

Kerangka Konseptual Asesmen Menulis Teks Argumentasi: Desain Rubrik Analitik Berbasis Literasi AI

Ayu Herawati¹, Esti Swatika Sari², Sudiati³

Magister Pendidikan Bahasa dan Sastra Indonesia, Fakultas Bahasa Seni dan Budaya,
Universitas Negeri Yogyakarta, Sleman, Daerah Istimewa Yogyakarta, Indonesia

ayuherawati.2025@student.uny.ac.id¹, esti_swastikasari@uny.ac.id², sudiati@uny.ac.id³

Abstrak: Integrasi *Generative Artificial Intelligence* (GenAI) dalam pendidikan bahasa menuntut redefinisi fundamental terhadap konstruk penilaian menulis. Tantangan utama saat ini adalah ketidakmampuan rubrik konvensional—yang menitikberatkan pada akurasi mekanis dan struktur permukaan—untuk membedakan antara kompetensi kognitif siswa dan kecanggihan algoritma. Artikel ini bertujuan menyusun kerangka konseptual dan desain prototipe rubrik analitik baru untuk asesmen teks argumentasi yang secara eksplisit mengintegrasikan dimensi literasi AI. Menggunakan metode studi konseptual, penelitian ini menyintesis referensi kunci yang mencakup teori asesmen bahasa, model argumentasi Toulmin, dan konsep *distributed agency* (agensi terdistribusi). Penelitian ini mengusulkan pergeseran paradigma asesmen dari *product-oriented* ke *process-oriented* melalui model "ko-kreasi". Hasil utama penelitian ini adalah desain rubrik analitik dengan empat dimensi utama beserta bobot usulannya: (1) Orisinalitas dan Sintesis Gagasan (35%), (2) Evaluasi Kritis Sumber dan Output AI (30%), (3) Logika Argumentasi (25%), dan (4) Mekanik (10%), ditambah dengan prasyarat Transparansi Etis. Kerangka ini tidak hanya merespons isu plagiarisme, tetapi juga dirancang untuk mendorong strategi *self-regulated learning*. Temuan ini diharapkan menjadi acuan bagi pendidik dan pengambil kebijakan dalam merancang ekosistem asesmen yang valid, etis, dan mampu menumbuhkan nalar kritis siswa di era disrupsi teknologi.

Kata kunci: Asesmen Menulis, Teks Argumentasi, Rubrik Analitik, Literasi AI

Conceptual Framework for Argumentative Writing Assessment: AI Literacy-Based Analytical Rubric Design

Abstract: The integration of *Generative Artificial Intelligence* (GenAI) in language education demands a fundamental redefinition of the construct of writing assessment. A major challenge currently is the inability of conventional rubrics—which emphasize mechanical accuracy and surface structure—to distinguish between students' cognitive competence and algorithmic sophistication. This article aims to develop a conceptual framework and prototype design of a new analytical rubric for the assessment of argumentative texts that explicitly integrates the AI literacy dimension. Using a conceptual study method, this research synthesizes key references including language assessment theory, the Toulmin argumentation model, and the concept of *distributed agency*. This research proposes a paradigm shift in assessment from *product-oriented* to *process-oriented* through a "co-creation" model. The main result of this research is the design of an analytical rubric with four main dimensions and their proposed weights: (1) Originality and Idea Synthesis (35%), (2) Critical Evaluation of AI Sources and Outputs (30%), (3) Argumentation Logic (25%), and (4) Mechanics (10%), coupled with the prerequisite of *Ethical Transparency*. This framework not only addresses the issue of plagiarism but is also designed to encourage *self-regulated learning* strategies. These findings are expected to serve as a reference for educators and policymakers in designing an assessment ecosystem that is valid, ethical, and capable of fostering critical thinking in students in an era of technological disruption.

Keywords: Writing Assessment, Argumentative Text, Analytical Rubric, AI Literacy.

1. Pendahuluan

Dunia pendidikan bahasa saat ini sedang berada di persimpangan jalan yang sangat kritis akibat ledakan teknologi kecerdasan buatan atau *Artificial Intelligence*

(AI). Sebagaimana dijelaskan oleh Warschauer & Xu (2024), peluncuran ChatGPT

menandai masuknya kita ke era baru. Era yang membuat cara siswa belajar dan menulis tidak akan pernah sama lagi.

