PERILAKU VERIFIKASI MAHASISWA TERHADAP REKOMENDASI *GENERATIVE AI* DALAM KEPUTUSAN AKADEMIK

Students' Verification Behaviors toward Generative AI Recommendations in Academic Decision-Making

Ika Safitri Windiarti [1)] Agung Prabowo [2)] Fathimah Azzahra [3)]
IImu Komputer [1)] - Universitas Muhammadiyah Palangkaraya
Sistem Informasi [2)] - STMIK Palangka Raya
Bachelor of Science (Multimedia) [3)] - Universiti Utara Malaysia

ABSTRAK

Generative Artificial Intelligence (GenAI) semakin digunakan mahasiswa untuk memperoleh rekomendasi dalam keputusan akademik, tetapi keluaran yang meyakinkan tidak selalu akurat atau didukung sumber yang sahih. Penelitian ini bertujuan memetakan konteks penggunaan GenAI, tindakan verifikasi, dan keputusan mahasiswa setelah verifikasi. Survei deskriptif potong lintang dilakukan secara daring pada 6–18 Oktober 2026. Dari 167 respons masuk, 147 respons memenuhi kriteria persetujuan dan penggunaan GenAI dalam tiga bulan terakhir. Data dianalisis menggunakan frekuensi, persentase, mean, simpangan baku, median, dan analisis sensitivitas. ChatGPT menjadi platform utama (59,2%). Rekomendasi paling sering digunakan untuk pemilihan sumber (24,5%) dan penyelesaian masalah (21,1%); 80,3% keputusan dilakukan dalam waktu kurang dari dua jam. Responden melaporkan rata-rata 6,96 dari sembilan tindakan verifikasi. Tindakan terbanyak ialah menilai kesesuaian dengan konteks (93,6%), membaca sumber asli (91,9%), meminta alasan (91,7%), dan mengajukan ulang pertanyaan (91,4%). Setelah pemeriksaan, 57,8% menyesuaikan rekomendasi dan 19,7% hanya menggunakan bagian yang sesuai, sedangkan 10,9% menggunakannya tanpa perubahan. Hasil menunjukkan kecenderungan verifikasi yang luas, tetapi belum membuktikan ketepatan keputusan atau warranted reliance karena kualitas keluaran dan keputusan akhir tidak dinilai secara objektif. Perguruan tinggi perlu menjadikan pemeriksaan sumber, pembandingan independen, dan dokumentasi alasan keputusan sebagai bagian eksplisit dari literasi AI.

Kata kunci: generative AI; verifikasi informasi; keputusan akademik; mahasiswa; warranted reliance

ABSTRACT

Generative Artificial Intelligence (GenAI) is increasingly used by students to obtain recommendations for academic decisions, although persuasive outputs are not necessarily accurate or supported by valid sources. This study mapped the contexts of GenAI use, students' verification actions, and their decisions after verification. A cross-sectional descriptive online survey was conducted from 6 to 18 October 2026. Of 167 submissions, 147 met the consent and recent-use eligibility criteria. Frequencies, percentages, means, standard deviations, medians, and a sensitivity analysis were used. ChatGPT was the main platform (59.2%). Recommendations were most often used for selecting sources (24.5%) and solving problems (21.1%), while 80.3% of decisions had to be made within two hours. Respondents reported a mean of 6.96 out of nine verification actions. The most frequent actions were assessing contextual fit (93.6%), reading original sources (91.9%), requesting reasons (91.7%), and re-asking the AI (91.4%). Following verification, 57.8% modified the recommendation and 19.7% used only suitable parts; 10.9% used it unchanged. These findings indicate extensive self-reported verification, but do not establish decision accuracy or warranted reliance because output quality and final decisions were not objectively assessed. Higher education should explicitly require source checking, independent comparison, and documentation of decision rationales as components of AI literacy.

