# PELATIHAN VERIFIKASI INFORMASI DAN PENCEGAHAN HALUSINASI *ARTIFICIAL INTELLIGENCE* (AI) BAGI MAHASISWA

**Tino Feri Efendi[1], Dewi Muliasari[2], M. Hasan Ma'ruf[3]**
[1,2,3]Institut Teknologi Bisnis AAS Indonesia
Email : tinoferi8@gmail.com

***Abstract***

*The emergence of generative artificial intelligence (AI) technologies such as ChatGPT, Claude, Gemini, and Copilot has transformed the learning landscape in higher education in a very short time. Informatics students are the group with the highest adoption rates, utilizing AI not only for academic writing but also for code generation and error detection. However, this technology has a structural limitation in the form of AI hallucinations, which is the tendency of models to generate information, data, citations, references, and even programming library names that appear convincing but lack factual basis. This community service activity aims to improve students' AI literacy, information verification skills, and academic ethics awareness. The activity, carried out as a two-day intensive training followed by four weeks of mentoring, was attended by 42 fourth- and sixth-semester Informatics students from the AAS Indonesia Institute of Business Technology. The implementation method included seven stages: needs analysis, pre-test, delivery of five material modules, demonstration of AI hallucinations on scientific references and program code, guided verification practice, focus group discussions, and a post-test with evaluation and reflection. The evaluation instruments included a 25-item knowledge test, a verification skills rubric, and a satisfaction questionnaire. The results of the activity showed an increase in the average knowledge score from 49.16 in the pre-test to 83.61 in the post-test, with an N-Gain of 0.68, in the moderate category. A paired-sample t-test showed a highly significant difference with a significance value less than 0.001 and a very large effect size. The ability to identify fabricated references increased from 26.2% to 90.5%, while the proportion of participants who received AI output without tracing decreased from 40.5% to 2.4%. The participant satisfaction rate reached 91.3%. The activity concluded that the training based on demonstrations of real-world errors combined with hands-on verification practices effectively developed students' critical thinking without rejecting the productive use of AI.*

***Keywords:*** *AI hallucinations; artificial intelligence literacy; information verification; academic integrity; informatics students*

## 1. PENDAHULUAN

Sejak diluncurkannya ChatGPT pada akhir 2022, kecerdasan buatan generatif berbasis model bahasa besar mengalami tingkat adopsi yang belum pernah terjadi pada teknologi digital mana pun sebelumnya. Dalam kurun kurang dari empat tahun, ekosistemnya berkembang menjadi lanskap yang padat, meliputi ChatGPT, Claude, Gemini, Copilot, serta berbagai model sumber

terbuka. Kasneci et al. (2023) mencatat bahwa kemampuan model bahasa besar menghasilkan teks yang koheren dan menyerupai tulisan manusia menempatkan teknologi ini pada posisi yang bersinggungan langsung dengan aktivitas inti pendidikan tinggi, yaitu membaca, menulis, dan berargumentasi. Berbeda dengan teknologi pendidikan sebelumnya yang bersifat pendukung, AI generatif masuk ke jantung proses kognitif akademik itu sendiri. Karakteristik teknis model bahasa besar menjelaskan sebagian besar persoalan yang dibahas dalam artikel ini. Model bekerja dengan memprediksi *token* berikutnya yang paling mungkin muncul berdasarkan pola statistik dari korpus teks berskala raksasa. Bender et al. (2021) menyebut sistem semacam ini sebagai *stochastic parrots*, yakni sistem yang menyusun rangkaian bahasa secara probabilistik tanpa representasi makna yang dapat diverifikasi. Implikasinya bersifat mendasar: model tidak membedakan pernyataan yang benar dari pernyataan yang sekadar terdengar benar, karena keduanya sama-sama merupakan rangkaian *token* berprobabilitas tinggi. Dwivedi et al. (2023) menegaskan bahwa kekeliruan memahami arsitektur ini melahirkan ekspektasi yang salah, yaitu memperlakukan model generatif sebagai mesin pencari fakta, padahal ia pada dasarnya adalah mesin penghasil bahasa.

Kecepatan adopsi di kalangan mahasiswa jauh melampaui kecepatan institusi merumuskan kebijakan. Chan dan Hu (2023), melalui survei terhadap 399 mahasiswa, menemukan bahwa mayoritas mahasiswa telah menggunakan AI generatif untuk curah gagasan, penyuntingan bahasa, dan pencarian ide, sekaligus mengungkapkan kekhawatiran mereka sendiri mengenai akurasi keluaran dan ketiadaan panduan institusional. Chan dan Lee (2023) menambahkan adanya kesenjangan generasional: mahasiswa menunjukkan penerimaan yang jauh lebih tinggi dibandingkan dosen, tetapi optimisme tersebut tidak disertai pemahaman teknis mengenai batas keandalan model. Kesenjangan antara antusiasme dan pemahaman inilah yang menjadi titik masuk kegiatan pengabdian ini. Bagi mahasiswa Informatika, persoalan tersebut memiliki dimensi tambahan. Selain menggunakan AI untuk keperluan penulisan akademik sebagaimana mahasiswa lain, mereka juga menggunakannya secara intensif untuk pembuatan kode, penelusuran kesalahan program, dan pemahaman dokumentasi teknis. Hal ini memperluas permukaan risiko: halusinasi tidak hanya berwujud referensi palsu, tetapi juga nama pustaka yang tidak pernah ada, fungsi *application programming interface* yang tidak terdapat dalam dokumentasi resmi, serta penjelasan algoritma yang keliru namun disajikan dengan sangat meyakinkan. Ironisnya, mahasiswa Informatika kerap merasa lebih kebal terhadap kesalahan AI justru karena kedekatan mereka dengan teknologi, padahal keakraban dengan alat tidak otomatis berarti pemahaman terhadap keterbatasannya.

Transformasi yang dibawa AI generatif tidak dapat direduksi menjadi persoalan kecurangan akademik semata. Farrokhnia et al. (2024), melalui analisis SWOT, mengidentifikasi kekuatan berupa kemampuan memberikan respons yang dipersonalisasi, mendukung pembelajaran mandiri, dan menurunkan hambatan bahasa bagi penutur non-Inggris. Namun transformasi tersebut memunculkan pergeseran pola kognitif yang perlu diwaspadai. Zhai et al. (2024), melalui tinjauan sistematis, menemukan bahwa penggunaan berlebihan berkorelasi dengan penurunan kemampuan berpikir analitis dan retensi pengetahuan jangka panjang. Hou et al. (2025) memperdalam temuan tersebut dengan menunjukkan bahwa disposisi berpikir kritis berfungsi sebagai variabel moderator:

mahasiswa dengan disposisi tinggi memperlakukan keluaran AI sebagai hipotesis yang harus diuji, sedangkan mahasiswa dengan disposisi rendah memperlakukannya sebagai jawaban final. Meta-analisis Wang dan Fan (2025) memberikan gambaran yang seimbang, yaitu bahwa pengaruh positif AI terhadap performa belajar cukup besar tetapi pengaruhnya terhadap kemampuan berpikir tingkat tinggi bersifat lebih kecil dan sangat bergantung pada desain instruksional yang menyertainya. Dengan kata lain, manfaat AI tidak muncul otomatis dari ketersediaan teknologi, melainkan dari cara pengintegrasiannya ke dalam praktik pedagogis.

Istilah *AI hallucination* merujuk pada keluaran model yang lancar dan meyakinkan tetapi tidak memiliki dasar faktual. Ji et al. (2023) mendefinisikannya sebagai keluaran yang tidak setia terhadap masukan atau tidak dapat diverifikasi terhadap pengetahuan dunia nyata. Huang et al. (2025) menyempurnakan taksonomi tersebut dengan membedakan halusinasi faktual dari halusinasi kesetiaan, sekaligus memetakan penyebabnya pada tahap data, pelatihan, dan inferensi. Sun et al. (2024) mengusulkan klasifikasi distorsi informasi yang mencakup fabrikasi entitas, distorsi hubungan kausal, dan pemalsuan atribusi. Emsley (2023) mengajukan kritik terminologis penting: istilah "halusinasi" berpotensi menyesatkan karena apa yang dihasilkan model bukanlah persepsi keliru melainkan fabrikasi dan falsifikasi, yaitu informasi yang dikarang lengkap dengan atribut yang membuatnya tampak sah. Bukti empiris mengenai skala persoalan ini terus bertambah. Alkaissi dan McFarlane (2023) mendokumentasikan referensi ilmiah yang sepenuhnya fiktif dalam penulisan bidang kedokteran. Athaluri et al. (2023) menemukan bahwa dari 178 referensi yang dihasilkan untuk proposal penelitian, sebagian besar tidak memiliki DOI yang valid. Walters dan Wilder (2023) melaporkan proporsi kutipan fabrikatif yang tinggi dengan pola kesalahan yang halus dan sulit dideteksi sekilas, sedangkan McGowan et al. (2023) menunjukkan bahwa fabrikasi kutipan terjadi secara spontan bahkan ketika model tidak diminta menghasilkan referensi. Persoalan ini tidak lagi terbatas pada tugas mahasiswa. Topaz et al. (2026), melalui audit terhadap lebih dari dua juta artikel biomedis, menemukan bahwa proporsi artikel yang memuat setidaknya satu referensi fiktif meningkat dari satu di antara 2.828 artikel pada 2023 menjadi satu di antara 458 artikel pada 2025. Temuan-temuan tersebut menunjukkan bahwa kemampuan verifikasi bukan keterampilan tambahan, melainkan kompetensi dasar yang harus dimiliki sejak jenjang sarjana.