Dalam kurikulum pendidikan di Indonesia, khususnya di tingkat Sekolah Menengah Atas (SMA), keterampilan menulis teks argumentasi memiliki posisi yang sangat strategis.

Keterampilan ini bukan sekadar soal merangkai kata-kata indah, melainkan indikator utama apakah seorang siswa mampu berpikir kritis (*Higher Order Thinking Skills*) (Wang & Newell, 2025). Sebuah argumen yang baik menuntut siswa untuk menyusun klaim yang kuat, menyajikan data yang relevan, dan membangun logika yang runtut (Toulmin, 2003).

Namun, realitas di lapangan menunjukkan bahwa menulis argumentasi adalah salah satu materi yang paling sulit dikuasai oleh siswa. Banyak siswa SMA merasa kesulitan untuk menuangkan ide abstrak mereka menjadi tulisan yang logis dan terstruktur (T. Jin et al., 2020). Kesulitan ini sering kali diperparah oleh kendala bahasa dan kurangnya wawasan. Di tengah kesulitan ini, kehadiran AI seperti ChatGPT dilihat sebagai "dewa penolong" (Pratiwi et al., 2024). Lebih lanjut, Pratiwi et al. (2024) memaparkan bahwa AI bagi siswa di Indonesia menawarkan bantuan instan untuk menyelesaikan tugas. Namun, jika penggunaan ini tidak dibimbing, ada risiko besar bahwa siswa hanya akan menjadi pengguna pasif yang malas berpikir, sehingga daya nalar kritis mereka justru tumpul.

Feng et al. (2024) melakukan tinjauan riset terhadap 112 studi dan mengonfirmasi bahwa kemampuan AI dalam menulis teks argumentasi sudah sangat berkembang. Hal ini diperkuat oleh temuan meta-analisis terhadap 61 studi terbaru yang menunjukkan bahwa pembelajaran bahasa berbasis AI memiliki efektivitas tinggi ($d = 1.18$) dalam meningkatkan kemampuan bahasa siswa dibandingkan metode konvensional (Lee & Lee, 2024).

Teknologi AI kini sudah sangat canggih; ia mampu menghasilkan esai argumentatif dengan bahasa yang rapi, struktur yang bagus, dan ejaan yang sempurna hanya dalam hitungan detik (Kasneci et al., 2023). Hal ini secara fundamental mendisrupsi validitas penilaian guru di sekolah. Bayangkan jika guru masih menggunakan rubrik penilaian lama yang memberi bobot tinggi untuk aspek permukaan seperti 'kerapian bahasa' dan 'struktur paragraf'—kemampuan yang kini dapat dilakukan AI dengan sempurna (Kasneci et al., 2023). Akibatnya, siswa yang hanya menyalin dari AI tanpa melakukan penalaran apapun berpotensi mendapat nilai sempurna. Kondisi ini mencederai prinsip validitas asesmen. Penilaian seharusnya mengukur kompetensi siswa yang sebenarnya, bukan kecanggihan alat yang mereka gunakan (Brown, 2004).

Selain itu, transformasi teknologi ini juga memunculkan perubahan epistemologis dalam praktik menulis itu sendiri. Aktivitas menulis

yang sebelumnya bergantung pada kemampuan internal individu kini berkembang menjadi proses ko-konstruktif yang melibatkan interaksi dinamis antara penulis dan sistem AI, sebagaimana digambarkan dalam konsep *distributed agency* (Godwin-Jones, 2024).

Pergeseran ini menantang asumsi lama tentang kompetensi menulis yang selama ini berorientasi produk, dan memperkuat argumen bahwa asesmen harus mengakomodasi dimensi proses serta kemampuan berpikir kritis siswa (Warschauer & Xu, 2024; Wang & Newell, 2025). Dalam konteks ini, kebutuhan terhadap instrumen penilaian baru yang mampu menangkap cara siswa berkolaborasi, mengevaluasi, dan mengolah masukan AI menjadi semakin mendesak, terutama karena rubrik konvensional terbukti tidak dapat membedakan antara kinerja kognitif siswa dan keluaran algoritmik (Kasneci et al., 2023; Feng et al., 2024).

