung proses evaluasi secara lebih efisien, konsisten, dan selaras dengan prinsip pedagogis serta etika dalam pendidikan tinggi.

C. Subjek Penelitian dan Teknik Sampling

Subjek penelitian terdiri dari dua kelompok responden, yaitu mahasiswa dan dosen, yang terlibat langsung dalam penggunaan sistem penilaian esai berbasis AI. Mahasiswa berperan sebagai penerima hasil penilaian, sedangkan dosen berperan sebagai validator dan pengambil keputusan nilai akhir dalam pendekatan *Human-in-the-Loop*.

Pemilihan responden dilakukan menggunakan teknik purposive sampling untuk memastikan bahwa data yang dikumpulkan relevan dengan fokus penelitian, khususnya pada aspek pedagogis dan etika penggunaan AI. Kriteria spesifik responden mahasiswa adalah mahasiswa program sarjana (S1) tingkat akhir (semester 6-8) yang telah mengikuti mata kuliah dengan mekanisme penilaian esai menggunakan sistem AI. Responden dosen adalah dosen pengampu mata kuliah yang secara aktif menggunakan sistem, melakukan validasi, serta memiliki pengalaman minimal satu semester dalam melakukan penilaian esai berbasis AI.

D. Instrumen Penelitian

Instrumen penelitian disusun untuk mengukur persepsi pedagogis dan etika penggunaan sistem penilaian esai berbasis AI dari dua perspektif pengguna, yaitu mahasiswa dan dosen. Pengembangan instrumen merujuk pada kajian literatur terkait *Technology Acceptance Model* (TAM), pedagogi penilaian, serta prinsip *Responsible AI* dalam

pendidikan. Seluruh item menggunakan skala Likert lima poin (1 = sangat tidak setuju hingga 5 = sangat setuju).

Instrumen mahasiswa ditujukan untuk mengukur persepsi mahasiswa sebagai penerima hasil penilaian esai berbasis AI. Dimensi yang diukur mencakup aspek kegunaan, kemudahan penggunaan, kepercayaan dan keadilan, transparansi, nilai pedagogis, penerimaan etis, serta niat penggunaan sistem di masa mendatang. Konstruk yang digunakan untuk mengukur persepsi mahasiswa disajikan pada Tabel 1.

Tabel 1. Konstruk dan Dimensi Instrumen Mahasiswa

Konstruk	Deskripsi Singkat
Perceived Usefulness (PU)	Manfaat sistem dalam membantu penilaian dan pembelajaran
Perceived Ease of Use (PEOU)	Kemudahan penggunaan sistem
Trust & Fairness	Persepsi keadilan dan konsistensi penilaian AI
Transparency	Kejelasan dasar pemberian skor dan umpan balik
Pedagogical Value	Dukungan terhadap proses pembelajaran
Ethical Acceptance	Kenyamanan penggunaan AI dengan keterlibatan dosen
Behavioral Intention	Niat menggunakan sistem di masa mendatang

Instrumen dosen dirancang untuk menggali persepsi dosen terhadap penggunaan AI dalam penilaian esai dari sudut pandang pedagogis dan etika. Fokus instrumen tidak pada aspek teknis sistem, melainkan pada peran dosen, tanggung jawab akademik, kepercayaan terhadap AI, serta penerapan pendekatan Human-in-the-Loop dalam proses penilaian.

Instrumen ini memosisikan dosen sebagai *expert users* yang memiliki otoritas epistemik dan moral dalam penilaian akademik, sehingga jawaban dosen diperlakukan sebagai sumber refleksi dan legitimasi pedagogis dari penggunaan AI. Dimensi instrumen yang digunakan untuk mengukur persepsi dosen disajikan pada Tabel 2

