Integrative Perspectives of Social and Science Journal (IPSSJ)
ipssj.com ippsj2024@gmail.com E-ISSN: 3064-4011

CHATBOT KECERDASAN BUATAN DALAM PEMBELAJARAN MAHASISWA INFORMATIKA: TINJAUAN LITERATUR KOMPARATIF

Noval Pramudia Ramadhan[1*], Tata Sutabri[2]

[1,2] Program Studi Teknik Informatika, Institut Teknologi Indonesia, Tangerang Selatan, Indonesia
*Corresponding email: novalpramudia102@gmail.com, tata.sutabri@gmail.com

Abstrak: Meluasnya kecerdasan buatan generatif membuat mahasiswa informatika kian sering menjadikan chatbot seperti ChatGPT sebagai teman belajar, sementara bukti mengenai efektivitasnya bersumber dari penelitian yang rancangan dan indikatornya tidak seragam. Artikel ini bertujuan membandingkan tiga studi Indonesia terbitan 2025–2026 yang menelaah pemanfaatan chatbot kecerdasan buatan oleh mahasiswa informatika. Pendekatan yang dipakai adalah kajian literatur komparatif: tiga artikel dipilih berdasarkan lima kriteria, datanya diekstraksi ke dalam matriks yang seragam, kemudian dianalisis melalui ringkasan deskriptif, perbandingan lintas studi, dan sintesis tematik. Survei melaporkan penggunaan ChatGPT sebesar 96,7%, kemudahan 81,31%, manfaat 78,63%, keberhasilan 77,30%, dan ketergantungan pada kategori netral (54,79%). Sebaliknya, eksperimen berpasangan tidak menemukan perbedaan signifikan pada kualitas kode (semua $p > 0,05$), dan tidak ada bug yang dilaporkan pada kode. Hasil pembahasan mengindikasikan bahwa ChatGPT lebih berfungsi sebagai pendamping pemahaman dan pemercepat pekerjaan ketimbang penentu kualitas luaran. Dengan demikian, pemanfaatannya perlu proporsional, kritis, dan berintegritas akademik.

Kata kunci: chatbot kecerdasan buatan; ChatGPT; kajian literatur komparatif; mahasiswa informatika; pembelajaran pemrograman

Abstract: *The rise of generative AI has prompted informatics students to rely on chatbots such as ChatGPT as study companions, although the available evidence on their effectiveness originates from studies with differing designs and indicators. This article compares three Indonesian studies published in 2025–2026 that examine how informatics students use AI chatbots. A comparative literature review was employed: three articles were chosen according to five criteria, their data were extracted into a uniform matrix, and the findings were examined through descriptive summaries, cross-study comparison, and thematic synthesis. The surveys report ChatGPT use of 96.7%, ease of use of 81.31%, benefits of 78.63%, success of 77.30%, and neutral dependence (54.79%). A paired experiment detected no significant difference in code quality (all $p > 0.05$), and no bugs were reported in the programs. ChatGPT therefore functions more as an aid to comprehension and productivity than as a demonstrated determinant of output quality, and its use should be proportionate, critical, and attentive to academic integrity.*

Keywords: *artificial intelligence chatbot; ChatGPT; comparative literature review; informatics students; programming education*

PENDAHULUAN

Hadirnya kecerdasan buatan generatif mengubah cara mahasiswa mencari penjelasan, mengasah keterampilan, dan mengerjakan tugas. ChatGPT merupakan chatbot berarsitektur *Generative Pre-trained Transformer* yang diluncurkan OpenAI pada akhir 2022; perangkat ini mampu menjawab pertanyaan, menguraikan konsep, serta membangkitkan teks maupun kode program yang sesuai dengan konteks (Hendrik dkk., 2025; Lo, 2023; Utami dkk., 2026). Kapabilitas tersebut relevan bagi mahasiswa informatika, sebab kurikulum mereka sarat mata kuliah yang bersifat abstrak dan teknis, misalnya algoritma, struktur data, dan pemrograman. Mata kuliah pemrograman tidak sekadar melatih penulisan kode; mata kuliah ini juga menumbuhkan kemampuan berpikir komputasional (*computational thinking*), yakni kemampuan bernalar secara algoritmik dan memecahkan masalah secara sistematis (Makna dkk., 2026). Berbagai tinjauan pada bidang ini memaparkan peluang, seperti umpan balik yang dipersonalisasi dan penjelasan interaktif, di samping ancaman berupa kecurangan dan melemahnya pemikiran kritis (Kasneci dkk., 2023; Rahman & Watanobe, 2023).

Pergeseran ini membawa konsekuensi yang besar bagi mahasiswa informatika. Asisten percakapan dapat diakses kapan saja dan mampu menanggapi pertanyaan yang masih setengah jadi mengenai pesan kesalahan lewat penjelasan, contoh, dan potongan kode yang sudah

diperbaiki. Namun, kemudahan serupa berpotensi memangkas proses bergulat yang produktif, padahal proses itulah yang lazimnya menjadi jalan bagi pemula untuk membangun keterampilan pemrograman. Sejak awal, para pendidik komputasi menyadari bahwa model pembangkit kode menggugat asumsi lama tentang cara pemrograman pengantar diajarkan, sehingga kehadirannya membawa tantangan sekaligus peluang (Denny dkk., 2024). Sebuah studi metode campuran yang melibatkan 47 responden survei dan 8 orang yang diwawancarai mendapati bahwa perangkat kecerdasan buatan generatif cepat diadopsi, tetapi belum sepenuhnya menggantikan sumber bantuan tradisional, dan manfaat yang diperoleh mahasiswa tampaknya bergantung pada kemampuan mereka memanfaatkan perangkat tersebut (Hou dkk., 2024).

Manfaat tersebut diiringi oleh kekhawatiran. Risiko ketergantungan dan penyalahgunaan ditunjukkan oleh Hendrik dkk. (2025); Makna dkk. (2026) mencatat adanya kecemasan terhadap plagiarisme dan menurunnya pemikiran kritis; sementara Utami dkk. (2026) memperlihatkan bahwa pekerjaan yang lebih cepat belum tentu menghasilkan kode yang lebih baik. Tinjauan yang lebih luas sepakat bahwa ketergantungan berlebihan pada sistem dialog kecerdasan buatan dapat memengaruhi pemikiran kritis, pengambilan keputusan, dan penalaran analitis (Zhai dkk., 2024), serta bahwa integritas akademik tetap menjadi perhatian di pendidikan tinggi (Cotton dkk., 2024). Oleh karena itu, efektivitas chatbot kecerdasan buatan tidak cukup dinilai dari persepsi pengguna saja: persepsi, pola penggunaan, dan kualitas hasil kerja perlu dikaji secara bersamaan.