Keywords: *generative AI; information verification; academic decision-making; university students; warranted reliance*

Pendahuluan

Generative Artificial Intelligence (GenAI) telah berkembang dari alat pembuat teks menjadi mitra yang memberi saran, alternatif, dan alasan untuk berbagai aktivitas akademik. Mahasiswa memanfaatkannya untuk memahami materi, menemukan sumber, menyusun langkah tugas, dan memecahkan masalah. Manfaat tersebut berjalan bersama risiko keluaran yang tidak akurat, referensi yang tidak dapat dilacak, bias, dan ketergantungan berlebihan. Karena itu, pemanfaatan GenAI di pendidikan perlu menempatkan

penilaian kritis dan pemeriksaan fakta sebagai kompetensi inti, bukan sekadar kemampuan menyusun prompt (Kasneci et al., 2023; Tlili et al., 2023).

Literatur mutakhir menunjukkan bahwa mahasiswa menginginkan ketepatan ilmiah, sitasi yang dapat diandalkan, transparansi, serta pencegahan halusinasi pada alat AI (von Garrel & Mayer, 2024). Niat untuk memverifikasi informasi juga berkaitan dengan penggunaan ChatGPT secara bertanggung jawab (Acosta-Enriquez et al., 2024). Namun, laporan diri tentang penggunaan AI belum tentu setara dengan kemampuan mengevaluasi keluaran secara mendalam. Perbedaan antara klaim mahasiswa dan penilaian asesor terhadap sensemaking kritis masih dapat terjadi (Fischer et al., 2024).

Dalam keputusan berbantuan AI, tujuan yang relevan bukan menolak atau menerima saran secara otomatis, melainkan warranted reliance: tingkat ketergantungan yang layak berdasarkan mutu rekomendasi, bukti, konteks, dan risiko keputusan. Saran AI berkualitas dapat membantu keputusan, sedangkan saran berkualitas rendah dapat merugikan; kepercayaan pengguna turut memengaruhi konsultasi dan adopsi saran (Grawitch et al., 2026). Penjelasan juga perlu diperlakukan secara hati-hati karena lebih banyak penjelasan tidak selalu meningkatkan penilaian pengguna dan kebutuhannya bergantung pada tugas serta risiko (Hauptman et al., 2024).

Kajian tentang GenAI di pendidikan banyak menyoroti penerimaan, manfaat, integritas akademik, dan sikap. Bukti rinci mengenai apa yang benar-benar dilakukan mahasiswa untuk memeriksa satu rekomendasi akademik terbaru, serta keputusan yang mereka ambil sesudahnya, masih terbatas. Penelitian ini mengisi celah tersebut melalui dua pertanyaan: (1) dalam konteks apa mahasiswa menggunakan rekomendasi GenAI untuk keputusan akademik; dan (2) bagaimana mahasiswa memverifikasi rekomendasi dan menentukan keputusan akhir?

Metodologi Penelitian

Desain dan Responden

Penelitian menggunakan desain survei kuantitatif deskriptif potong lintang. Kuesioner daring diisi pada 6–18 Oktober 2026. Kriteria inklusi ialah memberikan persetujuan berpartisipasi dan pernah memakai GenAI untuk keputusan akademik dalam tiga bulan terakhir. Dari 167 respons masuk, dua tidak memberikan persetujuan dan 18 tidak memenuhi kriteria penggunaan terkini, sehingga sampel analisis utama berjumlah 147. Sampel bersifat nonprobabilitas berbasis respons; hasil tidak ditujukan untuk mengestimasi prevalensi seluruh mahasiswa Indonesia.

Instrumen dan Analisis

Instrumen meminta responden mengingat satu pengalaman penggunaan GenAI yang paling baru. Bagian konteks mengukur platform, tujuan keputusan, bentuk rekomendasi, waktu yang tersedia, pentingnya keputusan, konsekuensi kesalahan, dan pengaruh AI. Bagian perilaku mencakup sembilan tindakan verifikasi dengan pilihan Ya, Tidak, atau Tidak berlaku. Bagian berikutnya merekam tindakan pertama, waktu verifikasi, jumlah sumber, hasil pemeriksaan, keputusan akhir, dan delapan kondisi yang memengaruhi pemeriksaan pada skala 1–5. Semua butir perilaku dirumuskan positif; pilihan Tidak berlaku tidak dihitung sebagai tindakan yang dilakukan.