Tabel 2. Konstruk dan Dimensi Instrumen Dosen

Konstruk	Deskripsi Singkat
Pedagogical Role	Peran dosen dalam penilaian berbasis AI
Academic Responsibility	Tanggung jawab dosen atas nilai akhir
Trust in AI	Tingkat kepercayaan dosen terhadap AI sebagai alat bantu
Ethical Perspective	Pandangan etika terhadap penggunaan AI
Acceptance of HIL	Penerimaan dosen terhadap pendekatan Human-in-the-Loop

Sebelum digunakan dalam pengumpulan data, instrumen penelitian melalui proses evaluasi isi (*content validity*) melalui kajian literatur dan penilaian ahli. Selanjutnya, uji

reliabilitas dilakukan menggunakan Cronbach's Alpha untuk memastikan konsistensi internal setiap konstruk, khususnya pada instrumen mahasiswa yang dianalisis secara kuantitatif.

E. Teknik Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui beberapa teknik untuk mendukung keterpaduan dan validitas temuan penelitian. Teknik yang digunakan melibatkan kuesioner tertutup, pertanyaan terbuka, serta dokumentasi sistem, dengan melibatkan dua kelompok responden utama, yaitu mahasiswa dan dosen.

Kuesioner digunakan sebagai instrumen utama untuk mengumpulkan data kuantitatif mengenai persepsi mahasiswa dan dosen terhadap penggunaan sistem penilaian esai berbasis AI. Kuesioner disebarikan secara daring setelah responden menggunakan sistem penilaian esai berbasis AI yang terintegrasi dengan LMS Moodle. Responden diminta memberikan penilaian berdasarkan pengalaman langsung mereka dalam menggunakan atau menerima hasil penilaian dari sistem tersebut.

Selain kuesioner tertutup, penelitian ini juga menggunakan pertanyaan terbuka untuk memperoleh data kualitatif. Pertanyaan terbuka ditujukan untuk menggali pandangan, pengalaman, serta refleksi pedagogis dan etika responden terkait penggunaan AI dalam penilaian esai, termasuk kekhawatiran terhadap keadilan, transparansi, serta peran dosen dalam proses evaluasi.

Sebagai data pendukung, penelitian ini juga memanfaatkan dokumentasi sistem, meliputi deskripsi alur kerja penilaian esai berbasis AI, mekanisme pemanfaatan konteks melalui pendekatan *Retrieval-Augmented Generation*, serta proses validasi dosen dalam pendekatan *Human-in-the-Loop*. Dokumentasi ini digunakan untuk memperkuat pemahaman kontekstual terhadap temuan penelitian dan memastikan kesesuaian antara desain sistem dan persepsi pengguna.

III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian yang diperoleh dari analisis data kuantitatif dan kualitatif berdasarkan kuesioner yang disebarikan kepada mahasiswa dan dosen. Penyajian hasil difokuskan pada persepsi pengguna terhadap aspek pedagogis dan etika penggunaan sistem penilaian esai berbasis AI yang terintegrasi dengan LMS Moodle serta didukung oleh pendekatan *Retrieval-Augmented Generation* dan *Human-in-the-Loop*.

A. Karakteristik Responden

Penelitian ini melibatkan 60 responden yang terdiri dari 36 mahasiswa dan 24 dosen. Seluruh responden telah menyelesaikan instrumen penelitian secara lengkap, sehingga data tersebut dinyatakan valid untuk dianalisis lebih lanjut. Ringkasan profil responden berdasarkan kelompok partisipan disajikan pada Tabel 3.

Tabel 3. Karakteristik Responden Penelitian

Kelompok Responden	Peran dalam Sistem	Jumlah Responden
--------------------	--------------------	------------------

Mahasiswa	Penerima hasil penilaian AI	36
Dosen	Validator & penentu nilai akhir	24
Total		60

Data pada Tabel 3 menunjukkan bahwa kelompok mahasiswa mendominasi jumlah partisipan sebagai penerima hasil penilaian, sementara kelompok dosen berpartisipasi dalam kapasitasnya sebagai validator serta pengambil keputusan nilai akhir.