Masalah penelitian ini bersumber dari kenyataan bahwa studi-studi Indonesia yang tersedia menerapkan pendekatan dan indikator yang beragam: survei berskala kecil (Hendrik dkk., 2025), survei dengan pengujian instrumen (Makna dkk., 2026), dan eksperimen kualitas kode (Utami dkk., 2026). Setiap studi menjawab pertanyaan yang berlainan, sehingga para pendidik informatika belum memperoleh gambaran terpadu mengenai apa yang sesungguhnya ditunjukkan oleh bukti. Tinjauan terdahulu telah menjangkau pendidikan secara umum (Albadarin dkk., 2024; Lo, 2023; Montenegro-Rueda dkk., 2023) maupun pendidikan pemrograman secara khusus (Alanazi dkk., 2025; Baig dkk., 2026; Manorat dkk., 2025). Penelitian ini melengkapinya dengan menjajarkan tiga studi Indonesia yang rancangannya berbeda. Nilai kajian ini terletak bukan pada banyaknya studi, melainkan pada kontras antarstudi: dua studi merekam apa yang dipikirkan dan dilakukan mahasiswa, sedangkan satu studi merekam apa yang dihasilkan mahasiswa, dan hanya pembacaan gabungan yang mampu memperlihatkan apakah kedua jenis bukti itu selaras.

Penyelesaian masalah tersebut direncanakan dalam tiga tahap. Tahap pertama mendeskripsikan ketiga studi dengan format yang seragam. Tahap kedua membandingkan studi-studi itu menurut rancangan, jenis bukti, temuan, kebaruan, dan keterbatasan melalui matriks perbandingan. Tahap ketiga mensintesis temuan secara tematik guna memperlihatkan titik temu dan titik beda antara persepsi dan bukti objektif. Fondasi teoretisnya bersumber dari penelitian tentang ChatGPT dalam pendidikan (Kasneci dkk., 2023; Lo, 2023), penelitian tentang kecerdasan buatan dalam pendidikan pemrograman (Baig dkk., 2026; Yilmaz & Karaoglan Yilmaz, 2023), serta penelitian tentang penerimaan teknologi (Habibi dkk., 2023; Strzelecki, 2024).

Kajian ini dipandu oleh empat pertanyaan penelitian. PP1: seberapa tinggi penggunaan chatbot kecerdasan buatan yang dilaporkan mahasiswa informatika, dan bagaimana persepsi mereka terhadapnya? PP2: bagaimana ketiga studi tersebut dibandingkan dari segi metodologi, objek, dan temuan? PP3: apa kekuatan dan kebaruan setiap studi? PP4: apa implikasi sintesis itu terhadap efektivitas chatbot kecerdasan buatan dan terhadap mahasiswa informatika? Penelitian ini bertujuan menjawab seluruh pertanyaan tersebut melalui kajian literatur komparatif.

Kajian ini diantisipasi memperlihatkan bahwa manfaat yang dirasakan tergolong kuat, sedangkan bukti objektif atas peningkatan kualitas luaran masih terbatas. Kegunaannya bersifat praktis: mahasiswa memperoleh gambaran tentang fungsi chatbot yang bermanfaat, dosen dapat merancang tugas yang terukur, institusi dapat menyusun pedoman berbasis bukti, dan peneliti dapat mengenali rancangan yang belum tersedia. Susunan artikel ini adalah sebagai berikut. Landasan Teori memaparkan kerangka teoretis dan literatur terkait, Metode menguraikan metodologi kajian, Hasil dan Pembahasan menyajikan sekaligus membahas hasil, dan Simpulan mengakhiri artikel.

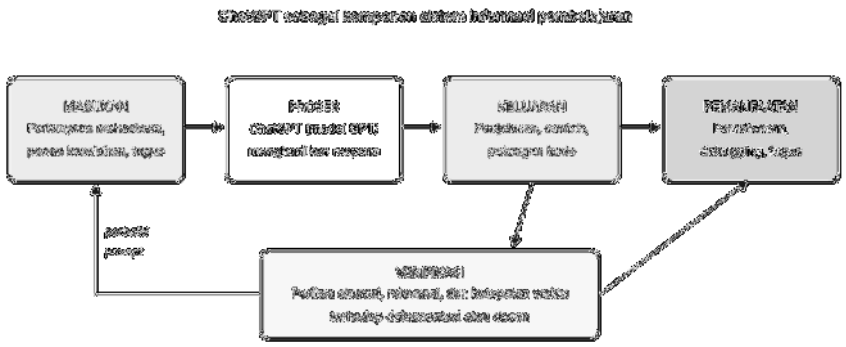
LANDASAN TEORI

Chatbot Kecerdasan Buatan Generatif dalam Pendidikan Tinggi

Tinjauan awal mengenai ChatGPT di bidang pendidikan melukiskannya sebagai asisten yang potensial bagi pengajar sekaligus tutor virtual bagi mahasiswa, namun mengingatkan bahwa perangkat ini dapat menghasilkan informasi yang keliru dan luput dari deteksi plagiarisme (Lo, 2023). Sebuah tinjauan sistematis atas dua belas studi melaporkan adanya dampak positif terhadap pengajaran dan pembelajaran, seraya menegaskan bahwa guru memerlukan pelatihan agar dapat memakai perangkat ini secara tepat (Montenegro-Rueda dkk., 2023). Tinjauan sistematis lainnya mencakup empat belas studi empiris mengenai cara mahasiswa dan guru memanfaatkan ChatGPT (Albadarin dkk., 2024). Sebuah makalah posisi berargumen bahwa pembelajar dan pengajar membutuhkan kompetensi dan literasi baru untuk memahami teknologi ini beserta keterbatasannya (Kasneci dkk., 2023), dan bahwa peluang seperti umpan balik yang dipersonalisasi dan akses yang lebih luas perlu ditimbang terhadap ancaman seperti kecurangan dan menurunnya pemikiran kritis (Rahman & Watanobe, 2023).