Menurut prinsip dasar asesmen, sebuah tes dikatakan valid jika ia benar-benar mengukur apa yang seharusnya diukur (Brown, 2004). Jika rubrik guru hanya mengukur kerapian bahasa (yang dikerjakan AI), maka asesmen itu tidak valid karena gagal mengukur kemampuan berpikir siswa. Kubiszyn dan Borich (2013) juga mengingatkan bahwa instrumen penilaian harus selalu diperbarui mengikuti perkembangan zaman. Di era digital ini, kemampuan menulis tidak lagi berdiri sendiri, tetapi sudah menyatu dengan kemampuan literasi digital dan etika penggunaan teknologi (Kubiszyn & Borich, 2013; Eaton, 2025).

Pada tahap awal integrasi teknologi ini, institusi pendidikan dihadapkan pada tantangan kompleks terkait integritas akademik (Eaton, 2025). Berbagai upaya telah dilakukan, mulai dari pengaturan akses hingga penggunaan perangkat lunak deteksi (*AI detectors*) untuk memverifikasi keaslian karya. Namun, literatur terkini menunjukkan bahwa pendekatan deteksi teknis semata memiliki keterbatasan. Riset oleh Liu et al. (2024) mengungkapkan bahwa akurasi alat pendeteksi AI seringkali tidak tepat, terutama ketika teks telah melalui proses penyuntingan atau diparafrasa oleh siswa, sehingga diperlukan strategi asesmen yang lebih komprehensif melampaui sekadar deteksi.

Studi lain menunjukkan bahwa siswa sebenarnya menginginkan panduan yang jelas mengenai batasan penggunaan AI (Turková et al., 2025). Siswa membutuhkan aturan etis yang transparan dan dapat diterapkan secara konsisten.

Pemerintah Indonesia melalui Kemendikbudristek telah merespons kebutuhan ini dengan menerbitkan *Panduan Penggunaan GenAI pada Pembelajaran di Perguruan Tinggi* (Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, 2024). Panduan tersebut menekankan pentingnya penggunaan AI yang etis, transparan, dan tidak menggantikan kemampuan berpikir kritis.

Selain lembaga nasional, lembaga internasional seperti APA telah merilis aturan cara mengutip AI (McAdoo, 2024). Walaupun dalam praktiknya, aturan nasional maupun internasional tersebut masih bersifat administratif—belum menyentuh aspek pedagogis seperti bagaimana AI seharusnya dinilai dalam proses pembelajaran. Sebagai contoh, APA menganjurkan sitasi AI sebagai berikut: *OpenAI. (2025). ChatGPT (versi model). <https://www.openai.com>*.

Oleh karena itu, artikel ini menawarkan solusi yang bersifat jalan tengah. Guru tidak dapat menolak kehadiran teknologi AI, tetapi juga tidak dapat membiarkan siswa menggunakannya secara tidak terarah. Guru perlu memastikan bahwa siswa menggunakan AI secara bertanggung jawab.

Hal ini menuntut perubahan paradigma penilaian, khususnya dalam keterampilan menulis teks argumentasi. Pendekatan yang digunakan dalam penelitian ini merujuk pada konsep agensi terdistribusi (*distributed agency*) (Godwin-Jones, 2024), yang memandang bahwa menulis pada era teknologi bukan lagi aktivitas individual semata, melainkan proses kolaboratif antara manusia dan teknologi pendukungnya.

Berdasarkan permasalahan tersebut, artikel ini bertujuan untuk merancang kerangka konseptual dan prototipe rubrik analitik untuk menilai kemampuan menulis teks argumentasi siswa SMA pada era penggunaan GenAI. Rubrik ini dirancang untuk mengukur kemampuan berpikir kritis, sintesis gagasan, evaluasi sumber, serta transparansi penggunaan AI. Selain memberikan kontribusi teoretis dalam pengembangan asesmen berbasis literasi AI, artikel ini juga menawarkan manfaat praktis bagi guru dan pengambil kebijakan dalam merancang instrumen penilaian yang valid, etis, dan adaptif terhadap disrupsi teknologi.

2. Metode Penelitian

Penelitian ini menggunakan desain Studi Konseptual (*Conceptual Study*). Berbeda dengan penelitian eksperimen atau survei yang membutuhkan data lapangan berupa angka, studi konseptual berfokus pada analisis dan sintesis

teori untuk menghasilkan model, kerangka, atau instrumen baru (Creswell, 2014). Pendekatan ini sangat relevan untuk menjawab masalah yang sifatnya teoretis dan mendesak, seperti kebutuhan akan instrumen asesmen baru di era AI. Validitas isi (*content validity*) harus dibangun terlebih dahulu melalui analisis logis sebelum uji coba lapangan (Kubiszyn & Borich, 2013).