B. Validitas dan Reliabilitas Instrumen

Instrumen penelitian dikembangkan melalui kajian literatur yang relevan dan dievaluasi kualitasnya menggunakan pendekatan validitas isi serta uji reliabilitas internal dengan koefisien Cronbach's Alpha. Hasil pengujian menunjukkan bahwa instrumen yang digunakan pada responden mahasiswa dan dosen memiliki konsistensi internal yang baik, sehingga dinyatakan andal dan layak digunakan untuk analisis lebih lanjut.

C. Temuan dan Pembahasan

Hasil analisis persepsi mahasiswa terhadap penggunaan sistem penilaian esai berbasis kecerdasan artifisial. Analisis didasarkan pada data kuesioner mahasiswa yang diisi oleh 36 responden. Mahasiswa merupakan pihak yang menerima secara langsung hasil penilaian esai berbasis AI, sehingga persepsi mereka merefleksikan dampak sistem terhadap proses pembelajaran dan evaluasi akademik.

Analisis statistik deskriptif dilakukan pada tujuh konstruk utama, yaitu *Perceived Usefulness*, *Perceived Ease of Use*, *Trust & Fairness*, *Transparency*, *Pedagogical Value*, *Ethical Acceptance*, dan *Behavioral Intention*. Nilai mean digunakan untuk menggambarkan kecenderungan persepsi responden, sedangkan standar deviasi digunakan untuk menunjukkan tingkat variasi jawaban mahasiswa.

Tabel 4. Statistik Deskriptif Persepsi Mahasiswa terhadap Sistem Penilaian Esai Berbasis AI (n = 36)

Konstruk	Mean	Standar Deviasi
Perceived Usefulness (PU)	3.83	0.95
Perceived Ease of Use (PEOU)	3.86	0.90
Trust & Fairness	3.68	0.87
Transparency	3.77	0.82
Pedagogical Value	4.01	0.86
Ethical Acceptance	4.18	0.76
Behavioral Intention	3.94	0.76

Skor menggunakan skala Likert 1–5. Nilai mean menunjukkan kecenderungan persepsi mahasiswa terhadap

setiap konstruk, sedangkan standar deviasi menunjukkan tingkat variasi jawaban responden.

Hasil pada Tabel 4 menunjukkan bahwa seluruh konstruk memiliki nilai rata-rata di atas titik tengah skala, yang mengindikasikan persepsi positif mahasiswa terhadap penggunaan sistem penilaian esai berbasis AI. Konstruk *pedagogical value* dan *ethical acceptance* memperoleh nilai rata-rata tertinggi, yang menunjukkan bahwa mahasiswa menilai AI bermanfaat dalam mendukung proses pembelajaran dan dapat diterima secara etis.

Nilai mean pada konstruk *ethical acceptance* yang relatif tinggi disertai dengan standar deviasi yang rendah menunjukkan adanya konsensus yang cukup kuat di kalangan mahasiswa mengenai penerimaan etis penggunaan AI, khususnya ketika sistem digunakan sebagai alat bantu dan tidak menggantikan peran dosen. Temuan ini menunjukkan bahwa keberterimaan etis menjadi faktor penting dalam adopsi AI di lingkungan akademik.

Konstruk *trust & fairness* serta *transparency* memperoleh nilai rata-rata yang lebih rendah dibandingkan konstruk lainnya, meskipun masih berada pada kategori positif. Hal ini mengindikasikan bahwa mahasiswa masih memiliki kehati-hatian terhadap kejelasan dasar penilaian dan keadilan sistem AI. Temuan ini sejalan dengan studi yang menyoroti bahwa kurangnya transparansi mengenai dasar pemberian skor merupakan tantangan utama yang dapat menurunkan kepercayaan mahasiswa terhadap objektivitas penilaian [16]. Nilai standar deviasi yang berada pada kategori sedang mencerminkan adanya variasi persepsi antar responden, yang wajar mengingat kompleksitas implikasi etika dan pedagogis penggunaan AI dalam penilaian akademik.