Penelitian yang berfokus pada mahasiswa memperkaya kajian dengan perspektif pengguna. Pendapat mahasiswa mengenai kemampuan dan keterbatasan ChatGPT ditelusuri melalui analisis tematik yang dilanjutkan dengan survei (Shoufan, 2023). Secara agregat, rangkaian studi ini mengindikasikan bahwa di pendidikan tinggi chatbot berperan secara bersamaan sebagai tutor, alat produktivitas, dan potensi jalan pintas. Peran yang menonjol bergantung pada jenis tugas, karakteristik mahasiswa, dan aturan mata kuliah; karena itulah satu indikator tunggal, seperti kepuasan atau adopsi, tidak dapat mewakili efektivitas.

Jika ditinjau sebagai komponen sistem informasi pembelajaran, chatbot menerima masukan dari mahasiswa (pertanyaan, pesan kesalahan, tugas), mengolahnya, lalu mengembalikan keluaran (penjelasan, contoh, potongan kode) yang selanjutnya dimanfaatkan oleh mahasiswa (Sutabri, 2012b, 2014). Bernilai-tidaknya keluaran itu ditentukan oleh akurasi, ketepatan waktu, dan relevansi informasinya, sehingga verifikasi oleh mahasiswa merupakan langkah yang wajib, bukan pilihan. Alur ini diringkas pada Gambar 1.



Gambar 1. Alur pemakaian ChatGPT sebagai bagian dari sistem informasi pembelajaran

Kecerdasan Buatan dalam Pendidikan Pemrograman

Sebuah tinjauan pelingkupan (*scoping review*) terhadap 59 artikel terbitan 2022–2025 mengidentifikasi peran ChatGPT sebagai asisten pemrograman serta sebagai alat penilaian dan umpan balik otomatis, disertai risiko ketergantungan berlebihan, menurunnya pemikiran kritis, keterbatasan akurasi, dan persoalan integritas akademik (Baig dkk., 2026). Cakupan yang lebih luas ditempuh oleh tinjauan atas 119 makalah dari 2012 hingga 2024, yang memetakan perangkat kecerdasan buatan dan pembelajaran mesin pada seluruh proses pedagogis mata kuliah pemrograman (Manorat dkk., 2025). Adapun meta-analisis terhadap 35 studi terkendali melaporkan waktu penyelesaian tugas yang lebih singkat dan skor kinerja yang lebih tinggi pada pembelajaran berbantuan kecerdasan buatan, tetapi tidak ditemukan keunggulan yang signifikan secara statistik pada keberhasilan belajar maupun kemudahan pemahaman (Alanazi dkk., 2025).

Eksperimen terkendali memperlihatkan kecenderungan serupa. Dibandingkan dengan kelompok kontrol, mahasiswa yang memakai ChatGPT dalam praktik pemrograman mingguan menampilkan *computational thinking*, efikasi diri pemrograman, dan motivasi yang lebih tinggi (Yilmaz & Karaoglan Yilmaz, 2023). Pemula yang diberi akses ke generator kode berbasis kecerdasan buatan mencapai kinerja penulisan kode yang lebih baik tanpa kemunduran dalam kemampuan memodifikasi kode secara manual (Kazemitabaar dkk., 2023). Penelitian empiris lainnya mengajukan pertanyaan apakah mode pemrograman yang difasilitasi ChatGPT memengaruhi perilaku, kinerja, dan persepsi pemrograman mahasiswa perguruan tinggi (Sun dkk., 2024). Mahasiswa sarjana yang memanfaatkan ChatGPT untuk tugas pengodean mengaku lebih produktif dan kurang frustrasi (da Silva dkk., 2024), sedangkan persepsi mahasiswa ilmu komputer ditelaah melalui studi deskriptif dan korelasional (Singh dkk., 2023). Dari perspektif kode, GitHub Copilot dapat menjadi aset bagi pengembang berpengalaman, tetapi menjadi beban bagi pemula yang tidak mampu menyaring solusi yang mengandung bug (Moradi Dakhel dkk., 2023).

Penerimaan Teknologi sebagai Lensa Interpretasi

Riset penerimaan teknologi menerangkan alasan mahasiswa mengadopsi suatu perangkat sekaligus alasan mengapa adopsi semata tidak membuktikan adanya peningkatan belajar. Studi penerimaan yang berlandaskan teori terpadu penerimaan dan penggunaan teknologi (UTAUT) melaporkan bahwa kondisi yang memfasilitasi menjadi penentu terkuat niat menggunakan ChatGPT pada mahasiswa perguruan tinggi Indonesia (Habibi dkk., 2023), dan bahwa ekspektasi kinerja berkaitan dengan niat perilaku, yang selanjutnya berkaitan dengan penggunaan aktual (Bazelais dkk., 2024). Model UTAUT2 yang diperluas turut menjelaskan penerimaan mahasiswa di konteks lain (Strzelecki, 2024), termasuk Vietnam (Maheshwari, 2024). Model-model tersebut berkutat pada niat dan penggunaan: ia menjelaskan siapa yang cenderung memakai chatbot beserta alasannya, tetapi tidak mengukur pengaruh chatbot terhadap kualitas hasil kerja mahasiswa. Kesenjangan inilah yang membedakan ketiga studi yang ditinjau, dan kesenjangan ini pula yang melandasi pembacaan gabungan atas bukti berbasis persepsi dan bukti objektif dalam artikel ini.

Bukti Berbasis Persepsi dan Bukti Objektif

Kajian ini membedakan bukti ke dalam dua jenis. Bukti berbasis persepsi mencakup penggunaan, sikap, dan pengalaman yang dilaporkan sendiri oleh responden melalui kuesioner dan pertanyaan terbuka. Jenis bukti ini bernilai untuk memetakan adopsi serta memahami alasan mahasiswa beralih ke chatbot, dan menjadi satu-satunya cara yang praktis untuk menangkap aspek afektif seperti kepercayaan diri. Kelemahannya, manfaat yang dirasakan bisa saja berbeda dari manfaat yang terukur. Bukti objektif, di sisi lain, berupa luaran yang diukur, misalnya metrik kualitas kode hasil analisis statis. Bukti ini lebih kecil kemungkinannya terpengaruh bias laporan diri, tetapi cakupannya cenderung lebih sempit, sebab sejumlah

metrik pada beberapa tugas tidak mampu merepresentasikan seluruh pengalaman belajar. Penilaian efektivitas chatbot yang berimbang dengan demikian memerlukan kedua jenis bukti, dan gagasan ini menjadi kerangka analisis komparatif pada bagian Hasil dan Pembahasan.