Prosedur pengembangan kerangka rubrik dalam artikel ini mengadaptasi langkah-langkah pengembangan tes standar yang disarankan oleh para ahli pengukuran pendidikan, seperti Oppenheim (1992) dan Kubiszyn & Borich (2013). Prosedur ini terdiri dari tiga tahap utama.

Tahap 1 adalah Analisis Konstruk (*Construct Analysis*). Pada tahap ini, peneliti meninjau ulang definisi "kemampuan menulis argumentasi" dalam konteks Kurikulum Merdeka di SMA. Peneliti mengidentifikasi kompetensi apa saja yang dituntut dari siswa, dan bagaimana kehadiran AI mengubah definisi kompetensi tersebut. Referensi utama pada tahap ini meliputi teori argumentasi Toulmin (2003), literasi digital, dan literasi AI.

Tahap 2 adalah Sintesis Dimensi Penilaian. Peneliti mengumpulkan artikel jurnal internasional dan buku teks terpilih (yang terbit dalam rentang 2023-2025 untuk jurnal) untuk mengidentifikasi indikator penting asesmen menulis pada era AI. Indikator-indikator kemudian dikelompokkan menjadi beberapa dimensi penilaian. Tahap ini memastikan bahwa rubrik yang dirumuskan memiliki dasar teoretis kuat, bukan sekadar asumsi intuitif.

Tahap 3 adalah Perancangan Prototipe Rubrik (*Drafting*). Berdasarkan analisis konstruk dan sintesis dimensi pada tahap sebelumnya, peneliti menyusun draf rubrik analitik. Pada tahap ini ditentukan bobot untuk setiap kriteria (misalnya, mengapa logika lebih penting daripada ejaan) dan deskripsi kinerja untuk setiap level nilai.

Pemilihan studi konseptual sebagai pendekatan metodologis didasarkan pada karakter fenomena GenAI yang bersifat disruptif dan menuntut rekonstruksi teori sebelum pengujian empiris, sebagaimana dianjurkan dalam pengembangan instrumen oleh Kubiszyn dan Borich (2013) serta Oppenheim (1992). Pendekatan ini memungkinkan pemetaan ulang konstruk menulis argumentasi di era AI pada level konseptual, sebelum dirumuskan sebagai prototipe rubrik.

Analisis artikel dilakukan dengan mempertimbangkan relevansinya terhadap literasi AI (misalnya, halusinasi dan validasi sumber; H. Jin, 2025), teori argumentasi

(Toulmin, 2003), dan kajian pedagogi AI mutakhir (Bearman & Ajjawi, 2023; Woo et al., 2024). Meskipun tidak menghasilkan data statistik, pendekatan ini memberikan landasan teoretis yang kuat untuk merespons kebutuhan asesmen baru secara sistematis, konsisten dengan panduan desain penelitian konseptual oleh Creswell (2014).

3. Hasil dan Pembahasan

Berdasarkan kajian teori dan analisis kebutuhan yang telah dilakukan, artikel ini mengusulkan sebuah desain rubrik analitik baru. Perbedaan paling mencolok dari rubrik ini adalah pergeseran bobot penilaian. Aspek mekanik (seperti ejaan) dikurangi drastis, sementara aspek kognitif tingkat tinggi (seperti evaluasi dan sintesis) diberikan porsi terbesar.

Mengacu pada prinsip validitas konstruk yang telah dibahas, instrumen asesmen perlu direkonstruksi agar mampu menangkap kompetensi esensial di era digital. Tabel berikut ini menyajikan usulan desain rubrik analitik yang telah dikembangkan. Desain ini secara spesifik memetakan empat dimensi utama penilaian beserta bobotnya, yang mencerminkan prioritas baru dalam pembelajaran menulis argumentasi berbantuan AI.

Penetapan dimensi dan distribusi bobot dalam rubrik usulan tersebut didasarkan pada sintesis mendalam terhadap literatur mengenai literasi AI, teori argumentasi, dan psikologi pendidikan. Rasionalisasi berikut menjelaskan argumen akademis di balik setiap keputusan desain, menyoroti bagaimana setiap elemen rubrik berfungsi sebagai mekanisme kontrol kualitas sekaligus alat pedagogis untuk membimbing siswa di tengah tantangan disrupsi teknologi.