Konstruk *behavioral intention* menunjukkan nilai rata-rata yang relatif tinggi dengan variasi jawaban yang rendah, yang mengindikasikan bahwa mahasiswa memiliki niat yang cukup kuat untuk terus menggunakan sistem penilaian esai berbasis AI di masa mendatang, terutama apabila sistem tersebut tetap melibatkan dosen dalam pengambilan keputusan akhir. Hal ini memperkuat argumen penelitian yang menyatakan bahwa penerimaan mahasiswa terhadap AI dalam ekosistem LMS sangat bergantung pada integrasi yang mempertahankan peran kendali pengajar [3]. Dengan demikian, hasil ini menegaskan bahwa mahasiswa lebih menerima pendekatan AI-assisted assessment dibandingkan penilaian otomatis sepenuhnya.

D. Persepsi Dosen terhadap Sistem Penilaian Esai Berbasis AI

Hasil analisis persepsi dosen terhadap penggunaan sistem penilaian esai berbasis kecerdasan artifisial. Data diperoleh dari 24 responden dosen yang memiliki peran sebagai pengambil keputusan akademik dan validator akhir penilaian. Oleh karena itu, temuan pada subbab ini merefleksikan pandangan dosen dari sisi pedagogis, etika teknologi, serta legitimasi epistemik penggunaan AI dalam evaluasi pembelajaran. Tabel 5 menyajikan ringkasan hasil statistik deskriptif persepsi dosen terhadap sistem penilaian esai berbasis AI berdasarkan empat dimensi utama, yaitu

Human-in-the-Loop, Pedagogis, Etika Teknologi, dan Filsafat Sains (Epistemologi & Otoritas Pengetahuan).

Tabel 5. Statistik Deskriptif Persepsi Dosen terhadap Sistem Penilaian Esai Berbasis AI (n = 24)

<i>Dimensi</i>	<i>Mean</i>	<i>Standar Deviasi</i>
Human-in-the-Loop (HIL)	4.35	0.50
Pedagogis	4.27	0.45
Etika Teknologi	4.61	0.35
Filsafat Sains (Epistemologi & Otoritas Pengetahuan)	4.61	0.42

Skor menggunakan skala Likert 1–5. Nilai mean menunjukkan kecenderungan persepsi dosen terhadap masing-masing dimensi, sedangkan standar deviasi menunjukkan tingkat variasi jawaban responden.

Berdasarkan Tabel 5, seluruh dimensi menunjukkan nilai rata-rata di atas 4,00 dengan standar deviasi yang relatif rendah. Hal ini mengindikasikan persepsi yang sangat positif dan konsensus yang kuat di kalangan dosen terhadap penggunaan sistem penilaian esai berbasis AI dalam kerangka yang terkontrol.

Dimensi *Human-in-the-Loop* memperoleh nilai rata-rata sebesar 4,35 (SD = 0,50), yang menunjukkan bahwa dosen secara konsisten menekankan pentingnya keterlibatan manusia, kendali terhadap nilai akhir, serta tanggung jawab akademik dalam proses penilaian. Temuan ini menegaskan bahwa AI diposisikan sebagai alat bantu yang mendukung efisiensi evaluasi, bukan sebagai pengganti peran dosen. Hasil ini sejalan dengan penelitian yang menegaskan bahwa keterlibatan aktif manusia (human role) merupakan elemen krusial untuk menjaga akuntabilitas etis dan legitimasi akademik dalam ekosistem penilaian berbasis AI[5].

Pada dimensi pedagogis, nilai rata-rata sebesar 4,27 (SD = 0,45) menunjukkan bahwa dosen menilai sistem AI berkontribusi positif terhadap konsistensi penilaian dan dukungan proses pembelajaran. Variasi jawaban yang rendah mengindikasikan adanya kesepakatan bahwa AI dapat membantu evaluasi awal, sementara penilaian mendalam tetap memerlukan interpretasi manusia.