METODE

Desain Penelitian

Kajian ini berupa kajian literatur komparatif yang menempuh pendekatan kualitatif-deskriptif. Data primer tidak dihimpun dari responden; yang dianalisis adalah data sekunder berupa hasil penelitian yang telah dipublikasikan. Tinjauan pada bidang ini pada umumnya berpedoman pada prosedur berbasis PRISMA (Alanazi dkk., 2025; Baig dkk., 2026), dan penulis mengadopsi prinsip transparansi dalam pernyataan PRISMA 2020 (Page dkk., 2021), yakni kriteria seleksi, butir yang diekstraksi, dan prosedur sintesis dinyatakan secara eksplisit. Karena tinjauan atas tiga studi tidak tergolong tinjauan sistematis penuh, diagram alir PRISMA tidak dilaporkan. Desain komparatif dipilih sebab masalah penelitian menyangkut kontras antarstudi yang rancangannya berbeda, bukan penggabungan efek; penggabungan semacam itu memang tidak sesuai untuk survei dan eksperimen yang mengukur konstruk berlainan.

Empat pertanyaan penelitian pada bagian Pendahuluan menjadi acuan penyusunan kajian ini: PP1 dijawab berdasarkan dua survei, PP2 melalui perbandingan lintas studi, PP3 melalui penilaian kekuatan dan kebaruan, dan PP4 melalui sintesis tematik.

Strategi Pencarian

Pengumpulan artikel dilakukan lewat pencarian daring pada repositori jurnal. Kata kunci pencarian yang dipakai ialah “ChatGPT”, “mahasiswa informatika”, “pembelajaran pemrograman”, dan “code assistant”. Kata kunci tersebut dipadukan menggunakan operator Boolean AND ke dalam tiga rangkaian pencarian: (1) “ChatGPT” AND “mahasiswa informatika”; (2) “ChatGPT” AND “pembelajaran pemrograman”; serta (3) “ChatGPT” AND “code assistant”. Hasil pencarian dibatasi pada artikel terbitan 2025–2026. Judul dan abstrak yang terjaring diseleksi menurut kriteria pada subbab Kriteria Seleksi, kemudian teks lengkap setiap artikel yang melewati penyaringan awal dibaca secara penuh.

Kriteria Seleksi

Kajian ini mengambil tiga artikel jurnal ilmiah berbahasa Indonesia sebagai objek. Seleksi artikel didasarkan pada lima kriteria: (1) artikel mengulas chatbot kecerdasan buatan atau ChatGPT dalam konteks pembelajaran; (2) respondennya merupakan mahasiswa informatika atau mahasiswa program studi teknologi informasi yang terkait; (3) artikel terbit pada 2025–2026; (4) metodologi dan hasilnya dinyatakan secara eksplisit; dan (5) artikel mewakili rancangan yang berbeda, yaitu survei berskala kecil, survei dengan pengujian instrumen, dan eksperimen. Kriteria eksklusi merupakan kebalikan dari kriteria inklusi. Kedua rangkaian kriteria tersebut, sebagaimana diterapkan, disajikan pada Tabel 1.

Tabel 1. Kriteria inklusi dan eksklusi yang digunakan dalam seleksi artikel

Kode	Kriteria inklusi	Kriteria eksklusi
1	Mengulas chatbot kecerdasan buatan atau ChatGPT dalam konteks pembelajaran	Tidak mengulas chatbot dalam konteks pembelajaran
2	Respondennya mahasiswa informatika atau mahasiswa program studi teknologi informasi yang terkait	Respondennya berasal dari bidang yang tidak terkait
3	Terbit pada 2025–2026	Terbit sebelum 2025
4	Metodologi dan hasilnya dinyatakan secara eksplisit	Metodologi atau hasilnya tidak dinyatakan secara eksplisit

Kode	Kriteria inklusi	Kriteria eksklusi
5	Merepresentasikan rancangan yang berlainan dari artikel terpilih lain (survei berskala kecil, survei dengan pengujian instrumen, atau eksperimen)	Merupakan pengulangan rancangan yang telah terwakili

Setelah kriteria tersebut diterapkan, tersisa tiga studi. Identitas ketiga jurnal terpilih disajikan pada Tabel 2. Di dalam teks, artikel dirujuk menurut penulis dan tahun, sedangkan rincian publikasi yang diperlukan untuk menelusurinya dicantumkan pada Tabel 2.

Tabel 2. Identitas jurnal rujukan

Aspek	Hendrik dkk. (2025)	Makna dkk. (2026)	Utami dkk. (2026)
Judul	Analisis Pemanfaatan AI ChatGPT dalam Membantu Proses Belajar Mahasiswa Pendidikan Informatika di Universitas Citra Bangsa	Persepsi Penggunaan ChatGPT sebagai Alat Bantu Belajar Mata Kuliah Pemrograman di Institut Widya Pratama	Analisis Dampak ChatGPT sebagai Code Assistant terhadap Kualitas Kode Mahasiswa Informatika UTS
Publikasi	JUPE: Jurnal Pendidikan Mandala, Vol. 10 No. 2, Juni 2025	Jurnal Teknologi Terpadu, Vol. 12 No. 1, 2026	Management of Information System Journal, Vol. 4 No. 3, Juli 2026
Lokasi penelitian	Program Studi Pendidikan Informatika, Universitas Citra Bangsa	Institut Widya Pratama, Pekalongan, Jawa Tengah	Program Studi Informatika, Universitas Teknologi Sumbawa
Objek	Pemanfaatan ChatGPT untuk memahami materi, *debugging*, dan menyelesaikan tugas; efektivitas yang dirasakan	Persepsi terhadap ChatGPT dalam mata kuliah pemrograman: kemudahan penggunaan, manfaat, ketergantungan, dan keberhasilan	Kualitas kode yang ditulis secara manual dibandingkan dengan kode berbantuan ChatGPT
Metodologi	Survei kuantitatif; kuesioner skala Likert 15 butir; analisis deskriptif	Survei kuantitatif; kuesioner skala Likert 5 poin dan pertanyaan terbuka; uji validitas dan reliabilitas; *purposive sampling*	Eksperimen kuantitatif dengan dua kondisi pada kelompok yang sama; SonarQube; uji Shapiro–Wilk, *paired t-test*, Wilcoxon
Sampel	30 mahasiswa (10 per angkatan)	91 mahasiswa dari populasi 956	Mahasiswa informatika; dua tugas pemrograman dasar