Mengapa dimensi *Orisinalitas & Sintesis* memiliki bobot tertinggi (35%)? Riset membuktikan bahwa kolaborasi antara manusia dan AI memang bisa menghasilkan tulisan yang lebih berkualitas (Woo et al., 2024). Namun, kualitas ini sangat bergantung pada seberapa besar usaha kognitif yang dikeluarkan siswa. Chen et al. (2025) memperingatkan adanya risiko pengikisan kemampuan berpikir jika siswa terlalu bergantung pada AI tanpa berpikir. Selain itu, rubrik ini juga mengakomodasi fase perkembangan interaksi siswa dengan AI, dari sekadar eksplorasi antusias menuju keterlibatan kritis (Zhou & Du Preez, 2025).

Oleh karena itu, rubrik ini dirancang untuk mencegah siswa menjadi pasif. Dengan bobot nilai yang tinggi pada aspek orisinalitas, siswa dipaksa untuk melakukan transformasi jika

menggunakan bantuan teks dari AI. Studi oleh Zhang et al. (2025b) menemukan bahwa siswa yang terlibat secara mendalam dengan respons AI—yaitu membaca, menganalisis, dan mengeditnya—akan memiliki pemahaman logika yang lebih baik. Sebaliknya, siswa yang hanya sibuk meminta jawaban tanpa membacanya,

Tabel 1. Usulan Rubrik Asesmen Argumentasi Berbasis Literasi AI (SMA)

Dimensi Penilaian	Bobot	Fokus Penilaian
1. Orisinalitas & Sintesis Gagasan	35%	Menilai apakah siswa memiliki "suara" penulis (<i>authorial voice</i>) sendiri. Apakah siswa mampu menggabungkan informasi dari AI dengan pengalaman pribadi, konteks lokal, atau wawasan uniknya sendiri? Mengukur kemampuan siswa melakukan transformasi ide, bukan sekadar menyalin.
2. Evaluasi Kritis & Validasi Sumber	30%	Menilai kemampuan siswa dalam melakukan cek fakta (<i>fact-checking</i>). Apakah siswa mampu mendeteksi jika AI memberikan informasi yang salah atau halusinasi? Apakah sumber yang dikutip valid dan nyata?
3. Logika Argumentasi (Model Toulmin)	25%	Menilai ketajaman struktur argumen. Apakah klaim yang disusun jelas? Apakah data mendukung klaim? Apakah ada referensi yang menghubungkan data dan klaim secara logis?
4. Mekanik & Kebahasaan	10%	Menilai ejaan, tanda baca, dan struktur kalimat sesuai EYD. Bobotnya kecil karena aspek ini sudah banyak dibantu oleh teknologi.
Prasyarat: Transparansi Etis	-	Siswa wajib menyertakan Deklarasi Penggunaan AI. - Jika menggunakan AI: Lampirkan Prompt Log (bukti chat). - Jika TIDAK menggunakan AI: Lampirkan pernyataan tertulis orisinalitas murni.

justru mengalami penurunan kemampuan. Rubrik ini menghargai proses "mengolah", bukan proses "meminta".

Dimensi kedua, *Evaluasi Kritis* (30%) adalah respons langsung terhadap kelemahan utama teknologi GenAI saat ini, yaitu halusinasi. Riset oleh F. Jin (2024) menunjukkan bahwa meskipun AI bisa menulis dengan lancar, ia sering mengarang fakta atau sumber referensi yang tidak ada. Di sinilah peran manusia menjadi tak tergantikan.

Rubrik ini menuntut siswa untuk berperan sebagai *editor* atau *validator*. Jika siswa memang menggunakan bantuan AI, maka siswa perlu melakukan validasi dengan mengecek kebenaran informasi dari AI. Hal ini sejalan dengan temuan Shi et al. (2024) yang menyatakan bahwa siswa perlu melewati fase kritis (*critical phase*) saat menerima masukan dari teknologi. Jika siswa menelan mentah-mentah informasi dari AI, nilainya di dimensi ini akan rendah.