Analisis menggunakan frekuensi, persentase, mean, simpangan baku, dan median. Persentase platform dapat melebihi 100% karena responden dapat memilih lebih dari satu. Analisis sensitivitas mengeluarkan 25 respons yang memiliki penanda konsistensi logis, menghasilkan 122 respons. Data sumber dipertahankan tanpa perubahan, dan identitas analisis dibuat anonim. Persetujuan berpartisipasi diperoleh secara elektronik. Penelitian ini tidak menguji hubungan kausal maupun menilai kebenaran rekomendasi menggunakan kunci jawaban atau ahli.

Hasil dan Pembahasan

Profil dan Konteks Penggunaan

Tabel 1. Profil responden (N = 147)

Karakteristik	n	%
Sarjana/S1	146	99,3
Semester 1–2	61	41,5
Semester 3–4	46	31,3
Semester 5–6	37	25,2
Semester 7–8	3	2,0
Ekonomi/bisnis	63	42,9

Karakteristik	n	%
Ilmu komputer/TI	44	29,9
Pendidikan	40	27,2

Hampir seluruh responden adalah mahasiswa sarjana. Komposisi didominasi semester awal dan rumpun ekonomi/bisnis. Oleh sebab itu, temuan terutama menggambarkan kelompok yang terjangkau dalam survei dan tidak boleh digeneralisasi tanpa kehati-hatian.

Tabel 2. Konteks keputusan akademik

Indikator utama	n	%
Platform: ChatGPT	87	59,2
Platform: Gemini	30	20,4
Platform: Claude	23	15,6
Memilih sumber/referensi	36	24,5
Menyelesaikan soal/masalah	31	21,1
Memilih topik	26	17,7
Menentukan langkah tugas	26	17,7
Waktu < 2 jam	118	80,3
AI memberi sumber	115	78,2

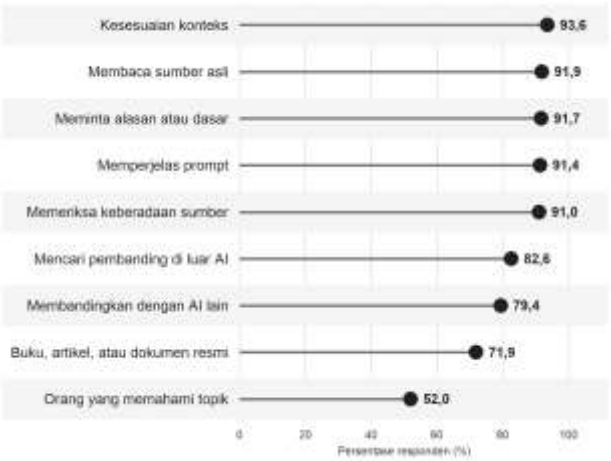
ChatGPT menjadi platform utama (59,2%). Rekomendasi paling banyak dipakai untuk memilih sumber (24,5%) dan menyelesaikan masalah (21,1%). Sebanyak 80,3% responden harus bertindak dalam waktu kurang dari dua jam. Statistik deskriptif variabel utama disajikan pada Tabel 3.

Tabel 3. Statistik deskriptif variabel utama (N = 147)

Variabel	M	SD	Md	Min–maks
Pentingnya keputusan	3,72	0,73	4	1–5
Konsekuensi kesalahan	3,14	0,86	3	1–5
Pengaruh AI	3,37	0,92	3	1–5
Jumlah tindakan	6,96	2,20	8	0–9

Nilai pentingnya keputusan, konsekuensi jika keliru, dan pengaruh AI diukur pada skala 1–5. Ketiga mean menunjukkan bahwa keputusan dipersepsikan cukup penting dan rekomendasi AI cukup memengaruhi keputusan akhir. Namun, data deskriptif ini tidak membuktikan bahwa tekanan waktu atau persepsi risiko menyebabkan reliance yang lebih tinggi.

Pola Verifikasi



Gambar 1. Persentase responden yang melakukan setiap tindakan verifikasi

Responden melaporkan rata-rata 6,96 dari sembilan tindakan (SD = 2,20; median = 8). Angka tersebut menunjukkan keluasan aktivitas, bukan skor kompetensi yang telah tervalidasi. Tindakan paling umum menggabungkan penilaian konteks, pembacaan sumber, dan dialog lanjutan dengan AI. Temuan ini selaras dengan Acosta-Enriquez et al. (2024), tetapi mengulang pertanyaan

atau meminta alasan kepada sistem yang sama belum menghasilkan bukti independen. Penjelasan yang lancar dapat membantu memahami asumsi, tetapi tidak menggantikan pemeriksaan sumber.