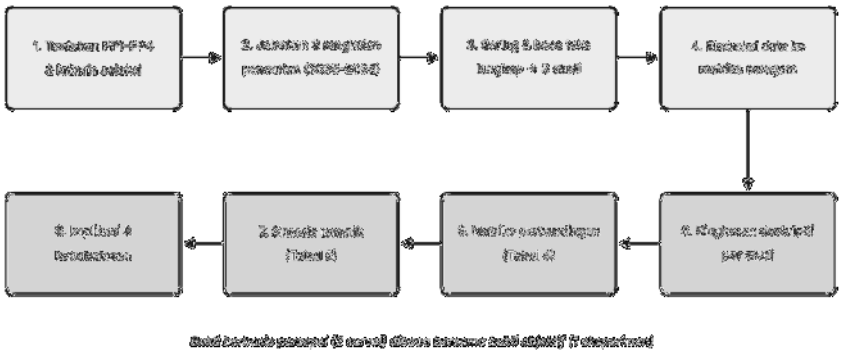
Ekstraksi Data

Tiap artikel dibaca secara menyeluruh, lalu informasinya diekstraksi ke dalam sebuah matriks yang seragam. Matriks itu memuat identitas, objek, lokasi penelitian, metodologi, sampel, instrumen, jenis bukti, temuan utama, kebaruan, dan keterbatasan. Hasil kuantitatif dicatat persis sebagaimana yang dilaporkan penulis aslinya, mencakup persentase, nilai signifikansi, dan ukuran efek; tidak ada angka yang dihitung ulang maupun dibulatkan dengan cara yang berbeda. Jika suatu artikel melaporkan penyesuaian metodologis, misalnya perubahan teknik pengambilan sampel, penyesuaian itu dicatat berdampingan dengan rencana awalnya.

Prosedur Analisis

Prosedur dijalankan secara kronologis. Langkah pertama adalah penetapan pertanyaan penelitian dan kriteria seleksi. Langkah kedua adalah pencarian dan penyaringan artikel berdasarkan kriteria tersebut, yang menyisakan tiga studi. Pada langkah ketiga, setiap artikel dibaca secara penuh dan informasinya diekstraksi ke dalam format seragam yang diuraikan di atas. Langkah keempat berupa penyusunan ringkasan deskriptif untuk setiap studi. Pada

langkah kelima, matriks perbandingan dibangun berdasarkan fokus, rancangan, jenis bukti, ukuran sampel, instrumen, dan temuan. Langkah keenam adalah sintesis tematik untuk mengidentifikasi pola lintas studi mengenai manfaat, risiko, dan kondisi yang memengaruhi efektivitas. Langkah terakhir adalah penurunan implikasi bagi mahasiswa, dosen, institusi, dan peneliti. Urutan ini, yang bergerak dari identifikasi masalah melalui pengumpulan data hingga analisis dan rekomendasi, sejalan dengan logika umum analisis sistem informasi (Sutabri, 2012a) dan digambarkan pada Gambar 2.



Gambar 2. Tahapan prosedur kajian literatur komparatif

Kerangka Sintesis dan Evaluasi Hasil

Evaluasi hasil kajian dilakukan melalui tiga cara. Cara pertama, setiap temuan yang dilaporkan digolongkan berdasarkan jenis buktinya, yakni berbasis persepsi (survei) atau objektif (kualitas kode yang diukur), lalu dibaca bersama ukuran sampel serta bukti validitas dan reliabilitas yang disajikan penulis aslinya. Cara kedua, kebaruan ditentukan dari ada tidaknya pertanyaan, rancangan, atau ukuran yang disumbangkan suatu studi dan tidak dimiliki dua studi lainnya. Cara ketiga, suatu tema dalam sintesis dinilai terdukung dengan baik hanya jika sekurang-kurangnya dua studi mengarah pada hal yang sama; tema yang bertumpu pada satu studi dilaporkan sebagai bukti satu studi. Pada tahap akhir, temuan dibandingkan dengan literatur yang lebih luas tentang ChatGPT dalam pendidikan dan pemrograman (Alanazi dkk., 2025; Baig dkk., 2026; Lo, 2023) guna menguji konsistensinya.

Mengingat instrumen risiko bias yang formal tidak dipakai, mutu setiap studi ditaksir dari indikator yang terlihat pada artikelnya sendiri, yaitu ukuran sampel, teknik pengambilan sampel, bukti validitas dan reliabilitas instrumen, serta kesesuaian uji statistik. Penaksiran ini berfungsi sebagai kualifikasi terhadap kekuatan setiap simpulan, bukan sebagai dasar untuk menyingkirkan suatu studi.

HASIL DAN PEMBAHASAN

Ringkasan Setiap Studi

Hendrik dkk. (2025). Hasil survei terhadap 30 mahasiswa pendidikan informatika di Universitas Citra Bangsa mengindikasikan bahwa 96,7% responden (29 dari 30) memakai ChatGPT, terutama untuk memahami materi, menyelesaikan tugas, dan belajar secara mandiri. Para mahasiswa mengaku dapat menghemat waktu, lebih mudah memperoleh klarifikasi konseptual, serta mendapat penjelasan yang tersusun lebih terstruktur. Penulisnya turut menyimpulkan bahwa interaksi dengan ChatGPT menumbuhkan kepercayaan diri mahasiswa untuk bertanya tanpa tekanan diskusi kelas, dan menganjurkan pemakaian yang proporsional dan kritis guna mencegah ketergantungan kognitif.

Makna dkk. (2026). Survei ini mengikutsertakan 91 mahasiswa Institut Widya Pratama dari populasi sebanyak 956 orang. Ukuran sampel semula ditentukan dengan rumus Slovin

pada margin kesalahan 10%, tetapi pada praktiknya disesuaikan menjadi *purposive sampling*, dan data dihimpun pada rentang 28 Juli hingga 4 Agustus 2025. Responden berjenis kelamin laki-laki mencapai 62,6%; sebanyak 45,1% berkuliah di Sistem Informasi, 44% di Teknik Informatika, dan 11% di Komputerisasi Akuntansi; mayoritas duduk di semester tujuh (65,94%). Seluruh butir instrumen dinyatakan valid (r-tabel 0,204) dan reliabel (alfa Cronbach 0,713 hingga 0,866). Tujuan penggunaan yang paling lazim adalah mencari referensi atau penjelasan ulang (41,8%), disusul *debugging* (26,4%), pemahaman konsep (19,8%), dan penyelesaian tugas pengodean (12,1%). Pertanyaan terbuka menyingkap kesadaran terhadap risiko: saran yang paling sering dikemukakan ialah memperlakukan ChatGPT sebagai alat bantu, bukan kebutuhan utama (20,88%), serta tetap belajar mandiri (19,78%).