Meskipun teknologi berubah, prinsip dasar argumentasi yang baik tetap sama. Oleh karena itu, dimensi *Logika Argumentasi* (25%) tetap dipertahankan menggunakan model klasik Toulmin. Interaksi dengan AI seharusnya membantu siswa memperkuat logika ini, misalnya dengan meminta AI mencari argumen lawan (*counter-argument*) untuk dibantah. Studi Zhang et al. (2025a) menegaskan bahwa interaksi yang terstruktur dapat meningkatkan level analisis siswa. Rubrik ini memastikan bahwa teknologi hanya menjadi alat bantu, sementara kendali logika tetap di tangan siswa.

Salah satu inovasi terpenting dalam desain rubrik ini adalah adanya *Prasyarat Transparansi*. Siswa wajib melampirkan Deklarasi Penggunaan AI. Ini untuk menghindari prasangka bahwa semua siswa pasti menggunakan AI. Bagi siswa yang menggunakan AI, mereka harus menyertakan bukti *Prompt Log* sebagai wujud *Prompt Literacy* yang kini diakui sebagai keterampilan akademis baru (Plaatjes & van Wyk, 2025). Bagi siswa yang memilih tidak menggunakan AI, deklarasi orisinalitas menjadi bukti integritas mereka.

Pendekatan ini sejalan dengan konsep Postplagiarism (Eaton, 2025). Di masa depan, integritas bukan berarti *menulis sendirian di kamar tertutup*, melainkan *jujur dan transparan tentang proses penulisan*. Dengan mekanisme deklarasi ini, guru mendidik siswa untuk bertanggung jawab atas pilihan alat mereka, sekaligus mengurangi bias penilaian guru (Henestrosa & Kimmerle, 2024).

Selain mendidik siswa, rubrik analitik yang rinci ini juga berfungsi untuk melindungi siswa dari bias subjektif guru. Dalam konteks penggunaan AI, penilaian manusia sering kali rentan terhadap bias. Riset eksperimental oleh Lermann Henestrosa dan Kimmerle (2024) menunjukkan bahwa manusia cenderung memberikan penilaian kredibilitas yang lebih rendah ketika mereka mengetahui (atau menduga) bahwa sebuah teks dihasilkan oleh AI, terlepas dari kualitas aktual teks tersebut. Tanpa panduan rubrik yang jelas, seorang guru mungkin secara tidak sadar memberi nilai buruk hanya karena *curiga* siswa menggunakan AI.

Dengan memecah penilaian ke dalam dimensi spesifik (seperti validasi sumber dan logika), rubrik ini memaksa guru untuk menilai berdasarkan bukti kinerja yang objektif, bukan berdasarkan asumsi asal-usul teks. Ini menjamin keadilan dalam asesmen, memastikan bahwa siswa dinilai berdasarkan kompetensi mereka dalam mengelola informasi, bukan dihukum karena menggunakan alat yang tersedia di zaman mereka.

Selain itu, penerapan rubrik ini juga selaras dengan tuntutan Kurikulum Merdeka yang menekankan pengembangan kemampuan bernalar kritis sebagai bagian dari Dimensi Profil Lulusan. Dengan memberikan bobot tinggi pada sintesis gagasan dan evaluasi kritis, rubrik ini mendorong siswa untuk menjalani proses berpikir tingkat tinggi sebagaimana direkomendasikan oleh Wang dan Newell (2025). Rubrik ini juga berfungsi sebagai alat pedagogis yang memandu siswa memasuki *critical phase* penggunaan AI, sebagaimana diuraikan oleh Shi et al. (2024), yaitu tahap ketika siswa harus mengevaluasi, memverifikasi, dan mengolah kembali masukan teknologi.

Dalam konteks pembelajaran nyata, rubrik ini mampu membedakan dua tipe siswa: siswa yang hanya menerima respons AI secara pasif—yang menurut Chen et al. (2025) rentan mengalami pengikisan usaha kognitif—dan siswa yang menggunakan AI sebagai titik awal untuk mengembangkan argumen mereka. Dengan demikian, rubrik ini tidak hanya menjadi alat penilaian, tetapi juga instrumen yang memperkuat integrasi AI secara etis dan transformatif dalam pembelajaran bahasa.