Perdebatan mengenai integritas akademik pun telah bergeser dari pertanyaan apakah penggunaan AI merupakan pelanggaran menuju pertanyaan bagaimana penggunaannya dapat diatur secara transparan. Perkins (2023) berargumen bahwa definisi konvensional plagiarisme tidak sepenuhnya menangkap persoalan yang ditimbulkan model bahasa besar, karena teks yang dihasilkan bersifat orisinal secara leksikal namun tidak merepresentasikan proses berpikir penulis. Cotton et al. (2024) menekankan bahwa respons institusional yang hanya mengandalkan perangkat deteksi bersifat rapuh karena akurasinya terbatas dan berpotensi menimbulkan tuduhan keliru. Pendekatan yang lebih konstruktif ditawarkan Perkins et al. (2024) melalui *AI Assessment Scale*, yaitu kerangka bertingkat yang menetapkan sejauh mana AI boleh digunakan pada setiap jenis penugasan. Chan (2023) mengusulkan kerangka kebijakan dengan tiga dimensi, yaitu literasi pengguna, tata kelola, dan penyelarasan kurikulum, sedangkan Bittle dan El-Gayar (2025) menyimpulkan bahwa risiko ketidakjujuran akademik dapat dimitigasi melalui kombinasi

pendidikan literasi, desain penilaian yang tahan-AI, dan kebijakan yang jelas. Dalam konteks Indonesia, kerangka normatif tersedia melalui Undang-Undang Nomor 12 Tahun 2012 dan Peraturan Menteri Pendidikan Nasional Nomor 17 Tahun 2010, yang disusun sebelum AI generatif tersedia luas sehingga belum mengatur secara spesifik penggunaan AI dalam karya akademik. Kekosongan ini menempatkan tanggung jawab pembinaan pada tingkat program studi.

Literasi AI dipahami sebagai kompetensi yang memungkinkan seseorang mengevaluasi dan berkolaborasi secara kritis dengan sistem kecerdasan buatan. UNESCO (2023) menekankan pendekatan berpusat pada manusia dan menempatkan kemampuan penilaian kritis pengguna sebagai lapis pertahanan utama terhadap misinformasi mesin, sedangkan kerangka kompetensi UNESCO (2024) memerinci empat dimensi yang melampaui cakupan literasi digital konvensional. Klingbeil et al. (2024), melalui studi eksperimental, menunjukkan bahwa kepercayaan berlebihan terhadap rekomendasi AI menimbulkan biaya nyata berupa keputusan keliru, dan bahwa pengguna cenderung tidak memverifikasi ketika keluaran disajikan dengan gaya bahasa yang percaya diri. Wang et al. (2025) menemukan pola serupa, yaitu bahwa kepercayaan mahasiswa dibentuk lebih oleh kefasihan keluaran daripada oleh akurasinya.

Penelusuran terhadap publikasi dan laporan pengabdian dalam kurun 2023–2026 menunjukkan tiga pola. Pertama, kajian mengenai halusinasi AI didominasi penelitian dasar dan audit teknis yang kuat secara konseptual tetapi tidak dirancang untuk diterjemahkan menjadi intervensi pendidikan. Kedua, kajian integritas akademik berfokus pada tataran kebijakan institusional sehingga menghasilkan rekomendasi struktural yang belum menyediakan model pelatihan siap terap. Ketiga, kegiatan pengabdian bertema AI umumnya menitikberatkan pengenalan alat dan pelatihan *prompt* untuk meningkatkan produktivitas, bukan pengembangan kemampuan verifikasi. Orientasi produktivitas semacam ini berpotensi kontraproduktif karena mempercepat produksi keluaran yang belum terverifikasi, sebagaimana diperingatkan Wach et al. (2023) mengenai sisi gelap AI generatif. Kesenjangan yang teridentifikasi adalah belum tersedianya model pelatihan yang mengintegrasikan tiga unsur sekaligus, yaitu pemahaman konseptual mengenai mekanisme halusinasi, praktik verifikasi berbasis alat penelusuran nyata, dan internalisasi etika dalam bentuk protokol pengungkapan. Kegiatan ini dirancang untuk mengisi kesenjangan tersebut, dengan urgensi yang bertumpu pada tiga argumen: jeda waktu antara adopsi teknologi dan penyesuaian kurikulum, sifat kumulatif kesalahan yang tidak terdeteksi, serta nilai kompetensi verifikasi yang jauh melampaui masa studi sebagaimana ditegaskan OECD (2026).

## 2. PERMASALAHAN MITRA

Mitra kegiatan adalah mahasiswa Program Studi Informatika Institut Teknologi Bisnis AAS Indonesia semester empat dan semester enam. Pemetaan permasalahan dilakukan melalui observasi terhadap tugas mahasiswa selama dua semester, wawancara dengan lima dosen pengampu, dan survei pendahuluan daring terhadap 71 mahasiswa. Triangulasi ketiga sumber tersebut mengidentifikasi enam kelompok permasalahan.

Kepercayaan penuh terhadap jawaban AI**.** Survei pendahuluan menunjukkan 68% responden jarang atau tidak pernah meragukan jawaban AI apabila tersusun rapi dan menggunakan istilah teknis.

Sebagian menyatakan menganggap AI lebih andal daripada mesin pencari konvensional karena langsung memberikan jawaban. Sikap ini konsisten dengan temuan Klingbeil et al. (2024) mengenai *overreliance*, di mana kefasihan penyajian menggantikan bukti sebagai dasar kepercayaan.

a. Tidak melakukan verifikasi referensi
   Hanya 26% responden menyatakan pernah memeriksa keberadaan referensi yang diperoleh dari AI. Sebagian besar tidak mengetahui fungsi DOI, tidak pernah menggunakan Crossref, dan tidak memahami perbedaan mesin pencari umum dengan pangkalan data ilmiah. Mereka yang mengaku sudah memverifikasi umumnya hanya menyalin judul ke mesin pencari dan menganggap kemunculan hasil apa pun sebagai konfirmasi.
b. Penggunaan referensi fiktif dalam karya akademik
   Pemeriksaan acak terhadap 30 laporan tugas mata kuliah menemukan 12 laporan memuat sekurang-kurangnya satu referensi yang tidak dapat ditelusuri. Pola kesalahan sejalan dengan dokumentasi Walters dan Wilder (2023), yaitu nama penulis nyata dipasangkan dengan judul yang tidak pernah ada, jurnal yang benar-benar terbit tetapi dengan nomor volume tidak sesuai, serta DOI berformat benar yang tidak teregistrasi. Mayoritas mahasiswa yang menyertakannya tidak berniat curang; mereka sekadar tidak menyadari bahwa referensi dapat difabrikasi.
c. Penerimaan kode program tanpa pengujian.
   Permasalahan yang khas pada mitra ini adalah kebiasaan menyalin kode hasil AI langsung ke dalam proyek tanpa membaca ulang atau menguji. Dosen pengampu melaporkan temuan berulang berupa pemanggilan pustaka yang tidak ada dalam ekosistem paket resmi, penggunaan parameter fungsi yang tidak terdapat dalam dokumentasi versi yang dipakai, serta penjelasan kompleksitas algoritma yang keliru namun disalin apa adanya ke dalam laporan. Ketika kode gagal dijalankan, mahasiswa cenderung meminta perbaikan kepada AI berulang kali alih-alih membaca dokumentasi resmi atau pesan galat.
d. Pemahaman etika penggunaan AI yang terbatas
   Mahasiswa tidak memiliki acuan mengenai bentuk penggunaan yang dapat diterima. Sebagian menganggap seluruh penggunaan AI sebagai kecurangan sehingga menyembunyikannya, sebagian lain menganggap seluruhnya wajar sehingga menyerahkan pengerjaan sepenuhnya kepada AI. Kedua sikap ekstrem ini mencerminkan ketiadaan kerangka rujukan sebagaimana diusulkan Perkins et al. (2024). Tidak seorang pun responden mengetahui konsep pengungkapan penggunaan AI.
e. Belum mengenal konsep halusinasi AI dan keterbatasan merumuskan *prompt*
   Sebanyak 73,8% responden belum pernah mendengar istilah *AI hallucination*, dan ketika dijelaskan, sebagian besar mengira kesalahan AI hanya berupa informasi yang belum diperbarui. Ketiadaan model mental mengenai cara kerja probabilistik membuat mahasiswa tidak memiliki dasar untuk memprediksi kapan AI cenderung keliru, padahal pola kesalahan model bersifat sistematis dan dapat diantisipasi (Huang et al., 2025). Hal ini diperparah kebiasaan merumuskan *prompt* yang sangat singkat dan terbuka, yang justru

memaksimalkan ruang fabrikasi karena tidak memberikan batasan sumber maupun konteks. Giray (2023) menunjukkan bahwa rekayasa *prompt* yang cermat dapat menurunkan tingkat kesalahan secara berarti.

## 3. SOLUSI YANG DITAWARKAN

Berdasarkan pemetaan tersebut, tim memilih bentuk intervensi berupa pelatihan intensif yang dilanjutkan pendampingan terstruktur, dengan empat pertimbangan.

Pertama, karakteristik permasalahan bersifat kompetensial, bukan struktural. Persoalan utama mitra bukan ketiadaan akses teknologi, melainkan ketiadaan kemampuan mengevaluasi keluarannya. Persoalan kompetensi paling efektif diatasi melalui penyampaian pengetahuan, demonstrasi, dan latihan terbimbing secara berurutan; sosialisasi satu arah tidak memadai karena keterampilan verifikasi bersifat prosedural.

Kedua, kebutuhan akan pengalaman diskonfirmasi langsung. Kepercayaan yang tidak berdasar terhadap keluaran AI tidak runtuh oleh penjelasan verbal, melainkan oleh pengalaman menemukan kesalahan sendiri. Karena itu demonstrasi halusinasi dirancang sebagai inti pengalaman belajar, bukan ilustrasi tambahan, dan dilaksanakan dengan perangkat yang dioperasikan langsung oleh peserta.

Ketiga, kebutuhan akan pendampingan pascapelatihan. Zhai et al. (2024) menunjukkan bahwa ketergantungan pada AI merupakan pola perilaku yang terbentuk melalui pengulangan, sehingga penggantiannya juga menuntut pengulangan. Pelatihan karena itu dilengkapi pendampingan empat pekan melalui konsultasi terjadwal dan penugasan terapan.

Keempat, kesesuaian dengan kerangka kompetensi internasional. Rancangan materi diselaraskan dengan empat dimensi kompetensi AI dari UNESCO (2024) dan prinsip berpusat pada manusia dalam UNESCO (2023), sehingga hasilnya berpotensi direplikasi di institusi lain.