Utami dkk. (2026). Pada eksperimen di Universitas Teknologi Sumbawa, peserta diminta mengerjakan dua tugas pemrograman dasar yang telah divalidasi pakar, yakni menentukan jenis bilangan dan membuat program kasir, masing-masing satu kali tanpa ChatGPT dan satu kali dengan ChatGPT. Kode dianalisis oleh SonarQube berdasarkan tiga metrik: *Maintainability Index*, *Cyclomatic Complexity*, dan *Reliability* (jumlah bug). Oleh karena data yang tersedia kurang dari 50, normalitas diperiksa dengan uji Shapiro–Wilk, kemudian dilanjutkan dengan *paired t-test* atau uji peringkat bertanda Wilcoxon beserta ukuran efeknya. Tidak satu pun metrik menunjukkan perbedaan yang signifikan, dan para penulis melaporkan nol bug pada kedua kondisi, walaupun tabel rekapitulasi mereka mencatat satu bug pada satu entri yang ditulis secara manual. Secara deskriptif, kode berbantuan ChatGPT lebih konsisten dalam hal struktur, komentar, dan format keluaran, serta memperlihatkan kompleksitas siklomatik yang lebih tinggi; namun, hal ini tidak mengubah kualitas secara signifikan. Para penulis menduga bahwa tugas yang relatif mudah dan penguasaan keterampilan dasar yang memadai pada peserta menjadi penjelas hasil tersebut.

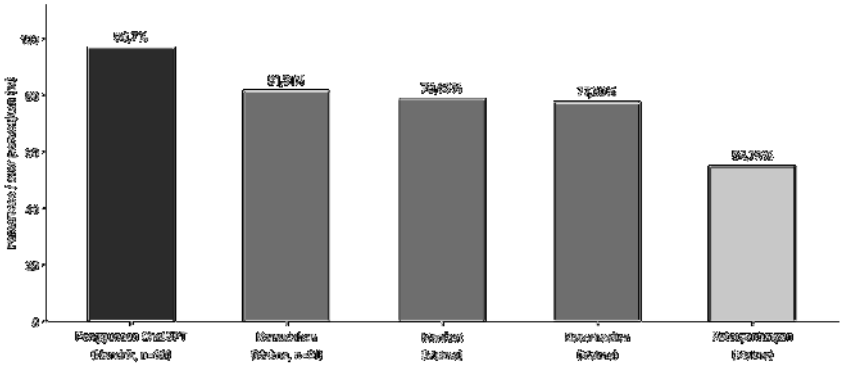
Tabel 3 merangkum angka-angka utama dari ketiga studi. Persentase pada survei merepresentasikan proporsi responden atau tingkat persetujuan, sedangkan nilai pada eksperimen berupa hasil uji signifikansi atas metrik kualitas kode; kedua kelompok angka itu karenanya menjawab pertanyaan yang berlainan dan tidak dapat dibandingkan secara langsung.

Tabel 3. Temuan pokok ketiga studi

Studi	Indikator/metrik	Hasil
Hendrik dkk. (2025)	Tingkat penggunaan ChatGPT	96,7% (29 dari 30 responden)
Hendrik dkk. (2025)	Manfaat yang dilaporkan	Efisiensi waktu, klarifikasi konsep, penjelasan terstruktur, kepercayaan diri untuk bertanya
Makna dkk. (2026)	Kemudahan penggunaan	81,31% (sangat setuju)
Makna dkk. (2026)	Manfaat dalam menyelesaikan tugas	78,63% (setuju)
Makna dkk. (2026)	Ketergantungan	54,79% (netral)
Makna dkk. (2026)	Keberhasilan penggunaan	77,30% (setuju)
Utami dkk. (2026)	*Maintainability Index*	p = 0,281 (Wilcoxon; tidak signifikan)
Utami dkk. (2026)	*Cyclomatic Complexity*	p = 0,837 (*paired t-test*; Cohen's d = 0,0562)
Utami dkk. (2026)	*Reliability* (jumlah bug)	p = 1,000 (Wilcoxon); tidak ada bug yang dilaporkan

Apabila Tabel 3 dicermati secara keseluruhan, hasil survei (PP1) melukiskan populasi mahasiswa yang sebagian besar telah mengadopsi ChatGPT dan menilainya mudah digunakan serta bermanfaat. Skor kemudahan penggunaan dan manfaat yang tinggi disertai skor ketergantungan pada kategori netral, yang mengisyaratkan bahwa mahasiswa tidak menggambarkan diri mereka sangat bergantung. Sebaliknya, hasil eksperimen tidak memperlihatkan perbedaan yang terukur. Ukuran efek yang sangat kecil pada *Cyclomatic*

Complexity (Cohen's d = 0,0562) serta nilai p sebesar 1,000 pada *Reliability*, yang merupakan akibat dari tidak adanya bug yang dilaporkan, berarti kedua kondisi praktis tidak dapat dibedakan pada metrik tersebut. Perbandingan indikator survei disajikan dalam bentuk diagram batang pada Gambar 3.



Gambar 3. Indikator survei menurut laporan Hendrik dkk. (2025) dan Makna dkk. (2026)

Analisis Komparatif

Ketiga studi dibandingkan pada beberapa dimensi dalam Tabel 4. Perbedaan yang paling mendasar ada pada jenis bukti: dua studi mengukur persepsi, sementara satu studi mengukur luaran secara objektif.