Implementasi rubrik ini menuntut prasyarat penting: kesiapan guru. Guru tidak bisa menilai *Evaluasi Kritis Sumber AI* jika guru tersebut tidak paham cara kerja halusinasi AI. Studi terbaru oleh Chan dan Tang (2024) menekankan bahwa integrasi AI yang efektif membutuhkan kerangka *Technological Pedagogical Content Knowledge*

(TPACK). Guru Bahasa Indonesia di SMA perlu memiliki pengetahuan teknis tentang keterbatasan AI (misalnya, data yang tidak real-time atau bias algoritma) agar dapat memberikan umpan balik yang bermakna pada *Prompt Log* siswa. Oleh karena itu, adopsi rubrik ini idealnya dibarengi dengan program pengembangan profesional guru yang berfokus pada literasi AI pedagogis, bukan sekadar pelatihan teknis operasional.

4. Simpulan dan Saran

Kehadiran AI tidak bisa dibendung, dan siswa di sekolah harus diawasi dalam menggunakan AI secara bertanggung jawab. Artikel ini menawarkan sebuah jalan tengah yang konstruktif bagi guru SMA agar menerima kehadiran teknologi sebagai mitra berpikir (*thinking partner*), namun menuntut standar kognitif yang lebih tinggi dari siswa.

Desain rubrik analitik yang diusulkan dalam artikel ini menggeser paradigma penilaian. Fokus penilaian tidak lagi pada *kebenaran tata bahasa* (yang mudah dilakukan mesin), melainkan pada *kedalaman sintesis, ketajaman evaluasi, dan kejujuran proses*. Rubrik ini dirancang bukan hanya sebagai alat ukur, tetapi sebagai instrumen pembelajaran yang mendorong siswa untuk berevolusi dari sekadar konsumen teknologi menjadi kurator informasi yang kritis.

Agar rubrik ini dapat diterapkan dengan efektif, ada beberapa prasyarat yang harus dipenuhi. Pertama, guru perlu diberikan pelatihan profesional. Studi terbaru menunjukkan bahwa kesiapan guru dalam hal teknis dan pedagogis adalah kunci keberhasilan integrasi AI di kelas (Chan & Tang, 2024). Tanpa pemahaman yang baik tentang cara kerja AI, guru akan kesulitan menilai validitas *Prompt Log* siswa. Kedua, sekolah perlu menyepakati kebijakan etika yang jelas agar siswa tidak bingung mengenai batasan penggunaan AI. Ke depannya, pengembangan rubrik ini perlu diuji secara empiris di berbagai sekolah untuk melihat efektivitasnya dalam meningkatkan keterampilan bernalar kritis siswa secara nyata.

Daftar Pustaka

- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a complex AI future. *British Journal of Educational Technology*, 54, 1160–1173. <https://doi.org/10.1111/bjet.13337>.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Longman.
- Chan, K. K. W., & Tang, W. K. W. (2024). Evaluating English teachers' artificial intelligence readiness and training needs with a TPACK-based model. *World Journal of English Language*, 15(1), 129-145. URL: <https://doi.org/10.5430/wjel.v15n1p129>.
- Chen, Y., et al. (2025). Effects of generative artificial intelligence on cognitive effort and task performance: Study protocol for a randomized controlled experiment among college students. *Trials*, 26(244). <https://doi.org/10.1186/s13063-025-08950-3>.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi. (2024). *Panduan penggunaan Generative Artificial Intelligence (GenAI) pada pembelajaran di perguruan tinggi*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Eaton, S. E. (2025). Global trends in education: Artificial intelligence, postplagiarism, and future-focused learning for 2025 and beyond. *International Journal for Educational Integrity*, 21(12). <https://doi.org/10.1007/s40979-025-00187-6>.
- Feng, H., Li, K., & Zhang, L. J. (2024). What does AI bring to second language writing? A systematic review (2014–2024). *Language Learning & Technology*, 29(1), 1–27. <https://hdl.handle.net/10125/73629>.
- Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5–31. <https://hdl.handle.net/10125/73570>.
- Gregory, R. J. (2015). *Psychological testing: History, principles, and applications* (7th ed.). Pearson Education.
- Jin, F., Lin, C.-H., & Lai, C. (2024). Modeling AI-assisted writing: How self-regulated learning influences writing outcomes. *Computers in Human Behavior*, 165, 108538. <https://doi.org/10.1016/j.chb.2024.108538>.
- Jin, H. (2025). When AI meets source use: Exploring ChatGPT's potential in L2 summary writing assessment. *System*, 133, 103737. <https://doi.org/10.1016/j.system.2025.103737>.
- Jin, T., Su, Y., & Lei, J. (2020). Exploring the blended learning design for argumentative