Solusi diwujudkan dalam lima komponen: penyusunan dan penyampaian lima modul terpadu; demonstrasi terstruktur halusinasi AI pada referensi ilmiah sekaligus pada kode program; praktik terbimbing verifikasi menggunakan DOI, Crossref, Google Scholar, dokumentasi resmi pustaka, dan Retraction Watch; penyusunan protokol etika bersama peserta; serta pembentukan komunitas literasi AI sebagai wadah keberlanjutan.

## 4. METODE PELAKSANAAN

**Lokasi dan waktu.** Kegiatan dilaksanakan di Laboratorium Komputer Kampus 2 Institut Teknologi Bisnis AAS Indonesia, Kartasura, Kabupaten Sukoharjo, Jawa Tengah. Laboratorium dipilih agar setiap peserta dapat mengoperasikan sendiri perangkat AI dan alat penelusuran, mengingat inti pelatihan terletak pada praktik langsung. Rangkaian kegiatan berlangsung sebelas pekan sejak awal Februari hingga pertengahan April 2026, dengan pelatihan inti pada 9 dan 10 Maret 2026 pukul 08.00–16.00 WIB. Rentang tersebut mencakup persiapan empat pekan, pelatihan inti dua hari, pendampingan empat pekan, serta evaluasi dan pelaporan tiga pekan.

**Peserta.** Peserta berjumlah 42 mahasiswa aktif Program Studi Informatika semester empat dan semester enam. Rekrutmen dilakukan melalui pengumuman terbuka dengan tiga kriteria inklusi:

berstatus mahasiswa aktif pada kedua semester tersebut, pernah menggunakan sekurang-kurangnya satu perangkat AI generatif untuk keperluan akademik, serta bersedia mengikuti seluruh rangkaian termasuk pendampingan. Dari 58 pendaftar, 42 orang memenuhi ketiga kriteria. Pemilihan semester empat dan enam didasarkan pada pertimbangan bahwa kedua angkatan tersebut telah menempuh mata kuliah pemrograman lanjut dan metodologi penelitian, sedang aktif menyusun laporan proyek, namun belum memasuki tahap tugas akhir sehingga masih terdapat waktu memadai untuk memperbaiki kebiasaan akademik. Peserta dibagi ke dalam tujuh kelompok beranggotakan enam orang dengan komposisi campuran antarsemester, masing-masing didampingi satu fasilitator.

**Tim pelaksana.** Tim terdiri atas satu ketua pelaksana, dua anggota dosen dari rumpun bahasa dan informatika, serta empat mahasiswa asisten yang telah memperoleh pembekalan dua sesi mengenai materi, prosedur verifikasi, dan teknik fasilitasi.

**Tahapan kegiatan.** Kegiatan dilaksanakan melalui tujuh tahap. *Tahap pertama*, analisis kebutuhan dan persiapan selama empat pekan, meliputi observasi karya mahasiswa, wawancara dosen, survei pendahuluan, penyusunan modul dan instrumen, serta penyiapan bank kasus. Bank kasus memuat 38 contoh halusinasi terdokumentasi dari interaksi nyata dengan empat model bahasa besar, mencakup referensi fiktif, salah atribusi konsep, kesalahan data statistik, nama pustaka pemrograman yang tidak ada, serta parameter fungsi yang tidak terdapat dalam dokumentasi resmi.

*Tahap kedua*, pembukaan dan prates pada 9 Maret 2026. Sebelum materi apa pun disampaikan, peserta mengerjakan prates selama 40 menit berupa 25 butir soal yang mencakup empat ranah: pengetahuan cara kerja AI generatif (6 butir), pemahaman konsep halusinasi (7 butir), pengetahuan prosedur verifikasi (7 butir), dan pemahaman etika akademik terkait AI (5 butir). Peserta juga mengerjakan tugas kinerja berupa penilaian keaslian lima referensi, tiga di antaranya fabrikatif.

*Tahap ketiga*, penyampaian lima modul materi melalui kombinasi ceramah interaktif, tanya jawab, dan peragaan langsung, masing-masing berdurasi 60 hingga 90 menit dengan jeda praktik singkat.

*Tahap keempat*, demonstrasi halusinasi AI selama 90 menit dalam tiga babak. Pada babak pertama, fasilitator meminta empat perangkat AI menghasilkan lima referensi ilmiah mengenai topik spesifik dalam bidang rekayasa perangkat lunak, yaitu penerapan pembelajaran mesin untuk deteksi kerentanan kode pada perangkat lunak sumber terbuka. Keluaran keempat perangkat ditampilkan berdampingan dan peserta diminta menebak berapa referensi yang benar-benar ada sebelum verifikasi; sebagian besar memperkirakan seluruhnya valid. Pada babak kedua, peserta menelusuri sendiri setiap referensi melalui DOI, Crossref, dan Google Scholar. Dari 20 referensi yang dihasilkan, 10 tidak dapat ditelusuri sama sekali, 4 memiliki penulis nyata dengan judul yang tidak pernah diterbitkan, 3 memuat kesalahan nomor volume atau tahun, dan hanya 3 yang seluruh unsurnya akurat. Pada babak ketiga, demonstrasi dialihkan ke ranah kode program: fasilitator meminta AI menuliskan potongan program menggunakan pustaka tertentu, lalu peserta memeriksa keberadaan pustaka pada repositori paket resmi dan memeriksa keberadaan fungsi pada dokumentasi versi yang digunakan. Ditemukan dua nama paket yang tidak terdaftar sama sekali dan empat pemanggilan fungsi yang tidak terdapat dalam dokumentasi. Babak ketiga ini juga membahas risiko keamanan berupa penyerang yang mendaftarkan paket dengan nama yang sering

dihalusinasikan model, sehingga kelalaian verifikasi dapat berubah menjadi celah keamanan rantai pasok perangkat lunak.

*Tahap kelima*, praktik terbimbing verifikasi selama 120 menit. Setiap peserta menerima lembar kerja berisi sepuluh butir keluaran AI yang terdiri atas empat referensi bibliografis, tiga pernyataan faktual atau teknis, dan tiga potongan kode beserta klaim penyertanya. Peserta menerapkan protokol verifikasi lima langkah, mendokumentasikan penelusuran, dan menetapkan status setiap butir sebagai terverifikasi, sebagian terverifikasi, atau tidak terverifikasi. Fasilitator memberikan bantuan teknis tanpa memberikan jawaban langsung.

*Tahap keenam*, diskusi kelompok terfokus selama 90 menit dengan tujuh tema berbeda, meliputi batas penggunaan AI dalam penyusunan tugas akhir dan proyek pemrograman, bentuk pengungkapan yang praktis, tanggung jawab mahasiswa atas kesalahan yang berasal dari AI, risiko AI terhadap kemampuan berpikir komputasional, peran dosen dalam pembinaan, serta usulan butir protokol etika program studi. Setiap kelompok menyusun rumusan tertulis dan mempresentasikannya, yang kemudian dikonsolidasikan menjadi draf Protokol Penggunaan AI Bertanggung Jawab.

*Tahap ketujuh*, pascates, evaluasi, dan refleksi. Pascates menggunakan instrumen setara dari bank soal yang sama dengan urutan dan pengecoh diacak, disertai tugas kinerja dengan set butir berbeda pada tingkat kesulitan setara. Peserta kemudian mengisi angket kepuasan berisi 15 pernyataan berskala Likert lima tingkat dan tiga pertanyaan terbuka, dilanjutkan refleksi terbuka selama 45 menit.

**Pendampingan pascapelatihan.** Pendampingan berlangsung empat pekan melalui tiga mekanisme: penugasan terapan berupa anotasi bibliografi delapan referensi yang seluruhnya disertai bukti verifikasi, konsultasi terjadwal setiap Jumat yang dimanfaatkan 33 dari 42 peserta, serta forum daring untuk berbagi temuan halusinasi baru yang menghasilkan 71 laporan selama masa pendampingan.

**Instrumen dan analisis data.** Tes pengetahuan 25 butir memiliki koefisien reliabilitas Kuder-Richardson 20 sebesar 0,81 berdasarkan uji coba pada 20 mahasiswa di luar peserta. Rubrik keterampilan verifikasi mencakup empat kriteria, yaitu ketepatan identifikasi status, ketepatan pemilihan alat penelusuran, kelengkapan dokumentasi, dan ketepatan penarikan kesimpulan. Angket kepuasan memiliki koefisien alfa Cronbach 0,88. Data kuantitatif dianalisis dengan statistik deskriptif dan uji *paired sample t-test* yang didahului uji normalitas Shapiro-Wilk, sedangkan besaran peningkatan dihitung menggunakan *Normalized Gain*. Data kualitatif dianalisis melalui pengodean tematik dengan verifikasi silang antaranggota tim.

## 5. MATERI PELATIHAN

### Materi 1: Pengenalan Kecerdasan Buatan Generatif

Modul ini bertujuan membangun model mental yang akurat mengenai cara kerja AI generatif, karena tanpa pemahaman mekanistik peserta cenderung memperlakukan kesalahan AI sebagai anomali yang akan hilang dengan pembaruan versi. Materi diawali penelusuran singkat dari sistem berbasis aturan hingga arsitektur *transformer*, dengan penekanan pada satu gagasan inti: model

dilatih memprediksi kata berikutnya berdasarkan pola statistik, bukan untuk menyimpan atau menelusuri fakta (Bender et al., 2021). Untuk peserta Informatika, penjelasan dapat disampaikan pada tingkat teknis yang lebih dalam, meliputi mekanisme perhatian, tokenisasi, temperatur pembangkitan, jendela konteks, dan batas pengetahuan. Dibahas pula perbedaan penting antara model yang beroperasi murni dari parameter internal dan model yang terhubung dengan penelusuran daring atau *retrieval-augmented generation*, karena perbedaan ini menentukan tingkat keandalan keluaran meskipun keduanya disajikan dengan gaya bahasa yang sama meyakinkan. Modul ditutup dengan pemetaan bersama peserta mengenai tugas yang sesuai dan tidak sesuai untuk didelegasikan kepada AI; penjelasan konsep, curah gagasan, dan penyuntingan bahasa masuk kategori sesuai, sedangkan pencarian referensi, penyediaan data statistik, dan penggunaan kode tanpa pengujian masuk kategori berisiko tinggi.