Tindakan pertama responden paling sering ialah membuka sumber (55,8%), kemudian mencari pembanding eksternal (19,0%). Sebanyak 38,1% menggunakan 5–15 menit dan 31,3% memeriksa dua sumber; 10,2% tidak memeriksa sumber. Aktivitas terstruktur yang meminta mahasiswa mengevaluasi teks AI memakai literatur ilmiah dilaporkan dapat mengembangkan evaluasi kritis (Oates & Johnson, 2025). Karena itu, verifikasi sebaiknya menjadi langkah tugas yang dapat diamati, bukan anjuran umum.

Verifikasi sebagai Berpikir Kritis dan Kewaspadaan Epistemik

Secara teoretis, verifikasi merupakan penerapan berpikir kritis karena mahasiswa tidak cukup memahami isi rekomendasi, tetapi perlu menilai kredibilitas premis, relevansi bukti, konsistensi alasan, dan kelayakan kesimpulan. Ennis (2018) menempatkan penilaian alasan dan pengambilan keputusan mengenai apa yang layak dipercaya atau dilakukan sebagai inti berpikir kritis. Dalam konteks GenAI, proses tersebut juga dapat dibaca sebagai kewaspadaan epistemik: perhatian tidak hanya diarahkan pada isi klaim, tetapi juga pada asal informasi, kompetensi sumber, dan kemungkinan bahwa bentuk bahasa yang meyakinkan menutupi ketidakpastian. Dengan demikian, membuka sumber asli dan mencari pembanding eksternal memiliki bobot epistemik lebih kuat daripada sekadar meminta sistem yang sama mengulangi atau menjelaskan jawabannya.

Hasil penelitian Indonesia memperkuat interpretasi tersebut. Rizaldi et al. (2024) menemukan bahwa mahasiswa dengan kemampuan berpikir kritis tinggi cenderung mempertanyakan kredibilitas informasi ChatGPT dan menyusun kesimpulan berdasarkan pemeriksaan, sedangkan kelompok dengan kemampuan lebih rendah lebih mudah menerima dan menyalin jawaban. Sabrina et al. (2024) juga melaporkan dampak ChatGPT yang bervariasi terhadap berpikir kritis dan pemecahan masalah. Artinya, ketersediaan AI tidak otomatis meningkatkan atau menurunkan berpikir kritis; hasilnya bergantung pada cara mahasiswa berinteraksi, kriteria evaluasi yang digunakan, dan tuntutan tugas.

Temuan bahwa 91,9% responden membaca sumber asli terlihat positif, tetapi laporan diri belum menunjukkan kualitas pembacaan tersebut. Pemeriksaan dapat berhenti pada keberadaan judul atau tautan, padahal verifikasi substantif memerlukan penilaian apakah sumber benar-benar mendukung klaim, apakah metode sumber memadai, dan apakah kesimpulan AI melampaui bukti. Oleh sebab itu, penelitian ini membedakan kuantitas tindakan dari kualitas verifikasi. Perbedaan ini penting karena mahasiswa dapat melakukan banyak langkah secara prosedural tanpa menemukan kesalahan yang relevan.

Verifikasi sebagai Regulasi Diri dalam Belajar

Urutan meminta alasan, mengajukan ulang pertanyaan, membaca sumber, dan menyesuaikan keputusan dapat dijelaskan sebagai siklus regulasi diri: mahasiswa merencanakan kebutuhan informasi, memantau kecukupan jawaban, lalu merefleksikan apakah rekomendasi layak digunakan. Ng et al. (2024) menunjukkan bahwa chatbot generatif dapat mendukung pengetahuan, keterlibatan, motivasi, dan perilaku regulasi diri ketika ditempatkan dalam desain pembelajaran yang memberikan dukungan adaptif. Dengan sudut pandang ini, nilai pendidikan GenAI tidak terletak pada kecepatan menghasilkan jawaban, tetapi pada kemampuannya memicu monitoring metakognitif dan revisi keputusan.