- writing. *Language Learning & Technology*, 24(2), 23–34. <http://hdl.handle.net/10125/44720>.
- 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>.
- Kubiszyn, T., & Borich, G. (2013). *Educational testing and measurement: Classroom application and practice* (10th ed.). John Wiley & Sons.
- Lee, H., & Lee, J. H. (2024). The effects of AI-guided individualized language learning: A meta-analysis. *Language Learning & Technology*, 28(2), 134–162. <https://doi.org/10.64152/10125/73575>.
- Henestrosa, A. L., & Kimmerle, J. (2024). The effects of assumed AI vs. human authorship on the perception of a GPT-generated text. *Journalism and Media*, 5, 1085–1097. <https://doi.org/10.3390/journalmedia5030069>.
- Liu, J. Q. J., Hui, K. T. K., Zoubi, F. A., Zhou, Z. X., Samartzis, D., Yu, C. C. H., ... & Wong, A. Y. L. (2024). The great detectives: Humans versus AI detectors in catching large language model-generated medical writing. *International Journal for Educational Integrity*, 20(8). <https://doi.org/10.1007/s40979-024-00155-6>.
- Oppenheim, A. N. (1992). Questionnaire design, interviewing and attitude measurement. Continuum.
- Plaatjies, B. O., & van Wyk, M. M. (2025). Prompt literacy as an enhancer of students' academic writing in higher education institutions: A systematic literature review. *Journal of Teaching and Learning*, 19(4), 114–134. <https://doi.org/10.22329/jtl.v19i4.10017>.
- Pratiwi, N. K., Yulianto, B., Mintowati, M., Supratno, H., Sodiq, S., & Mulyono, M. (2024). Persepsi mahasiswa terhadap penggunaan ChatGPT: Peluang dan tantangan bagi pembelajaran Bahasa Indonesia sebagai mata kuliah wajib pada kurikulum perguruan tinggi. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, 10(3), 2727–2742. <https://doi.org/10.30605/onoma.v10i3.3931>.
- Shi, Z., Liu, F., Lai, C., & Jin, T. (2024). Enhancing the use of evidence in argumentative writing through collaborative processing of content-based automated writing evaluation feedback. *Language Learning & Technology*, 26(2), 106–128. <https://doi.org/10.125/73481>.
- Toulmin, S. E. (2003). *The uses of argument* (Updated ed.). Cambridge University Press.
- Turková, K., Krásničan, V., Prázová, I., Turčinek, P., & Foltýnek, T. (2025). Adapting to the future: The use of AI tools and applications in university education and a call for transparent rules and guidelines. *International Journal for Educational Integrity*, 21(29). <https://doi.org/10.1007/s40979-025-00203-9>.
- Wang, Q., & Newell, G. E. (2025). Teaching and learning argumentative writing as critical thinking in an EFL composition classroom. *Learning, Culture and Social Interaction*, 54, 100891. <https://doi.org/10.1016/j.lcsi.2025.100891>.
- Warschauer, M., & Xu, Y. (2024). Generative AI for Language learning: Entering a new era. *Language Learning & Technology*, 28(2), 1–4. <http://hdl.handle.net/10125/73569>.
- Woo, D. J., Susanto, H., Yeung, C. H., Guo, K., & Fung, A. K. Y. (2024). Exploring AI-Generated text in student writing: How does AI help? *Language Learning & Technology*, 28(2), 183–209. <https://hdl.handle.net/10125/73577>.
- Zhang, R., Zou, D., Xie, H., & Wang, F. L. (2025). Engagements with GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic and their impacts. *Computers and Education: Artificial Intelligence*, 9, 100489. <https://doi.org/10.1016/j.caeai.2025.100489>.
- Zhang, Y., Qi, W., Xia, C., Sun, H., & Chen, L. (2025). Exploring the effect of cooperative learning on senior high school students' critical thinking in EFL writing: An intervention study. *Thinking Skills and Creativity*, 56, 101765. <https://doi.org/10.1016/j.tsc.2025.101765>.
- Zhou, S., & Du Preez, G. D. (2025). Bringing Cinderella into spotlight: GenAI-assisted grammar acquisition in academic writing. *Language Learning & Technology*, 29(1), 1–25. <https://doi.org/10.64152/10125/73639>.