**Materi 2: Halusinasi AI**

Modul terpanjang ini dibagi menjadi empat bagian. Pada bagian definisi, disampaikan pembedaan Ji et al. (2023) antara halusinasi intrinsik dan ekstrinsik, kerangka Huang et al. (2025) mengenai halusinasi faktual dan halusinasi kesetiaan, klasifikasi distorsi informasi Sun et al. (2024), serta kritik terminologis Emsley (2023). Pada bagian penyebab, pembahasan disusun mengikuti tiga tahap siklus hidup model: pada tahap data mencakup ketidaklengkapan korpus dan kelangkaan data untuk topik atau bahasa tertentu, yang menjelaskan mengapa halusinasi lebih sering muncul pada pertanyaan bertopik lokal Indonesia atau pustaka pemrograman yang jarang digunakan; pada tahap pelatihan mencakup keterbatasan arsitektural dalam menyimpan pengetahuan terstruktur serta efek penyelarasan berbasis umpan balik manusia yang secara tidak sengaja memberi penghargaan pada jawaban yang terdengar percaya diri; pada tahap inferensi mencakup sifat probabilistik pemilihan *token*, akumulasi kesalahan dalam pembangkitan berurutan, dan pengaruh *prompt* yang mengandung praanggapan keliru.

Pada bagian contoh nyata, disajikan temuan Alkaissi dan McFarlane (2023), Athaluri et al. (2023), Walters dan Wilder (2023), McGowan et al. (2023), serta data mutakhir Topaz et al. (2026) mengenai peningkatan tajam kutipan fabrikatif dalam literatur ilmiah. Ditambahkan pula kasus berbahasa Indonesia dari bank kasus tim serta kasus khas bidang informatika berupa nama paket yang tidak terdaftar, fungsi yang tidak ada dalam dokumentasi, dan penjelasan kompleksitas algoritma yang keliru. Pada bagian dampak, pembahasan dilakukan pada empat tingkatan: individu berupa penurunan nilai dan risiko tuduhan pelanggaran akademik; institusi berupa penurunan mutu repositori; ilmu pengetahuan berupa pencemaran rantai bukti sebagaimana ditunjukkan Topaz et al. (2026); serta masyarakat berupa penguatan persoalan misinformasi yang lebih luas (Wach et al., 2023). Khusus untuk bidang informatika, ditambahkan dimensi dampak keamanan, yaitu potensi masuknya kerentanan ke dalam sistem produksi akibat kode yang tidak diverifikasi.

**Materi 3: Teknik Verifikasi Informasi**

Modul paling operasional ini disusun sebagai protokol lima langkah. *Langkah pertama*, klasifikasi jenis klaim menjadi klaim bibliografis, faktual, numerik, dan interpretatif, karena setiap kategori menuntut metode verifikasi berbeda dan klaim interpretatif tidak dapat dinilai benar-salah melainkan dinilai kekuatan argumentasinya. *Langkah kedua*, verifikasi keberadaan melalui empat

jalur berjenjang, yaitu penelusuran DOI langsung, penelusuran metadata melalui Crossref, penelusuran judul lengkap dalam tanda kutip melalui Google Scholar, dan pemeriksaan laman resmi penerbit. Ditekankan bahwa kemunculan hasil serupa tidak cukup; seluruh unsur berupa penulis, tahun, judul, nama jurnal, volume, nomor, dan halaman harus cocok. Untuk klaim teknis, jalur setara adalah pemeriksaan repositori paket resmi dan dokumentasi versi yang digunakan. *Langkah ketiga*, verifikasi isi, yaitu memastikan bahwa sumber yang terbukti ada benar-benar memuat klaim yang dikemukakan AI, dengan membaca abstrak dan bagian hasil secara langsung. *Langkah keempat*, triangulasi minimal dua sumber independen untuk klaim faktual dan numerik penting, dengan pengarahan pada sumber primer resmi untuk data Indonesia dan pada tinjauan sistematis untuk klaim ilmiah. *Langkah kelima*, penilaian mutu dan status sumber melalui pemeriksaan reputasi penerbit, keterindeksan jurnal, dan status pencabutan artikel melalui Retraction Watch, termasuk pengenalan ciri jurnal predator. Protokol ini diringkas dalam lembar kerja saku yang dibagikan kepada seluruh peserta.

**Materi 4: Rekayasa *Prompt* untuk Jawaban yang Lebih Akurat**

Modul ini berangkat dari premis bahwa kualitas keluaran sangat dipengaruhi kualitas masukan, meskipun rekayasa *prompt* tidak dapat menghilangkan risiko halusinasi sepenuhnya (Giray, 2023). Enam prinsip diajarkan: pemberian konteks dan peran yang spesifik; pemberian izin eksplisit bagi model untuk menyatakan ketidaktahuan alih-alih mengarang; penyediaan sumber oleh pengguna, yaitu membalik alur kerja dengan memperoleh artikel atau dokumentasi terlebih dahulu lalu meminta AI membantu memahaminya; penguraian tugas menjadi langkah kecil yang dapat diperiksa bertahap; permintaan penalaran eksplisit agar dasar jawaban dapat dinilai; serta verifikasi silang antarmodel dengan memperlakukan perbedaan jawaban sebagai penanda perlunya verifikasi manual. Modul ditutup dengan penegasan berulang bahwa rekayasa *prompt* merupakan upaya mitigasi, bukan jaminan, dan tidak ada rumusan *prompt* yang membuat verifikasi menjadi tidak perlu.

**Materi 5: Etika Penggunaan AI di Perguruan Tinggi**

Modul ini menempatkan seluruh keterampilan teknis dalam kerangka normatif, diawali penegasan prinsip tanggung jawab akhir bahwa tanggung jawab atas isi karya sepenuhnya berada pada mahasiswa sebagai penulis, apa pun proses yang ditempuh. Berikutnya dibahas kerangka bertingkat penggunaan AI dengan mengadaptasi *AI Assessment Scale* (Perkins et al., 2024), yang menetapkan tingkatan mulai dari penugasan tanpa AI sama sekali hingga penggunaan penuh dengan pengungkapan. Peserta kemudian dilatih menyusun pernyataan pengungkapan yang memuat perangkat yang digunakan, bagian pekerjaan yang dibantu, bentuk bantuan, dan langkah verifikasi yang telah dilakukan, serta diperkenalkan pada kebijakan penerbit ilmiah yang menetapkan bahwa sistem AI tidak dapat dicantumkan sebagai penulis. Bagian berikutnya mengaitkan praktik tersebut dengan kerangka nasional, yaitu bahwa meskipun Undang-Undang Nomor 12 Tahun 2012 dan Peraturan Menteri Pendidikan Nasional Nomor 17 Tahun 2010 belum mengatur AI generatif secara khusus, prinsip kejujuran, orisinalitas, dan pertanggungjawaban di dalamnya tetap berlaku penuh. Bagian terakhir membahas dimensi etika yang lebih luas meliputi kerahasiaan data ketika memasukkan kode atau data proyek ke layanan AI publik, persoalan lisensi kode yang dihasilkan

model, potensi bias keluaran, dan kesenjangan akses (Cotton et al., 2024; UNESCO, 2023). Modul ditutup dengan penyusunan bersama draf protokol etika yang dilanjutkan pada sesi diskusi kelompok.

## 6. HASIL DAN PEMBAHASAN

### 6.1 Gambaran Umum Pelaksanaan

Seluruh rangkaian terlaksana sesuai rencana dengan tingkat kehadiran 100% pada kedua hari pelatihan inti. Tingkat penyelesaian tugas pendampingan mencapai 88,1%, yaitu 37 dari 42 peserta menyerahkan anotasi bibliografi terverifikasi sesuai tenggat. Angka ini tergolong tinggi untuk kegiatan sukarela dan menunjukkan keterlibatan peserta yang baik.

Dinamika kegiatan menunjukkan pola yang konsisten dengan rancangan pedagogis. Pada sesi pembukaan suasana relatif pasif dengan sedikit pertanyaan. Perubahan tajam terjadi setelah sesi demonstrasi, terutama pada babak ketiga ketika peserta mendapati bahwa nama pustaka pemrograman pun dapat difabrikasi. Fasilitator mencatat 21 peserta melaporkan menemukan sekurang-kurangnya satu referensi atau pemanggilan pustaka mencurigakan pada tugas mereka sendiri pada hari itu juga. Momen ini menegaskan bahwa pengalaman diskonfirmasi langsung memiliki daya ubah yang jauh melampaui penjelasan konseptual.

### 6.2 Karakteristik Peserta

**Tabel 1. Karakteristik Peserta Pelatihan (n = 42)**

| Kategori | Klasifikasi | Frekuensi | Persentase (%) |
|---|---|---|---|
| Jenis kelamin | Laki-laki | 26 | 61,9 |
| | Perempuan | 16 | 38,1 |
| Semester | Semester 4 | 24 | 57,1 |
| | Semester 6 | 18 | 42,9 |
| Frekuensi penggunaan AI | Setiap hari | 24 | 57,1 |
| | Dua hingga tiga kali sepekan | 12 | 28,6 |
| | Sekali sepekan | 4 | 9,5 |
| | Jarang | 2 | 4,8 |
| Perangkat utama | ChatGPT | 24 | 57,1 |
| | Gemini | 8 | 19,0 |
| | Copilot | 7 | 16,7 |
| | Claude | 3 | 7,1 |
| Tujuan penggunaan dominan | Membantu penulisan laporan | 15 | 35,7 |
| | Membuat atau memperbaiki | 19 | 45,2 |

| Kategori | Klasifikasi | Frekuensi | Persentase (%) |
|---|---|---|---|
| | kode | | |
| | Mencari referensi | 6 | 14,3 |
| | Lainnya | 2 | 4,8 |
| Pengetahuan awal istilah halusinasi AI | Pernah mendengar | 11 | 26,2 |
| | Belum pernah mendengar | 31 | 73,8 |

Tabel 1 memuat beberapa informasi bermakna. Pertama, tingkat penetrasi AI sangat tinggi: 85,7% peserta menggunakan AI sekurang-kurangnya dua hingga tiga kali sepekan dan lebih dari separuh menggunakannya setiap hari. Angka ini lebih tinggi dibandingkan pola umum yang dilaporkan Chan dan Hu (2023), yang dapat dijelaskan oleh sifat pekerjaan mahasiswa Informatika yang berkelindan langsung dengan alat bantu pemrograman berbasis AI. Kedua, terdapat kesenjangan mencolok antara intensitas penggunaan dan pemahaman risiko: meskipun 85,7% menggunakan AI secara rutin, hanya 26,2% pernah mendengar istilah halusinasi AI. Kesenjangan sebesar ini merupakan indikator kerentanan serius, karena mahasiswa berinteraksi intensif dengan teknologi yang keterbatasan mendasarnya tidak mereka ketahui — kondisi yang menurut Klingbeil et al. (2024) paling kondusif bagi terbentuknya ketergantungan berlebihan.