Tabel 4. Matriks komparasi ketiga studi

Dimensi	Hendrik dkk. (2025)	Makna dkk. (2026)	Utami dkk. (2026)
Fokus	Pemanfaatan, manfaat, dan efektivitas yang dirasakan	Kemudahan penggunaan, manfaat, ketergantungan, dan keberhasilan	Pengaruh terhadap kualitas kode
Rancangan	Survei deskriptif	Survei deskriptif dengan pengujian instrumen	Eksperimen berpasangan
Jenis bukti	Persepsi dan pengalaman	Persepsi dan jawaban atas pertanyaan terbuka	Metrik objektif SonarQube
Ukuran sampel	30	91	Terbatas (15 entri; 14 pasangan lengkap)
Simpulan inti	Positif; perlu mewaspadai ketergantungan kognitif	Positif; ketergantungan netral; sebagai alat bantu, bukan pengganti	Tidak terdapat perbedaan kualitas yang signifikan
Kebaruan	Dimensi kepercayaan diri dan tingkat adopsi	Kerangka empat indikator dan pengujian instrumen	Pengukuran luaran kualitas secara objektif
Keterbatasan	Sampel kecil; instrumen tidak diuji	*Purposive sampling*; hanya persepsi	Sampel dan tugas yang terbatas

Matriks tersebut memunculkan tiga pengamatan. Pertama, rancangan-rancangan itu bersifat saling melengkapi, bukan saling mengulang. Survei deskriptif Hendrik dkk. (2025) menyuguhkan keluasan adopsi pada kelompok kecil; survei Makna dkk. (2026) menambah kualitas pengukuran lewat sampel yang lebih besar dan pengujian instrumen; adapun eksperimen Utami dkk. (2026) menghadirkan luaran yang objektif. Kedua, ukuran sampel berselisih cukup jauh, sehingga bobot setiap simpulan pun berbeda. Ketiga, simpulan inti ketiganya selaras alih-alih bertentangan: survei menyimpulkan bahwa ChatGPT dipersepsikan positif dan perlu dipakai dengan hati-hati, sedangkan eksperimen menyimpulkan bahwa tidak ditemukan perbedaan kualitas pada tugas dasar. Ketidakselarasan itu hanya tampak di

permukaan, sebab survei tidak mengklaim adanya peningkatan kualitas dan eksperimen tidak mengklaim bahwa mahasiswa tidak merasakan manfaat.

Kekuatan dan Kebaruan

Hendrik dkk. (2025) menekankan dimensi afektif: ChatGPT dipandang sebagai ruang yang relatif aman untuk bertanya sehingga kepercayaan diri meningkat. Aspek ini tidak dibahas oleh dua studi lainnya. Studi ini pun memetakan tingkat adopsi secara langsung (96,7%). Pengamatan tersebut sejalan dengan temuan bahwa berkurangnya tekanan sosial termasuk salah satu pendorong mahasiswa komputasi memakai kecerdasan buatan generatif untuk mencari bantuan (Hou dkk., 2024). Keterbatasannya meliputi sampel yang kecil, analisis yang murni deskriptif, ketiadaan uji validitas dan reliabilitas, serta klaim adanya lompatan metodologis yang tidak ditopang oleh hasil belajar yang terukur.

Makna dkk. (2026) tergolong baru karena menggabungkan empat indikator dalam satu studi, memiliki sampel terbesar, dan menguji instrumennya. Pengukuran ketergantungan secara eksplisit menghasilkan gambaran yang lebih berimbang, dan temuan bahwa tujuan terlazim adalah mencari penjelasan ulang, bukan menyalin jawaban, mengisyaratkan pola penggunaan yang condong ke arah belajar. Keterbatasannya adalah beralihnya teknik sampling ke *purposive sampling* yang menurunkan daya generalisasi, data yang hanya berupa persepsi, dan responden yang berasal dari tiga program studi.

Utami dkk. (2026) menjadi satu-satunya studi yang mengukur luaran secara objektif, yaitu lewat analisis statis dan uji berpasangan pada orang yang sama, sehingga perbedaan kemampuan antarpeserta dapat diminimalkan. Temuannya mengimbangi persepsi positif yang muncul pada survei: manfaat yang dirasakan belum tentu terlihat sebagai peningkatan kualitas yang terukur. Keterbatasannya mencakup jumlah peserta yang sedikit, tugas yang berkompleksitas dasar, metrik yang hanya berjumlah tiga, serta kenyataan bahwa hasil yang tidak signifikan pada sampel kecil tidak berarti efek itu tidak ada, mengingat daya statistiknya rendah.

Sintesis Efektivitas bagi Mahasiswa Informatika

Pertama, sebagai pendamping pemahaman. Kedua survei memperlihatkan bahwa mahasiswa memanfaatkan ChatGPT untuk menjelaskan ulang materi, memahami konsep, dan memperbaiki kode (Hendrik dkk., 2025; Makna dkk., 2026). Pola tersebut selaras dengan peran kecerdasan buatan sebagai *scaffolding* yang membantu mahasiswa melewati kesulitan pada materi teknis (Makna dkk., 2026), serta dengan peran asisten pemrograman yang teridentifikasi dalam literatur yang lebih luas (Baig dkk., 2026).

Kedua, sebagai pemercepat kerja. Efisiensi waktu tercatat sebagai salah satu manfaat (Hendrik dkk., 2025), dan mayoritas responden sepakat bahwa tugas dapat diselesaikan lebih cepat (Makna dkk., 2026). Waktu tidak diukur dalam eksperimen, tetapi para penulisnya menafsirkan ChatGPT sebagai alat pemercepat, bukan pengungkit kualitas, sebab kualitas kode setara pada tugas dasar (Utami dkk., 2026). Di luar Indonesia, mahasiswa sarjana yang memakai ChatGPT untuk tugas pemrograman turut mengaku lebih produktif dan kurang frustrasi (da Silva dkk., 2024).

Ketiga, efektivitas terhadap kualitas belum terbukti. Persepsi yang positif tidak serta-merta berarti kualitas luaran meningkat; eksperimen tidak mendapati peningkatan maupun penurunan yang signifikan (Utami dkk., 2026). Klaim bahwa ChatGPT "meningkatkan pemahaman" masih bertumpu pada persepsi, sedangkan klaim mengenai kualitas hanya berlaku pada tugas sederhana. Utami dkk. (2026) juga merujuk penelitian lain yang menemukan duplikasi dan masalah keterpeliharaan yang lebih banyak pada aplikasi berbantuan kecerdasan buatan, serta hasil yang bergantung pada mutu *prompt* dan kompleksitas masalah, sehingga hasilnya mungkin berbeda pada tugas yang lebih kompleks. Bukti dari luar studi yang ditinjau beragam: meta-analisis melaporkan skor kinerja yang lebih tinggi dan waktu

penyelesaian yang lebih singkat, tetapi tanpa keunggulan yang signifikan pada keberhasilan belajar atau kemudahan pemahaman (Alanazi dkk., 2025), sementara eksperimen terkendali melaporkan peningkatan *computational thinking* (Yilmaz & Karaoglan Yilmaz, 2023) dan kinerja penulisan kode (Kazemitabaar dkk., 2023).