Namun, waktu yang singkat dapat menggeser regulasi diri menjadi jalan pintas kognitif. Sebanyak 80,3% responden mengambil keputusan dalam waktu kurang dari dua jam, sementara pemeriksaan paling sering berlangsung 5–15 menit. Pola ini dapat menjelaskan mengapa tindakan yang mudah dan segera—misalnya bertanya ulang kepada AI—sangat populer. Kajian Zhai et al. (2024) memperingatkan bahwa ketergantungan berlebihan pada sistem dialog AI dapat mengurangi keterlibatan dalam analisis mandiri ketika efisiensi menggantikan evaluasi. Data penelitian ini tidak membuktikan penurunan kemampuan kognitif, tetapi menunjukkan perlunya prosedur verifikasi ringkas yang tetap menuntut bukti eksternal.

Hasil Pemeriksaan dan Keputusan Akhir

Tabel 4. Hasil verifikasi dan keputusan akhir

Kategori	%
Rekomendasi didukung	44,2
Didukung sebagian/perlu penyesuaian	40,8
Menyesuaikan rekomendasi	57,8
Menggunakan sebagian	19,7
Menggunakan tanpa perubahan	10,9
Memilih alternatif	9,5
Menolak rekomendasi	2,0



Gambar 2. Keputusan akhir setelah proses verifikasi

Mayoritas responden mengubah rekomendasi atau memakai hanya bagian yang dinilai sesuai. Distribusi ini menunjukkan seleksi setelah pemeriksaan, tetapi perubahan tidak selalu benar dan penerimaan tanpa perubahan tidak selalu salah. Warranted reliance baru dapat dinilai apabila respons pengguna dibandingkan dengan mutu saran dan bukti objektif. Hal ini sejalan dengan eksperimen Grawitch et al. (2026), yang menunjukkan bahwa dampak saran AI bergantung pada kualitas saran dan bahwa kepercayaan memengaruhi adopsi.

Kalibrasi Ketergantungan pada Rekomendasi AI

Secara teoretis, keputusan yang baik menuntut kalibrasi, bukan skeptisisme mutlak. Ketergantungan dikatakan tepat apabila pengguna menerima saran yang benar dan menolak atau memperbaiki saran yang keliru. Schemmer et al. (2023) merumuskan appropriate reliance sebagai kesesuaian antara ketepatan saran AI dan keputusan pengguna pada setiap kasus. Konsep ini menjelaskan mengapa persentase penyesuaian yang tinggi tidak dapat langsung disebut keberhasilan: tanpa tolok ukur kebenaran, penyesuaian dapat memperbaiki ataupun justru merusak rekomendasi awal.

Perbedaan antara penjelasan dan bukti juga penting bagi kalibrasi. Meminta AI menjelaskan alasan dapat meningkatkan pemahaman mengenai alur rekomendasi, tetapi penjelasan yang koheren dapat menciptakan rasa memahami yang berlebihan. Hauptman et al. (2024) menunjukkan bahwa kebutuhan terhadap penjelasan bergantung pada kompleksitas, risiko, dan kebutuhan audit, sedangkan Schemmer et al. (2023) menunjukkan bahwa penjelasan tidak dengan sendirinya menjamin reliance yang tepat. Karena itu, penjelasan internal perlu dipasangkan dengan validasi eksternal dan indikator ketidakpastian.

Dalam konteks Indonesia, Abdilah (2023) menegaskan tanggung jawab manusia untuk meninjau keluaran ChatGPT secara kritis karena sistem dapat mencampuradukkan fakta dan fiksi serta memberikan saran yang tidak konsisten. Hasil survei ini memberikan gambaran perilakunya: mahasiswa cenderung menyesuaikan atau menggunakan sebagian rekomendasi setelah pemeriksaan. Akan tetapi, 10,9% tetap menggunakan rekomendasi tanpa perubahan dan 10,2% tidak memeriksa sumber. Kelompok ini penting bagi intervensi pendidikan karena penerimaan langsung dapat berasal dari rekomendasi yang memang benar, tetapi juga dapat menunjukkan kepercayaan yang belum terkalibrasi.