Ketiga, dominasi penggunaan untuk pembuatan dan perbaikan kode sebesar 45,2% menegaskan pentingnya perluasan materi verifikasi ke ranah teknis, yang membedakan kegiatan ini dari pelatihan literasi AI pada umumnya. Keempat, dominasi ChatGPT sebesar 57,1% menunjukkan konsentrasi penggunaan pada satu perangkat, yang berimplikasi pada kecenderungan peserta menggeneralisasi karakteristik satu model kepada seluruh model. Hal ini menjadi salah satu alasan tim menyajikan demonstrasi lintas perangkat.

### 6.3 Hasil Prates dan Pascates

**Tabel 2. Perbandingan Hasil Prates dan Pascates per Aspek (n = 42)**

| Aspek pengukuran | Butir | Rerata prates | SB prates | Rerata pascates | SB pascates |
|---|---|---|---|---|---|
| Pengetahuan cara kerja AI generatif | 6 | 62,50 | 12,74 | 88,90 | 7,86 |
| Pemahaman konsep halusinasi AI | 7 | 41,80 | 12,05 | 85,30 | 7,64 |
| Pengetahuan prosedur verifikasi informasi | 7 | 40,50 | 11,63 | 84,10 | 8,22 |
| Pemahaman etika akademik terkait AI | 5 | 55,60 | 13,91 | 74,20 | 10,47 |
| **Rerata keseluruhan** | **25** | **49,16** | **11,05** | **83,61** | **7,93** |

**Tabel 3. Distribusi Kategori Nilai Prates dan Pascates (n = 42)**

| Rentang nilai | Kategori | Prates | % | Pascates | % |
|---|---|---|---|---|---|
| 80–100 | Sangat baik | 0 | 0,0 | 29 | 69,0 |
| 70–79 | Baik | 3 | 7,1 | 10 | 23,8 |
| 60–69 | Cukup | 8 | 19,0 | 3 | 7,1 |
| < 60 | Kurang | 31 | 73,8 | 0 | 0,0 |
| **Total** | | **42** | **100,0** | **42** | **100,0** |

Hasil prates menegaskan urgensi kegiatan. Rerata keseluruhan 49,16 berada jauh di bawah batas ketuntasan 70 yang ditetapkan tim, dan 73,8% peserta berada pada kategori kurang tanpa seorang pun mencapai kategori sangat baik. Pola antaraspek memberikan informasi diagnostik yang berharga. Aspek pengetahuan cara kerja AI memperoleh skor tertinggi pada prates, yaitu 62,50, yang wajar mengingat latar belakang keilmuan peserta. Namun skor tersebut ternyata tidak berbanding lurus dengan pemahaman mengenai halusinasi, yang justru memperoleh skor terendah kedua yaitu 41,80. Temuan ini menarik dan penting: penguasaan pengetahuan teknis mengenai arsitektur model tidak otomatis disertai kesadaran mengenai implikasi praktis keterbatasannya. Analisis butir menunjukkan kesalahan terbanyak terjadi pada soal mengenai penyebab halusinasi, di mana mayoritas peserta memilih pengecoh berupa keterbatasan data pelatihan yang belum diperbarui, mencerminkan model mental yang memperlakukan AI sebagai basis data yang perlu disegarkan alih-alih sebagai sistem probabilistik.

Aspek prosedur verifikasi memperoleh skor terendah, yaitu 40,50. Hanya 14,3% peserta mengetahui fungsi DOI dan hanya 7,1% mengetahui keberadaan Crossref. Rendahnya angka ini menunjukkan bahwa persoalan bukan pada kemauan memverifikasi melainkan pada ketiadaan pengetahuan mengenai alat yang tersedia. Sebaliknya, aspek etika akademik memperoleh skor 55,60 yang relatif lebih baik karena peserta telah memperoleh sosialisasi mengenai plagiarisme pada mata kuliah metodologi penelitian. Namun analisis butir menunjukkan pengetahuan tersebut bersifat umum: 85,7% peserta menjawab keliru pada butir mengenai pengungkapan penggunaan AI dan 78,6% keliru pada butir mengenai pihak yang bertanggung jawab atas kesalahan faktual yang berasal dari AI. Perbandingan distribusi pada Tabel 3 menunjukkan pergeseran yang sangat jelas. Kategori kurang yang semula menampung 73,8% peserta menjadi kosong sepenuhnya, sementara kategori sangat baik yang semula kosong kini menampung 69,0% peserta. Tidak seorang pun peserta mengalami penurunan skor. Tiga peserta yang tetap berada pada kategori cukup seluruhnya berasal dari semester empat dengan pengalaman penulisan akademik yang masih terbatas; ketiganya memperoleh pendampingan tambahan dan pada akhir masa pendampingan menunjukkan kemampuan verifikasi yang memadai berdasarkan penilaian tugas terapan.

### 6.4 Analisis Peningkatan

**Tabel 4. Persentase Peningkatan dan Nilai *N-Gain* per Aspek (n = 42)**

| Aspek pengukuran | Prates | Pascates | Selisih | Peningkatan (%) | *N-Gain* | Kategori |
|---|---|---|---|---|---|---|
| Pengetahuan cara kerja AI generatif | 62,50 | 88,90 | 26,40 | 42,24 | 0,70 | Tinggi |
| Pemahaman konsep halusinasi AI | 41,80 | 85,30 | 43,50 | 104,07 | 0,75 | Tinggi |
| Pengetahuan prosedur verifikasi informasi | 40,50 | 84,10 | 43,60 | 107,65 | 0,73 | Tinggi |
| Pemahaman etika akademik terkait AI | 55,60 | 74,20 | 18,60 | 33,45 | 0,42 | Sedang |
| **Keseluruhan** | **49,16** | **83,61** | **34,45** | **70,08** | **0,68** | **Sedang** |

**Tabel 5. Hasil Uji *Paired Sample t-Test***

| Statistik | Nilai |
|---|---|
| Rerata prates | 49,16 |
| Rerata pascates | 83,61 |
| Rerata selisih | 34,45 |
| Simpangan baku selisih | 9,02 |
| Nilai t hitung | 24,75 |
| Derajat kebebasan | 41 |
| Nilai signifikansi (2-arah) | < 0,001 |
| Cohen's *d* | 3,82 |

Uji normalitas Shapiro-Wilk terhadap data selisih menghasilkan nilai signifikansi 0,228 sehingga asumsi normalitas terpenuhi. Hasil uji *paired sample t-test* menunjukkan perbedaan yang sangat signifikan, dengan nilai *Cohen's d* sebesar 3,82 yang menandakan ukuran efek sangat besar.
Pola peningkatan antaraspek memberikan wawasan pedagogis penting. Peningkatan relatif terbesar terjadi pada aspek prosedur verifikasi sebesar 107,65% dan aspek pemahaman halusinasi sebesar 104,07%, keduanya dengan *N-Gain* kategori tinggi. Besarnya peningkatan pada kedua aspek ini dapat dijelaskan oleh skor awal yang rendah sekaligus, dan lebih penting, oleh fakta bahwa keduanya merupakan aspek yang paling intensif diperkuat melalui demonstrasi dan praktik langsung. Peserta tidak sekadar diberi tahu bahwa AI dapat berhalusinasi, tetapi menyaksikan dan membuktikannya sendiri; peserta juga tidak sekadar diberi tahu mengenai keberadaan DOI dan Crossref, tetapi mengoperasikannya secara langsung.

Sebaliknya, aspek etika akademik menunjukkan *N-Gain* terendah sebesar 0,42. Terdapat dua penjelasan yang saling melengkapi. Penjelasan statistik menunjuk pada skor awal yang lebih

tinggi sehingga ruang peningkatan lebih sempit. Penjelasan substantif yang lebih penting adalah bahwa pemahaman etika bukan sekadar pengetahuan deklaratif melainkan disposisi nilai yang terbentuk melalui pembiasaan jangka panjang. Satu sesi pelatihan dapat memperkenalkan kerangka bertingkat dan praktik pengungkapan (Perkins et al., 2024), tetapi internalisasi memerlukan penguatan berulang melalui kebijakan penilaian dan keteladanan dosen. Temuan ini sejalan dengan simpulan Bittle dan El-Gayar (2025) bahwa mitigasi risiko integritas akademik menuntut kombinasi pendidikan, desain penilaian, dan kebijakan institusional, bukan pelatihan tunggal.