Dimensi etika teknologi memperoleh nilai rata-rata tertinggi, yaitu 4,61 (SD = 0,35). Temuan ini menunjukkan penerimaan etis yang sangat kuat terhadap penggunaan AI selama berada di bawah pengawasan dosen dan tidak menghilangkan peran moral serta profesional pendidik. Transparansi dan akuntabilitas dipandang sebagai prasyarat utama penerapan AI yang bertanggung jawab.

Pada dimensi filsafat sains, nilai rata-rata sebesar 4,61 dengan standar deviasi 0,42 menunjukkan adanya konsensus yang tinggi bahwa penilaian esai merupakan proses interpretatif yang tidak dapat sepenuhnya diotomatisasi. AI dipersepsikan mampu membantu analisis awal, namun tidak memiliki otoritas epistemik dalam pengambilan keputusan akademik. Argumen ini didukung oleh Selvam dan Vallejo yang menegaskan bahwa model *Human-in-the-Loop* dalam penilaian berbasis AI tidak hanya berperan dalam

meningkatkan efektivitas proses penilaian, tetapi juga menjadi mekanisme pengawasan manusia untuk menjaga aspek etika, keadilan, transparansi, dan integritas akademik dalam evaluasi esai yang kompleks.[17]. Dengan demikian, pendekatan *Human-in-the-Loop* dipandang sebagai model penggunaan AI yang selaras dengan nilai dan praktik pendidikan tinggi.

E. Integrasi Temuan dan Implikasi Pedagogis–Etis Penggunaan AI dalam Penilaian Esai

Temuan penelitian dengan mengintegrasikan hasil persepsi mahasiswa dan dosen terhadap penggunaan sistem penilaian esai berbasis kecerdasan artifisial. Integrasi ini bertujuan untuk memberikan pemahaman yang lebih komprehensif mengenai implikasi pedagogis, etika teknologi, serta legitimasi akademik penerapan AI dalam konteks pendidikan tinggi.

Hasil penelitian menunjukkan adanya perbedaan orientasi penerimaan antara mahasiswa dan dosen. Mahasiswa cenderung menilai sistem AI dari sisi fungsional dan pengalaman penggunaan, yang tercermin dari persepsi positif terhadap kegunaan, kemudahan penggunaan, dan nilai pedagogis sistem. AI dipersepsikan membantu mempercepat umpan balik dan mendukung konsistensi penilaian, sehingga memberikan manfaat langsung dalam proses pembelajaran. Namun demikian, persepsi mahasiswa terhadap kepercayaan, keadilan, dan transparansi penilaian masih berada pada tingkat sedang, yang mengindikasikan adanya kehati-hatian terhadap otonomi sistem AI dalam menentukan nilai akademik.

Sebaliknya, dosen menunjukkan penerimaan yang sangat kuat terhadap penggunaan AI dalam penilaian esai sepanjang sistem diterapkan dalam kerangka yang terkontrol dan bertanggung jawab. Persepsi positif yang konsisten pada dimensi *Human-in-the-Loop*, pedagogis, etika teknologi, dan filsafat sains menunjukkan bahwa dosen menempatkan AI sebagai alat bantu pendukung, bukan sebagai pengganti peran pendidik. Temuan ini menegaskan bahwa legitimasi penilaian akademik tetap bergantung pada pertimbangan manusia sebagai pemegang otoritas epistemik.

Integrasi temuan mahasiswa dan dosen menunjukkan bahwa pendekatan *Human-in-the-Loop* berperan sebagai titik temu utama antara efisiensi teknologi dan nilai-nilai pendidikan. Bagi mahasiswa, keterlibatan dosen dalam proses penilaian meningkatkan rasa keadilan dan kepercayaan terhadap hasil evaluasi. Bagi dosen, pendekatan ini memastikan bahwa tanggung jawab akademik, etika profesional, dan kualitas pedagogis tetap terjaga.