Kondisi yang Memengaruhi dan Implikasi

Pada skala 1–5, kondisi dengan nilai tertinggi adalah pentingnya keputusan (M = 3,56), diikuti ketersediaan sumber dari AI dan pemahaman terhadap topik (masing-masing M = 3,51), serta kemudahan akses sumber (M = 3,50). Ketersediaan waktu memperoleh nilai terendah, meskipun tetap tinggi (M = 3,29). Temuan tersebut mendukung prosedur verifikasi yang proporsional terhadap risiko: keputusan lebih penting memerlukan sumber yang lebih dapat ditelusuri, pembanding independen, dan alasan keputusan yang terdokumentasi.

Implikasi bagi Literasi AI dan Desain Pembelajaran

Literasi AI yang dibutuhkan mahasiswa melampaui keterampilan operasional menggunakan platform. Literasi tersebut mencakup kemampuan mengenali batas pengetahuan model, membedakan penjelasan dari bukti, menilai otoritas sumber, membandingkan informasi secara lateral, serta mempertanggungjawabkan keputusan akhir. Fischer et al. (2024) menemukan bahwa klaim mahasiswa mengenai penggunaan GenAI belum tentu sejalan dengan kedalaman sensemaking yang dinilai oleh asesor. Karena itu, evaluasi pembelajaran sebaiknya tidak hanya menanyakan apakah mahasiswa melakukan verifikasi, tetapi juga meminta jejak bukti yang menunjukkan bagaimana verifikasi mengubah penilaian mereka.

Studi lokal menunjukkan bahwa integrasi ChatGPT yang disertai arahan dapat mendukung pembelajaran. Rizaldi et al. (2024) memperoleh skor berpikir kritis lebih tinggi pada kelas berbantuan ChatGPT daripada kelas kontrol, sedangkan Zein et al. (2024) menunjukkan potensi ChatGPT sebagai tutor virtual sekaligus menemukan jawaban yang tidak konsisten dan kebutuhan verifikasi

mandiri. Kedua temuan ini mendukung pendekatan pedagogis bersyarat: AI dapat menjadi scaffolding ketika tugas mewajibkan pertanyaan, pembuktian, dan refleksi, tetapi dapat menjadi jalan pintas apabila jawaban akhir lebih dihargai daripada proses evaluasi.

Perguruan tinggi dapat mengoperasionalkan prinsip tersebut melalui tugas pemeriksaan klaim. Mahasiswa diminta memilih satu klaim AI, melacak sumber aslinya, membandingkannya dengan sumber independen, menilai kesesuaian metode dan konteks, kemudian mendokumentasikan alasan menerima, merevisi, atau menolak rekomendasi. Model penugasan ini konsisten dengan Oates dan Johnson (2025), yang memperlihatkan manfaat evaluasi teks AI menggunakan literatur ilmiah, sekaligus menjawab seruan etis Abdilah (2023) agar manusia tetap bertanggung jawab atas keluaran yang digunakan.

Arah tersebut juga konsisten dengan kajian Alam dan Windiarti (2025), yang menempatkan AI sebagai pendukung pembelajaran adaptif, personal, dan interaktif, tetapi tetap memerlukan mitigasi risiko serta kebijakan pendidikan yang bertanggung jawab. Pada tingkat implementasi lokal, Windiarti et al. (2023) menunjukkan bahwa penerapan Generative AI memerlukan peningkatan pemahaman dan keterampilan praktis tenaga pendidikan. Dengan demikian, penguatan verifikasi mahasiswa perlu disertai kesiapan dosen dan institusi untuk merancang aktivitas, rubrik, serta dukungan belajar yang sesuai.

Analisis sensitivitas pada 122 respons menghasilkan pola yang stabil: perbedaan maksimum persentase tindakan verifikasi 5,0 poin persentase dan perbedaan maksimum mean kondisi 0,03. Meski demikian, laporan diri rentan terhadap kesalahan ingatan dan social desirability bias. Keterbatasan lain ialah sampel nonprobabilitas, dominasi mahasiswa sarjana dan semester awal, desain potong lintang, serta tidak adanya penilaian objektif atas mutu sumber, keluaran AI, dan keputusan akhir.