### 6.5 Peningkatan Kemampuan Verifikasi

**Tabel 6. Perbandingan Kinerja Tugas Verifikasi (n = 42)**

| Indikator kinerja | Prates | Pascates | Peningkatan |
|---|---|---|---|
| Mengidentifikasi seluruh referensi fabrikatif dengan tepat | 11 (26,2%) | 38 (90,5%) | +64,3 poin |
| Menggunakan DOI sebagai alat verifikasi utama | 6 (14,3%) | 41 (97,6%) | +83,3 poin |
| Menggunakan Crossref atau pangkalan data penerbit | 3 (7,1%) | 35 (83,3%) | +76,2 poin |
| Memeriksa kesesuaian seluruh unsur bibliografis | 8 (19,0%) | 37 (88,1%) | +69,1 poin |
| Memeriksa keberadaan pustaka pada repositori paket resmi | 5 (11,9%) | 39 (92,9%) | +81,0 poin |
| Memverifikasi kesesuaian isi sumber dengan klaim | 4 (9,5%) | 30 (71,4%) | +61,9 poin |
| Memeriksa status pencabutan artikel | 1 (2,4%) | 24 (57,1%) | +54,7 poin |
| Menerima keluaran AI tanpa penelusuran apa pun | 17 (40,5%) | 1 (2,4%) | −38,1 poin |

Tabel 6 menunjukkan perubahan perilaku verifikasi yang substansial. Indikator terakhir merupakan yang paling bermakna: proporsi peserta yang menerima keluaran AI tanpa penelusuran turun dari 40,5% menjadi 2,4%, menandakan pergeseran sikap dasar dari penerimaan pasif menuju skeptisisme metodis. Peningkatan tertinggi terjadi pada indikator penggunaan DOI sebesar 83,3 poin persentase dan pemeriksaan repositori paket resmi sebesar 81,0 poin. Kedua angka ini menunjukkan bahwa hambatan utama bersifat informasional semata: peserta tidak melakukan verifikasi karena tidak mengetahui adanya alat yang tersedia, bukan karena enggan. Begitu alat diperkenalkan dan didemonstrasikan, adopsinya nyaris menyeluruh.

Dua indikator menunjukkan peningkatan lebih moderat. Indikator verifikasi kesesuaian isi hanya mencapai 71,4% dan pemeriksaan status pencabutan hanya 57,1%. Keduanya menuntut upaya lebih besar, yaitu pembacaan sumber secara sungguh-sungguh dan penggunaan alat tambahan yang belum familier. Temuan ini menunjukkan adanya gradasi kesulitan dalam praktik verifikasi, di mana langkah cepat dan mekanis lebih mudah diadopsi dibandingkan langkah yang menuntut pembacaan mendalam. Implikasinya bagi program lanjutan adalah perlunya alokasi waktu praktik yang lebih besar khusus untuk verifikasi isi.

Hasil penilaian tugas terapan pascapendampingan memperkuat temuan tersebut. Dari 37 peserta yang menyerahkan anotasi bibliografi, 34 peserta atau 91,9% menyerahkan dokumen dengan seluruh referensi terverifikasi dan terdokumentasi, dan tidak ditemukan satu pun referensi

fabrikatif dalam keseluruhan tugas yang diserahkan. Kondisi ini berbeda tajam dengan keadaan awal, ketika 12 dari 30 laporan sampel memuat referensi yang tidak tertelusur.

**6.6 Peningkatan Literasi AI dan Kesadaran Etika**

Analisis tematik terhadap jawaban pertanyaan terbuka dan catatan refleksi mengidentifikasi tiga tema. *Pertama*, pergeseran model mental. Sebelum pelatihan, peserta dominan menggambarkan AI sebagai sumber jawaban atau mesin pencari yang lebih pintar; setelah pelatihan, deskripsi bergeser menjadi mitra berpikir yang harus diperiksa. Pergeseran metafora ini bermakna karena metafora yang digunakan seseorang menentukan cara ia memperlakukan teknologi tersebut, dan sejalan dengan pendekatan berpusat pada manusia yang dianjurkan UNESCO (2023). *Kedua*, kesadaran akan pola kerentanan. Peserta mulai mampu memprediksi kapan AI cenderung berhalusinasi, dengan menyebutkan topik lokal Indonesia, peraturan perundangan, data statistik spesifik, serta pustaka pemrograman yang jarang digunakan atau baru dirilis. Kemampuan memprediksi ini menunjukkan terbangunnya model mental yang fungsional, bukan sekadar hafalan peringatan umum, dan merupakan indikator kompetensi teknis dalam kerangka UNESCO (2024). *Ketiga*, perubahan alur kerja. Sebelum pelatihan, alur umum adalah meminta AI mencari referensi lalu menuliskannya, serta menyalin kode langsung ke proyek; setelah pelatihan, alur yang dilaporkan adalah mencari artikel terlebih dahulu melalui pangkalan data lalu menggunakan AI untuk membantu memahaminya, serta memeriksa dokumentasi resmi sebelum menggunakan kode. Perubahan alur kerja ini merupakan hasil paling bernilai karena secara struktural menghilangkan peluang fabrikasi.

Pada aspek etika, meskipun peningkatan kuantitatif tergolong paling rendah, terdapat indikator kualitatif yang bermakna. Indikator pertama adalah dihasilkannya draf Protokol Penggunaan AI Bertanggung Jawab yang disusun sendiri oleh peserta, memuat sembilan butir kesepakatan antara lain kewajiban memverifikasi setiap klaim faktual, larangan menyerahkan karya yang seluruhnya dihasilkan AI, kewajiban mencantumkan pernyataan pengungkapan, kewajiban menguji setiap kode sebelum digunakan, serta larangan memasukkan kode atau data proyek yang bersifat rahasia ke layanan AI publik. Fakta bahwa protokol disusun peserta sendiri meningkatkan kemungkinan kepatuhan karena adanya rasa kepemilikan. Indikator kedua adalah perubahan sikap terhadap keterbukaan: kesediaan mengungkapkan penggunaan AI kepada dosen meningkat dari 14,3% menjadi 81,0%. Perubahan ini penting karena keterbukaan merupakan prasyarat pembinaan; selama penggunaan disembunyikan, dosen tidak memiliki kesempatan membimbing (Cotton et al., 2024). Indikator ketiga adalah pergeseran pemahaman mengenai lokus tanggung jawab, dari 78,6% menjawab keliru pada prates menjadi 88,1% menjawab tepat pada pascates bahwa tanggung jawab sepenuhnya berada pada penulis karya.

Analisis kualitatif juga mengungkap dua kerentanan yang tersisa. Sebagian kecil peserta menunjukkan skeptisisme berlebihan dengan menyatakan keinginan berhenti menggunakan AI sepenuhnya, sikap yang tidak diinginkan karena tujuan kegiatan adalah penggunaan kritis, bukan penolakan. Selain itu, terdapat kecenderungan menganggap perangkat yang menyertakan tautan sumber sebagai bebas kesalahan, padahal sistem yang menyertakan rujukan pun dapat

menampilkan atribusi yang tidak akurat. Kedua kerentanan ini menjadi catatan penyempurnaan modul.

### 6.7 Kepuasan Peserta

**Tabel 7. Hasil Angket Kepuasan Peserta (n = 42)**

| Aspek penilaian | Skor rerata (1–5) | Persentase (%) | Kategori |
|---|---|---|---|
| Kesesuaian materi dengan kebutuhan | 4,69 | 93,8 | Sangat puas |
| Kejelasan penyampaian materi | 4,60 | 91,9 | Sangat puas |
| Kebermanfaatan sesi praktik verifikasi | 4,76 | 95,2 | Sangat puas |
| Kecukupan alokasi waktu | 4,15 | 82,9 | Puas |
| Sarana dan prasarana pelatihan | 4,48 | 89,5 | Sangat puas |
| Kompetensi dan responsivitas fasilitator | 4,72 | 94,3 | Sangat puas |
| **Rerata keseluruhan** | **4,57** | **91,3** | **Sangat puas** |

Tingkat kepuasan tinggi tercatat pada seluruh aspek dengan rerata 91,3%. Aspek kebermanfaatan sesi praktik memperoleh skor tertinggi 95,2%, menguatkan temuan bahwa komponen praktik langsung merupakan bagian yang paling dihargai peserta dan konsisten dengan pola *N-Gain* tertinggi pada aspek yang paling intensif dipraktikkan. Aspek kecukupan waktu memperoleh skor terendah 82,9%, dengan keluhan terkonsentrasi pada sesi praktik dan diskusi. Sebanyak 25 peserta menyatakan sesi praktik terlalu singkat untuk menuntaskan seluruh sepuluh butir lembar kerja. Masukan ini diterima sebagai catatan perbaikan: pada penyelenggaraan berikutnya alokasi sesi praktik akan ditingkatkan dari 120 menjadi 180 menit, dengan pengurangan durasi ceramah Materi 1 yang sebagian dialihkan menjadi bahan bacaan mandiri prapelatihan. Usulan yang paling sering muncul pada pertanyaan terbuka adalah permintaan agar pelatihan serupa diselenggarakan bagi seluruh mahasiswa baru, disampaikan 31 dari 42 peserta, disusul permintaan modul cetak dan pelatihan lanjutan khusus untuk tugas akhir.

### 6.8 Dokumentasi Kegiatan

**Gambar 1. Pembukaan Pelatihan.** Suasana Ruang Seminar pada 9 Maret 2026 pukul 08.15 WIB, menampilkan pimpinan institut menyampaikan sambutan di hadapan 42 peserta yang duduk berkelompok sesuai pembagian kelompok kerja, dengan layar proyeksi dan spanduk kegiatan di bagian depan ruangan.

**Gambar 2. Penyampaian Materi.** Sesi Materi 2 mengenai halusinasi AI, menampilkan pemateri di samping layar yang memuat diagram tiga tahap penyebab halusinasi, dengan beberapa peserta mengangkat tangan untuk bertanya.

**Gambar 3. Praktik Verifikasi Informasi.** Suasana Laboratorium Komputer pada sesi praktik hari kedua, menampilkan peserta mengoperasikan perangkat masing-masing dengan layar yang menampilkan laman Crossref, Google Scholar, dan dokumentasi pustaka pemrograman secara berdampingan, serta seorang fasilitator yang mendampingi peserta menelusuri DOI.

**Gambar 4. Diskusi Kelompok.** Salah satu dari tujuh kelompok yang membahas batas penggunaan AI dalam proyek pemrograman, dengan enam peserta duduk melingkar dan kertas plano berisi rumusan butir protokol hasil kesepakatan kelompok.

**Gambar 5. Foto Bersama.** Seluruh peserta, tim pelaksana, dan pimpinan institut berfoto bersama di halaman kampus pada akhir hari kedua, dengan peserta memegang sertifikat keikutsertaan dan lembar protokol saku verifikasi lima langkah.