Dari perspektif pedagogis, hasil penelitian menunjukkan bahwa penggunaan AI lebih tepat diposisikan sebagai bagian dari *AI-assisted assessment*, yaitu sebagai sarana pendukung evaluasi awal dan pemberian umpan balik cepat, sementara evaluasi mendalam tetap dilakukan oleh dosen. Pendekatan ini memungkinkan optimalisasi manfaat teknologi tanpa mengorbankan kualitas pembelajaran dan interaksi akademik.

Dari sisi etika dan epistemologi, temuan penelitian menegaskan bahwa AI tidak memiliki otoritas epistemik dalam penilaian esai, melainkan berfungsi sebagai sistem pendukung pengambilan keputusan. Oleh karena itu, penerapan AI dalam penilaian pembelajaran perlu disertai dengan prinsip transparansi, akuntabilitas, dan pengawasan manusia agar selaras dengan nilai dan tujuan pendidikan tinggi.

VI. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa sistem penilaian esai berbasis AI dalam paradigma AI-assisted assessment diterima secara positif oleh mahasiswa meski terdapat kekhawatiran terkait transparansi, sementara dosen memberikan dukungan penuh selama sistem tetap menempatkan manusia sebagai pemegang kendali (Human-in-the-Loop). Meskipun temuan ini memberikan kontribusi signifikan terhadap pemahaman integrasi teknologi dalam penilaian akademik, penelitian ini memiliki keterbatasan pada ukuran sampel yang spesifik dan cakupan konteks institusional yang terbatas, sehingga belum sepenuhnya merepresentasikan keragaman populasi pendidikan tinggi secara luas. Oleh karena itu, penelitian selanjutnya disarankan untuk melibatkan sampel yang lebih representatif, mengevaluasi dampak jangka panjang penggunaan AI terhadap capaian belajar, serta mendalami aspek teknis lebih lanjut, seperti mitigasi bias algoritmik dan pengembangan fitur *explainability* untuk meningkatkan kepercayaan dan keadilan bagi pengguna. Penelitian ini menyimpulkan bahwa sistem penilaian esai berbasis AI dalam paradigma AI-assisted assessment diterima secara positif oleh mahasiswa meski terdapat kekhawatiran terkait transparansi, sementara dosen memberikan dukungan penuh selama sistem tetap menempatkan manusia sebagai pemegang kendali (*Human-in-the-Loop*). Meskipun temuan ini memberikan kontribusi signifikan terhadap pemahaman integrasi teknologi dalam penilaian akademik, penelitian ini memiliki keterbatasan pada ukuran sampel yang spesifik dan cakupan konteks institusional yang terbatas, sehingga belum sepenuhnya merepresentasikan keragaman populasi pendidikan tinggi secara luas. Oleh karena itu, penelitian selanjutnya disarankan untuk melibatkan sampel yang lebih representatif, mengevaluasi dampak jangka panjang penggunaan AI terhadap capaian belajar, serta mendalami aspek teknis lebih lanjut, seperti mitigasi bias algoritmik dan pengembangan fitur *explainability* untuk meningkatkan kepercayaan dan keadilan bagi pengguna.