Simpulan Dan Saran

Penelitian ini menyimpulkan bahwa mahasiswa tidak selalu menerima rekomendasi Generative AI secara pasif. Pada satu pengalaman akademik terbaru, responden melaporkan rata-rata 6,96 dari sembilan tindakan verifikasi, dan sebagian besar menyesuaikan rekomendasi (57,8%) atau hanya menggunakan bagian yang dianggap sesuai (19,7%). Penilaian kesesuaian dengan konteks, pembacaan sumber asli, permintaan alasan, dan pengajuan ulang pertanyaan menjadi tindakan yang paling sering dilaporkan. Pola tersebut menunjukkan adanya upaya evaluatif setelah menerima rekomendasi AI, terutama pada keputusan yang dipersepsikan penting dan harus dilakukan dalam waktu relatif singkat.

Meskipun demikian, keluasan tindakan tidak dapat disamakan dengan kualitas verifikasi atau ketepatan keputusan. Meminta penjelasan dan mengulang pertanyaan kepada sistem yang sama dapat membantu memahami alasan, tetapi tidak menyediakan bukti independen. Demikian pula, perubahan rekomendasi belum tentu memperbaiki keputusan apabila sumber pembanding tidak kredibel atau pengetahuan awal pengguna tidak memadai. Kontribusi penelitian ini terletak pada pemetaan hubungan antara konteks penggunaan, tindakan pemeriksaan, hasil verifikasi, dan keputusan akhir. Namun, karena menggunakan laporan diri, sampel nonprobabilitas, dan tidak memiliki tolok ukur objektif, penelitian ini belum dapat menyimpulkan bahwa ketergantungan mahasiswa pada AI telah terkalibrasi dengan benar atau mewakili seluruh populasi mahasiswa.

Perguruan tinggi disarankan menjadikan verifikasi sebagai aktivitas akademik yang dapat diamati dan dinilai. Prosedur minimum dapat mencakup pemeriksaan keberadaan serta identitas bibliografis sumber, pembacaan sumber asli untuk menilai dukungan terhadap klaim, perbandingan dengan sedikitnya satu sumber independen, dan pencatatan alasan menerima, mengubah, menggunakan sebagian, atau menolak rekomendasi. Dosen dapat meminta lampiran jejak verifikasi berupa prompt utama, sumber yang diperiksa, ketidaksesuaian yang ditemukan, dan alasan keputusan akhir. Pendekatan tersebut perlu didukung pelatihan literasi AI bagi mahasiswa dan dosen, rubrik penilaian proses, serta kebijakan penggunaan AI yang membedakan tingkat pemeriksaan berdasarkan pentingnya dan risiko keputusan. Rekomendasi ini sejalan dengan kebutuhan peningkatan kompetensi praktis tenaga pendidikan yang dikemukakan Windiarti et al. (2023) serta pengembangan pembelajaran AI yang adaptif dan bertanggung jawab (Alam & Windiarti, 2025).

Penelitian berikutnya disarankan menggunakan tugas keputusan yang memiliki jawaban acuan atau penilaian ahli agar kualitas verifikasi dan ketepatan keputusan dapat dibandingkan secara langsung. Desain eksperimen dapat memanipulasi kualitas rekomendasi, validitas sumber, tingkat penjelasan, tekanan waktu, pengetahuan awal, dan konsekuensi kesalahan. Pengukuran tidak hanya perlu mencatat jumlah tindakan, tetapi juga keberhasilan menemukan kesalahan, relevansi sumber pembanding, akurasi keputusan akhir, perubahan kepercayaan, dan waktu pemeriksaan. Sampel yang lebih beragam, desain longitudinal, serta kombinasi log interaksi, think-aloud, dan wawancara diperlukan untuk menjelaskan bagaimana kebiasaan verifikasi terbentuk dan apakah pelatihan menghasilkan perubahan yang bertahan.