### 6.9 Pembahasan Terpadu

Secara keseluruhan, hasil kegiatan mendukung tiga simpulan analitis. *Pertama*, pengalaman diskonfirmasi langsung merupakan komponen paling berdaya ubah. Perbandingan antaraspek menunjukkan bahwa aspek yang diperkuat melalui demonstrasi dan praktik memperoleh *N-Gain* kategori tinggi, sedangkan aspek yang lebih banyak disampaikan secara konseptual memperoleh *N-Gain* kategori sedang. Implikasinya bagi perancangan program serupa adalah bahwa alokasi waktu sebaiknya condong pada demonstrasi dan praktik, bukan pada ceramah. Temuan ini sejalan dengan argumen Klingbeil et al. (2024) bahwa ketergantungan berlebihan berakar pada kepercayaan yang tidak teruji, sehingga intervensi paling efektif adalah yang menguji kepercayaan tersebut secara langsung.

*Kedua*, penguasaan teknis tidak identik dengan literasi kritis. Temuan bahwa peserta memperoleh skor prates tertinggi pada aspek cara kerja AI namun skor rendah pada aspek halusinasi dan verifikasi menunjukkan bahwa kedekatan dengan teknologi tidak otomatis menghasilkan sikap kritis terhadapnya. Temuan ini penting untuk membantah asumsi umum bahwa mahasiswa bidang teknologi tidak memerlukan pelatihan literasi AI. Justru karena intensitas penggunaan mereka paling tinggi dan cakupan penggunaannya paling luas, kelompok ini memerlukan pembekalan verifikasi yang paling menyeluruh.

*Ketiga*, pengetahuan prosedural lebih cepat berubah daripada disposisi nilai. Kesenjangan antara *N-Gain* aspek verifikasi sebesar 0,73 dan aspek etika sebesar 0,42 menunjukkan bahwa pelatihan singkat efektif membangun keterampilan tetapi terbatas dalam membentuk disposisi. Karena itu pelatihan tidak dapat berdiri sendiri dan harus disertai perubahan pada tingkat kebijakan penilaian serta praktik pembelajaran sehari-hari, sebagaimana direkomendasikan Chan (2023). Hasil yang paling menjanjikan dari kegiatan ini pada akhirnya bukan skor pascates, melainkan laporan peserta mengenai pembalikan urutan kerja dari pola "AI dahulu, sumber kemudian" menjadi "sumber dahulu, AI kemudian", karena perubahan struktural pada alur kerja bersifat lebih tahan lama daripada kedisiplinan yang harus diperbarui setiap kali.

## 7. DAMPAK PENGABDIAN

**Bagi mahasiswa,** dampak paling langsung berupa peningkatan kompetensi yang terukur. Namun dampak yang lebih bermakna bersifat kualitatif, yaitu diperolehnya kerangka berpikir yang memungkinkan peserta mengevaluasi teknologi baru yang akan mereka hadapi di masa depan, bukan hanya perangkat yang ada saat ini. Protokol verifikasi lima langkah bersifat agnostik terhadap perangkat sehingga tetap relevan ketika model baru bermunculan. Dampak kedua adalah

berkurangnya risiko sanksi akademik, sebagaimana ditunjukkan hasil tugas terapan yang tidak memuat satu pun referensi fabrikatif. Dampak ketiga bersifat psikologis: beberapa peserta menyampaikan bahwa mereka sebelumnya merasa bersalah menggunakan AI karena ketiadaan kejelasan batas, dan kejelasan kerangka yang diperoleh mengurangi kegelisahan tersebut.

**Bagi perguruan tinggi,** kegiatan menghasilkan tiga dampak kelembagaan: tersedianya draf Protokol Penggunaan AI Bertanggung Jawab sebagai bahan masukan penyusunan kebijakan institusional; terbentuknya kader mahasiswa yang dapat berperan sebagai penggerak literasi AI; serta peningkatan mutu karya akademik mahasiswa yang berdampak pada mutu repositori institusional dan pemenuhan standar penjaminan mutu sebagaimana diamanatkan Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 53 Tahun 2023.

**Bagi dosen,** kegiatan memberikan pemutakhiran kompetensi mengenai teknologi yang digunakan mahasiswa sehari-hari, sekaligus menyediakan bahan konkret untuk merancang penugasan yang lebih tahan terhadap penyalahgunaan AI. Sejumlah dosen yang mengamati kegiatan berencana menambahkan persyaratan bukti verifikasi pada tugas penulisan dan persyaratan demonstrasi kode yang dapat dijalankan pada tugas pemrograman. Persyaratan sederhana semacam ini terbukti efektif menghilangkan keluaran fabrikatif tanpa memerlukan perangkat deteksi apa pun.

**Bagi budaya akademik,** kegiatan berkontribusi pada normalisasi dua hal yang sebelumnya tidak lazim, yaitu sikap meragukan sumber dan keterbukaan mengenai penggunaan alat bantu. Peningkatan kesediaan mengungkapkan penggunaan AI dari 14,3% menjadi 81,0% menunjukkan bahwa keterbukaan dapat dibangun apabila mahasiswa memperoleh kerangka yang jelas dan tidak merasa terancam.

**Terhadap penggunaan AI yang bertanggung jawab,** data menunjukkan bahwa peserta tidak mengurangi penggunaan AI secara keseluruhan, melainkan mengubah jenis penggunaannya. Survei pendampingan pada pekan keempat menunjukkan penggunaan AI untuk pencarian referensi menurun dari 71,4% menjadi 19,0%, sedangkan penggunaan untuk membantu memahami sumber dan dokumentasi yang telah dimiliki meningkat dari 26,2% menjadi 73,8%. Pergeseran ini merupakan bukti paling langsung bahwa tujuan kegiatan tercapai, yaitu bukan pengurangan pemanfaatan melainkan pemindahan pemanfaatan ke wilayah yang aman dan produktif.

## 8. FAKTOR PENDUKUNG DAN KENDALA

**Faktor pendukung.** Pertama, dukungan institusional berupa penyediaan ruang, akses laboratorium, dan izin bagi mahasiswa mengikuti kegiatan pada hari kerja, yang menghilangkan hambatan administratif. Kedua, relevansi tema dengan kebutuhan nyata peserta, terlihat dari tingkat kehadiran penuh, intensitas pertanyaan, dan tingkat penyelesaian tugas pendampingan yang tinggi tanpa memerlukan insentif tambahan. Ketiga, ketersediaan infrastruktur teknologi berupa laboratorium dengan koneksi memadai serta kesediaan peserta membawa perangkat pribadi. Keempat, ketersediaan bank kasus autentik berjumlah 38 contoh terdokumentasi yang dipilih khusus dari bidang studi peserta, yang terbukti jauh lebih berdaya ubah dibandingkan contoh generik dari literatur asing. Kelima, komposisi tim lintas bidang yang memadukan keahlian bahasa dan

informatika, sehingga materi teknis dapat disampaikan dengan bahasa yang dapat dipahami sekaligus tetap akurat secara substansi.

**Kendala.** Pertama, keterbatasan alokasi waktu, terutama pada sesi praktik dan diskusi; rancangan dua hari memadai untuk menyampaikan materi tetapi terlalu padat untuk pendalaman praktik. Kendala ini ditangani sementara melalui perpanjangan sesi konsultasi dan akan ditangani secara struktural melalui penambahan durasi. Kedua, keragaman kemampuan awal antara peserta semester empat dan semester enam, yang ditangani melalui komposisi kelompok heterogen sehingga terjadi tutor sebaya serta pendampingan tambahan bagi peserta yang tertinggal. Ketiga, keterbatasan akses terhadap pangkalan data berlangganan, sehingga tidak seluruh peserta dapat mengakses naskah lengkap artikel; kendala ini ditangani melalui pengenalan repositori akses terbuka dan portal jurnal nasional, meskipun tetap membatasi kedalaman praktik verifikasi isi sebagaimana tercermin pada capaian indikator tersebut yang hanya 71,4%. Keempat, sifat teknologi yang berubah cepat, di mana sejumlah kasus dalam bank kasus tidak lagi dapat direproduksi karena model telah diperbarui; kendala ini ditangani melalui penyusunan bank kasus yang bersifat dinamis dan pelibatan peserta dalam pengumpulan kasus baru. Kelima, kecenderungan skeptisisme berlebihan pada sebagian kecil peserta, yang ditangani melalui penegasan ulang pada sesi refleksi bahwa tujuan kegiatan adalah penggunaan kritis, bukan penolakan.

## 9. KEBERLANJUTAN PROGRAM

Tim menyusun lima rencana tindak lanjut. **Pembentukan Komunitas Literasi AI** dengan keanggotaan awal 42 peserta ditambah empat mahasiswa asisten, menjalankan tiga fungsi yaitu pemeliharaan dan pengembangan bank kasus halusinasi berbahasa Indonesia, penyelenggaraan pertemuan bulanan untuk berbagi temuan, serta penyediaan layanan konsultasi sebaya. **Penyusunan modul cetak dan digital** yang dilengkapi latihan mandiri, kunci jawaban, dan tautan sumber, direncanakan memperoleh nomor ISBN dan didaftarkan sebagai hak cipta agar dapat digunakan institusi lain, dengan versi digital tersedia terbuka melalui repositori institusi. **Penyelenggaraan lokakarya lanjutan** dalam dua bentuk, yaitu lokakarya bagi mahasiswa semester enam ke atas mengenai penggunaan AI secara bertanggung jawab dalam tugas akhir dengan penekanan pada verifikasi isi dan penulisan pernyataan pengungkapan, serta lokakarya bagi dosen mengenai perancangan penugasan yang tahan terhadap penyalahgunaan AI. **Integrasi ke dalam pembelajaran** melalui penambahan satu pertemuan khusus mengenai halusinasi AI dan verifikasi pada mata kuliah metodologi penelitian dan penulisan ilmiah, serta penambahan persyaratan bukti verifikasi pada rubrik penilaian tugas; melalui integrasi ini materi tidak lagi bergantung pada penyelenggaraan kegiatan pengabdian berkala melainkan menjadi bagian kurikulum reguler. **Perluasan jangkauan dan replikasi** kepada mahasiswa baru sebagai bagian dari kegiatan pengenalan kehidupan kampus, serta kerja sama dengan perguruan tinggi lain di wilayah Surakarta untuk mereplikasi model pelatihan dengan penyesuaian bank kasus sesuai bidang keilmuan masing-masing.