DAFTAR PUSTAKA

- [1] D. Ifenthaler *et al.*, “Artificial Intelligence in Education: Implications for Policymakers, Researchers, and Practitioners,” *Technology, Knowledge and Learning*, vol. 29, no. 4, pp. 1693–1710, Dec. 2024, doi: 10.1007/s10758-024-09747-0.
- [2] Mohammed Siddique Kadwa, “Written Assignments and the Ethical Considerations of Artificial Intelligence in Higher Education,” *International Journal of Linguistics, Literature and Translation*, vol. 8, no. 9, pp. 254–257, Sep. 2025, doi: 10.32996/ijllt.2025.8.9.27.
- [3] R. Manhica, A. Santos, and J. Cravino, “The use of artificial intelligence in learning management systems in the context of higher education: Systematic literature review,” in *2022 17th Iberian Conference on Information Systems and Technologies (CISTI)*, IEEE, Jun. 2022, pp. 1–6. doi: 10.23919/CISTI54924.2022.9820205.
- [4] I. Ogunsakin, “Unlocking The Potential of Artificial Intelligence:,” *International Journal of Research in STEM Education*, vol. 6, no. 2, pp. 37–49, Nov. 2024, doi: 10.33830/ijrse.v6i2.1698.
- [5] M. Ninaus and M. Sailer, “Closing the loop – The human role in artificial intelligence for education,” *Front. Psychol.*, vol. 13, Aug. 2022, doi: 10.3389/fpsyg.2022.956798.
- [6] Rianto and Widya Syafitri, “EXPLORING THE ROLE OF FEEDBACK IN AI AUTOMATED ESSAY SCORING,” *ICMIE Proceedings*, vol. 1, pp. 243–251, Jun. 2025, doi: 10.30983/8km3y331.
- [7] A. El Hajoui, O. Y. Alaoui, O. El Kharki, H. Boulaassal, M. Wahbi, and M. Maatouk, “Automated Generation of Personalized Evaluation Subjects in Higher Education,” in *2025 International Conference on Circuit, Systems and Communication (ICCSC)*, IEEE, Jun. 2025, pp. 1–5. doi: 10.1109/ICCSC66714.2025.11135129.
- [8] E. Emirtekin, “Large Language Model-Powered Automated Assessment: A Systematic Review,” *Applied Sciences*, vol. 15, no. 10, p. 5683, May 2025, doi: 10.3390/app15105683.
- [9] M. Kostic, H. F. Witschel, K. Hinkelmann, and M. Spahic-Bogdanovic, “LLMs in Automated Essay Evaluation: A Case Study,” *Proceedings of the AAAI Symposium Series*, vol. 3, no. 1, pp. 143–147, May 2024, doi: 10.1609/aaais.v3i1.31193.
- [10] J. Y. H Bai *et al.*, “Automated essay scoring (AES) systems: Opportunities and challenges for open and distance education,” 2014.
- [11] J. Du and N. R. binti M. Nordin, “A Systematic Review of Automated Writing Evaluation (AWE) Systems on University Students’ English Writing Performance,” *Forum for Linguistic Studies*, vol. 7, no. 11, Oct. 2025, doi: 10.30564/fls.v7i11.11764.
- [12] G. Boccuzzi, F. Manganello, and A. Nico, “Delegated Authority and Algorithmic Power. A Rapid Review of Ethical Issues in AI-Based Educational Assessment,” in *Proceedings of the 2025 International Conference on Information Technology for Social Good*, New York, NY, USA: ACM, Sep. 2025, pp. 1–5. doi: 10.1145/3748699.3749766.
- [13] R. Adha *et al.*, “Integrasi Gemini AI Berbasis Retrieval-Augmented Generation (RAG) pada Moodle untuk Penilaian Esai Otomatis dengan Pendekatan Human-in-the-Loop di Pendidikan Tinggi,” *Academic Journal of Computer Science Research*, vol. 8, no. 1, p. 67, Jan. 2026, doi: 10.38101/ajcsr.v8i1.16253.
- [14] A. Granić, “Educational Technology Adoption: A systematic review,” *Educ. Inf. Technol. (Dordr.)*,

- vol. 27, no. 7, pp. 9725–9744, Aug. 2022, doi: 10.1007/s10639-022-10951-7.
- [15] J. W. . Creswell and V. L. . Plano Clark, *Designing and conducting mixed methods research*. SAGE, 2018.
- [16] B. Memarian and T. Doleck, “Human-in-the-loop in artificial intelligence in education: A review and entity-relationship (ER) analysis,” *Computers in Human Behavior: Artificial Humans*, vol. 2, no. 1, p. 100053, Jan. 2024, doi: 10.1016/j.chbah.2024.100053.
- [17] M. Selvam and R. González Vallejo, “Human-in-the-Loop Models for Ethical AI Grading: Combining AI Speed with Human Ethical Oversight,” *EthAIca*, vol. 4, p. 413, Aug. 2025, doi: 10.56294/ai2025413.