Pustaka Acuan

Abdilah, Y. A. (2023). Tinjauan sistematis etika penggunaan ChatGPT di perguruan tinggi. Integralistik, 34(2), 69–78. https://doi.org/10.15294/integralistik.v34i2.50278

Acosta-Enriquez, B. G., Arbulú Ballesteros, M. A., Huamaní Jordan, O., López Roca, C., & Saavedra Tirado, K. (2024). Analysis of college students' attitudes toward the use of ChatGPT in their academic activities: Effect of intent to use, verification of information and responsible use. BMC Psychology, 12, 255. https://doi.org/10.1186/s40359-024-01764-z

Alam, T. H. I., & Windiarti, I. S. (2025). The future of artificial intelligence in interactive learning: Trends, challenges, opportunities. Engineering Proceedings, 84(1), 87. https://doi.org/10.3390/engproc2025084087

Ennis, R. H. (2018). Critical thinking across the curriculum: A vision. Topoi, 37(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4

Fischer, I., Sweeney, S., Lucas, M., & Gupta, N. (2024). Making sense of generative AI for assessments: Contrasting student claims and assessor evaluations. The International Journal of Management Education, 22(3), 101081. https://doi.org/10.1016/j.ijme.2024.101081

Grawitch, M. J., Winton, S. L., & Mudigonda, S. P. (2026). Who listens to ChatGPT and when should they? A two-study examination of AI-assisted decision making. Computers in Human Behavior: Artificial Humans, 7, 100283. https://doi.org/10.1016/j.chbah.2026.100283

Hauptman, A. I., Schelble, B. G., Duan, W., Flathmann, C., & McNeese, N. J. (2024). Understanding the influence of AI autonomy on AI explainability levels in human-AI teams using a mixed methods approach. Cognition, Technology & Work, 26, 435–455. https://doi.org/10.1007/s10111-024-00765-7

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & 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

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. British Journal of Educational Technology, 55, 1328–1353. https://doi.org/10.1111/bjet.13454

Oates, A., & Johnson, D. (2025). ChatGPT in the classroom: Evaluating its role in fostering critical evaluation skills. International Journal of Artificial Intelligence in Education, 35, 1793–1824. https://doi.org/10.1007/s40593-024-00452-8

Rizaldi, M., Sasalia, P., & Pancarita, P. (2024). Efektivitas ChatGPT untuk mendorong berpikir kritis mahasiswa dalam menyelesaikan masalah aljabar. EDU-MAT: Jurnal Pendidikan Matematika, 12(1), 168–181. https://doi.org/10.20527/edumat.v12i1.18668

Sabrina, E., Yulianti, R., Ambiyar, & Rizal, F. (2024). Mengeksplor dampak interaksi siswa dengan ChatGPT terhadap berpikir kritis dan pemecahan masalah. The Indonesian Journal of Computer Science, 13(3), 4875–4883. https://doi.org/10.33022/ijcs.v13i3.4073

Schemmer, M., Kühl, N., Benz, C., Bartos, A., & Satzger, G. (2023). Appropriate reliance on AI advice: Conceptualization and the effect of explanations. Proceedings of the 28th International Conference on Intelligent User Interfaces, 410–422. https://doi.org/10.1145/3581641.3584066

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10, 15. https://doi.org/10.1186/s40561-023-00237-x

von Garrel, J., & Mayer, J. (2024). Which features of AI-based tools are important for students? A choice-based conjoint analysis. Computers and Education: Artificial Intelligence, 7, 100311. https://doi.org/10.1016/j.caeai.2024.100311

Windiarti, I. S., Bahri, S., & Prabowo, A. (2023). Melangkah maju dengan teknologi Generative AI: Peningkatan kompetensi kepala sekolah SMP di Kota Palangkaraya. Parta: Jurnal Pengabdian Kepada Masyarakat, 4(1), 46–52. https://doi.org/10.38043/parta.v4i1.4344

Zein, S., Yaniawati, R. P., & Mudrikah, A. (2024). The role and evaluation of ChatGPT as a virtual tutor in improving students' creative and critical abilities reviewed from probing-prompting abilities. Kreano, Jurnal Matematika Kreatif-Inovatif, 15(2), 501–517. https://doi.org/10.15294/a2g5x690

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. Smart Learning Environments, 11, 28. https://doi.org/10.1186/s40561-024-00316-7