## 10. KESIMPULAN

Kegiatan pengabdian kepada masyarakat berjudul "Pelatihan Verifikasi Informasi dan Pencegahan Halusinasi *Artificial Intelligence* bagi Mahasiswa" telah dilaksanakan dengan hasil yang memenuhi seluruh tujuan yang ditetapkan. Kegiatan diikuti 42 mahasiswa Program Studi Informatika Institut Teknologi Bisnis AAS Indonesia semester empat dan semester enam, dilaksanakan melalui tujuh tahap yang mencakup analisis kebutuhan, prates, penyampaian lima modul, demonstrasi halusinasi AI pada referensi ilmiah sekaligus pada kode program, praktik terbimbing verifikasi, diskusi kelompok, serta pascates dengan evaluasi dan refleksi, dilanjutkan pendampingan empat pekan.

Rerata skor pengetahuan meningkat dari 49,16 pada prates menjadi 83,61 pada pascates, dengan *N-Gain* 0,68 kategori sedang serta hasil uji *paired sample t-test* yang signifikan pada taraf lebih kecil dari 0,001 dan ukuran efek yang sangat besar. Peningkatan tertinggi terjadi pada aspek prosedur verifikasi dengan *N-Gain* 0,73 dan pemahaman halusinasi dengan *N-Gain* 0,75, sedangkan terendah pada aspek etika akademik dengan *N-Gain* 0,42. Kemampuan mengidentifikasi referensi fabrikatif meningkat dari 26,2% menjadi 90,5%, kemampuan memeriksa keberadaan pustaka pemrograman pada repositori resmi meningkat dari 11,9% menjadi 92,9%, dan proporsi peserta yang menerima keluaran AI tanpa penelusuran turun dari 40,5% menjadi 2,4%. Tingkat kepuasan peserta mencapai 91,3%.

Tiga simpulan analitis dapat ditarik. Pertama, pengalaman diskonfirmasi langsung melalui demonstrasi dan praktik verifikasi merupakan komponen yang paling berdaya ubah, jauh melampaui penyampaian konseptual. Kedua, penguasaan teknis mengenai cara kerja AI tidak identik dengan literasi kritis terhadapnya, sehingga mahasiswa bidang teknologi justru memerlukan pembekalan verifikasi yang paling menyeluruh mengingat intensitas dan cakupan penggunaan mereka yang paling tinggi. Ketiga, keterampilan prosedural lebih cepat terbentuk dibandingkan disposisi etis, sehingga pelatihan perlu disertai perubahan kebijakan penilaian dan praktik pembelajaran agar dampaknya bertahan.

Kegiatan ini merekomendasikan agar literasi AI dan keterampilan verifikasi informasi diposisikan sebagai kompetensi dasar yang diintegrasikan ke dalam kurikulum, bukan kegiatan tambahan yang bersifat insidental; agar pengelola perguruan tinggi menyusun kebijakan penggunaan AI yang bertingkat dan transparan; serta agar dosen menambahkan persyaratan bukti verifikasi pada penugasan sebagai mekanisme pencegahan yang sederhana namun efektif. Keterbatasan kegiatan terletak pada jumlah peserta yang terbatas, ketiadaan kelompok pembanding, dan rentang pengukuran dampak yang hanya empat pekan pascapelatihan. Kegiatan lanjutan disarankan menggunakan rancangan dengan kelompok pembanding dan pengukuran jangka panjang untuk menguji ketahanan perubahan perilaku yang teramati.

## 11. UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Rektor Institut Teknologi Bisnis AAS Indonesia atas dukungan kelembagaan dan izin pelaksanaan kegiatan, kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat Institut Teknologi Bisnis AAS Indonesia atas dukungan

pendanaan dan fasilitasi administratif, serta kepada Ketua Program Studi Informatika atas dukungan koordinasi dengan mahasiswa. Ucapan terima kasih juga disampaikan kepada rekan-rekan dosen yang berperan sebagai fasilitator, kepada mahasiswa asisten yang membantu kelancaran teknis, dan terutama kepada 42 mahasiswa peserta yang mengikuti seluruh rangkaian kegiatan dengan antusiasme dan keterbukaan yang tinggi.

## DAFTAR PUSTAKA

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus, 15*(2), e35179. https://doi.org/10.7759/cureus.35179

Athaluri, S. A., Manthena, S. V., Kesapragada, V. S. R. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. Cureus, 15(4), e37432. https://doi.org/10.7759/cureus.37432

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. Information, 16(4), 296. https://doi.org/10.3390/info16040296

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20(1), 38. https://doi.org/10.1186/s41239-023-00408-3

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? Smart Learning Environments, 10(1), 60. https://doi.org/10.1186/s40561-023-00269-3

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Currie, G. M. (2023). Academic integrity and artificial intelligence: Is ChatGPT hype, hero or heresy? Seminars in Nuclear Medicine, 53(5), 719–730. https://doi.org/10.1053/j.semnuclmed.2023.04.008

Day, T. (2023). A preliminary investigation of fake peer-reviewed citations and references generated by ChatGPT. The Professional Geographer, 75(6), 1024–1027. https://doi.org/10.1080/00330124.2023.2190373

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., … Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Emsley, R. (2023). ChatGPT: These are not hallucinations – they're fabrications and falsifications. Schizophrenia, 9(1), 52. https://doi.org/10.1038/s41537-023-00379-4

European Commission, Directorate-General for Education, Youth, Sport and Culture. (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators. Publications Office of the European Union. https://data.europa.eu/doi/10.2766/153756

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. Innovations in Education and Teaching International, 61(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Giray, L. (2023). Prompt engineering with ChatGPT: A guide for academic writers. Annals of Biomedical Engineering, 51(12), 2629–2633. https://doi.org/10.1007/s10439-023-03272-4

Hou, C., Zhu, G., & Sudarshan, V. (2025). The role of critical thinking on undergraduates' reliance behaviours on generative AI in problem-solving. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13618

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., … Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. ACM Transactions on Information Systems, 43(2), 1–55. https://doi.org/10.1145/3703155

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., … Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), 1–38. https://doi.org/10.1145/3571730

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

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia. (2023). Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 53 Tahun 2023 tentang Penjaminan Mutu Pendidikan Tinggi. Kemendikbudristek RI.

Kementerian Pendidikan Nasional Republik Indonesia. (2010). Peraturan Menteri Pendidikan Nasional Nomor 17 Tahun 2010 tentang Pencegahan dan Penanggulangan Plagiat di Perguruan Tinggi. Kemendiknas RI.

Klingbeil, A., Grützner, C., & Schreck, P. (2024). Trust and reliance on AI — An experimental study on the extent and costs of overreliance on AI. Computers in Human Behavior, 160, 108352. https://doi.org/10.1016/j.chb.2024.108352

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), 410. https://doi.org/10.3390/educsci13040410

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? Library Hi Tech News, 40(3), 26–29. https://doi.org/10.1108/LHTN-01-2023-0009

McGowan, A., Gui, Y., Dobbs, M., Shuster, S., Cotter, M., Selloni, A., … Corcoran, C. M. (2023). ChatGPT and Bard exhibit spontaneous citation fabrication during psychiatry literature search. Psychiatry Research, 326, 115334. https://doi.org/10.1016/j.psychres.2023.115334

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. Education Sciences, 13(9), 856. https://doi.org/10.3390/educsci13090856

OECD. (2026). OECD digital education outlook 2026: Exploring effective uses of generative AI in education. OECD Publishing. https://doi.org/10.1787/062a7394-en

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching and Learning Practice, 20(2). https://doi.org/10.53761/1.20.02.07

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. Journal of University Teaching and Learning Practice, 21(06). https://doi.org/10.53761/q3azde36

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. Applied Sciences, 13(9), 5783. https://doi.org/10.3390/app13095783

Sun, Y., Sheng, D., Zhou, Z., & Wu, Y. (2024). AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content. Humanities and Social Sciences Communications, 11, 1278. https://doi.org/10.1057/s41599-024-03811-x

Tiller, N. B., Marcon, A. R., Zenone, M., Kidd, K. E., Jeukendrup, A. E., Master, Z., & Caulfield, T. (2026). Generative artificial intelligence-driven chatbots and medical misinformation: An accuracy, referencing and readability audit. BMJ Open, 16(4). https://doi.org/10.1136/bmjopen-2025-112695

Topaz, M., Roguin, N., Gupta, P., Zhang, Z., & Peltonen, L.-M. (2026). Fabricated citations: An audit across 2·5 million biomedical papers. The Lancet, 407(10541), 1779–1781. https://doi.org/10.1016/S0140-6736(26)00603-3

UNESCO. (2023). Guidance for generative AI in education and research (F. Miao & W. Holmes, Eds.). United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000386693

UNESCO. (2024). AI competency framework for students. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000391105

Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 tentang Pendidikan Tinggi. (2012). Lembaran Negara Republik Indonesia Tahun 2012 Nomor 158.

Wach, K., Duong, C. D., Ejdys, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszkiewicz, J., & Ziemba, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. Entrepreneurial Business and Economics Review, 11(2), 7–30. https://doi.org/10.15678/EBER.2023.110201

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. Scientific Reports, 13, 14045. https://doi.org/10.1038/s41598-023-41032-5

Wang, F., Li, N., Cheung, A. C. K., & Wong, G. K. W. (2025). In GenAI we trust: An investigation of university students' reliance on and resistance to generative AI in language learning. International Journal of Educational Technology in Higher Education, 22(1), 59. https://doi.org/10.1186/s41239-025-00547-9

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. Humanities and Social Sciences Communications, 12, 621. https://doi.org/10.1057/s41599-025-04787-y

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural

perspectives. International Journal of Educational Technology in Higher Education, 21(1), 21. https://doi.org/10.1186/s41239-024-00453-6

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(1), 28. https://doi.org/10.1186/s40561-024-00316-7