Keempat, risiko ketergantungan dan integritas akademik. Skor ketergantungan pada kategori netral (54,79%) mengindikasikan bahwa mahasiswa masih memanfaatkan buku, dosen, dan latihan mandiri (Makna dkk., 2026). Kendati demikian, ketiga studi sependapat bahwa risiko ketergantungan, menurunnya pemikiran kritis, dan ketidakjujuran akademik perlu dikelola, dan mahasiswa sendiri menyarankan agar keluaran ChatGPT diverifikasi (Makna dkk., 2026). Hal ini sejalan dengan tinjauan yang mengingatkan bahwa ChatGPT dapat menghasilkan informasi keliru dan lolos dari pendeteksi plagiarisme (Lo, 2023), serta bahwa ketergantungan berlebihan dapat melemahkan kemampuan kognitif (Zhai dkk., 2024).

Kelima, konteks menentukan hasil. Efektivitas ditentukan oleh jenis tugas, kemampuan awal mahasiswa, dan cara bertanya (*prompting*). Hasil nol bug dan ketiadaan perbedaan signifikan pada eksperimen kemungkinan terkait dengan peserta yang telah menguasai dasar pemrograman (Utami dkk., 2026), sementara survei memperlihatkan bahwa banyak mahasiswa beralih ke ChatGPT justru saat membutuhkan penjelasan ulang (Makna dkk., 2026). Riset penerimaan teknologi menguatkan peran konteks: kondisi yang memfasilitasi menjadi penentu niat terkuat pada mahasiswa perguruan tinggi Indonesia (Habibi dkk., 2023), dan hubungan antara ekspektasi kinerja, niat, dan penggunaan telah dikonfirmasi oleh studi lain (Bazelais dkk., 2024). Pengamatan bahwa mahasiswa yang mampu mengevaluasi respons model dan memadukannya ke dalam pemecahan masalah mereka sendiri tampak paling diuntungkan (Hou dkk., 2024) menunjuk ke arah yang sama: chatbot yang sama bukanlah intervensi yang sama bagi setiap mahasiswa.

Sintesis dirangkum pada Tabel 5 dengan menerapkan aturan yang telah dinyatakan pada subbab Kerangka Sintesis dan Evaluasi Hasil, yakni bahwa suatu tema dinilai terdukung dengan baik hanya jika sekurang-kurangnya dua studi mengarah pada hal yang sama. Tabel ini memperlihatkan bahwa tema-tema yang paling positif bersandar pada bukti berbasis persepsi dari dua survei, sedangkan tema yang menyangkut kualitas luaran bersandar pada satu eksperimen sehingga perlu dibaca sebagai bukti satu studi.

Tabel 5. Hasil sintesis tematik atas ketiga studi

Tema	Studi pendukung	Jenis bukti	Kekuatan dukungan
Pendamping pemahaman	Hendrik dkk. (2025); Makna dkk. (2026)	Berbasis persepsi	Didukung dua studi; sebatas persepsi
Pemercepat kerja	Hendrik dkk. (2025); Makna dkk. (2026); Utami dkk. (2026) sebagai dukungan tidak langsung	Berbasis persepsi dan objektif	Didukung oleh dua survei; eksperimen selaras pada tugas dasar
Belum terbuktinya peningkatan kualitas	Utami dkk. (2026)	Objektif (SonarQube)	Bertumpu pada satu studi
Risiko ketergantungan dan integritas akademik	Hendrik dkk. (2025); Makna dkk. (2026); Utami dkk. (2026)	Berbasis persepsi dan rekomendasi	Disepakati ketiga studi
Konteks menentukan hasil	Utami dkk. (2026); Makna dkk. (2026)	Objektif dan berbasis persepsi	Didukung dua studi; sifatnya interpretatif

Konsistensi dengan Literatur yang Lebih Luas

Secara umum, sintesis ini konsisten dengan literatur yang lebih luas. Gambaran penggunaan yang luas dan menguntungkan di kalangan mahasiswa sejalan dengan tinjauan yang melukiskan ChatGPT sebagai tutor virtual dan asisten pemrograman (Baig dkk., 2026;

Lo, 2023), serta dengan adopsi cepat yang dilaporkan pada mahasiswa komputasi di tempat lain (Hou dkk., 2024). Ketiadaan perbedaan kualitas yang terukur senada dengan temuan meta-analisis bahwa pembelajaran berbantuan kecerdasan buatan memangkas waktu pengerjaan tugas dan menaikkan skor kinerja, tetapi tidak menghasilkan keunggulan yang signifikan pada keberhasilan belajar maupun kemudahan pemahaman (Alanazi dkk., 2025). Di sisi lain, eksperimen yang ditinjau tidak mereproduksi peningkatan yang dilaporkan oleh studi terkendali tentang *computational thinking* dan penulisan kode (Kazemitabaar dkk., 2023; Yilmaz & Karaoglan Yilmaz, 2023). Perbedaan rancangan menyediakan penjelasan yang masuk akal. Studi-studi tersebut mengukur konstruk belajar yang spesifik, misalnya *computational thinking* dan kinerja penulisan kode, sedangkan eksperimen yang ditinjau membandingkan kualitas kode pada dua tugas singkat menggunakan tiga metrik analisis statis. Dengan demikian, perbandingan ini tidak berarti bahwa peningkatan yang ditemukan sebelumnya keliru; yang ditunjukkannya adalah bahwa metrik kualitas kode pada tugas dasar mungkin tidak peka terhadap jenis manfaat yang dilaporkan mahasiswa.

Implikasi

Sintesis tersebut melahirkan sejumlah implikasi (analisis penulis) yang dirangkum pada Tabel 6. Mahasiswa sebaiknya memanfaatkan ChatGPT untuk menjelaskan konsep, menelusuri kesalahan, dan membandingkan pendekatan, namun berupaya menyelesaikan masalah lebih dulu dan memverifikasi keluarannya, sebab pemakaian model bahasa yang baik menuntut literasi dan kesadaran akan keterbatasannya (Kasneci dkk., 2023). Dosen disarankan merancang tugas yang menuntut penjelasan penalaran, pengujian, dan *code review*, dengan dukungan pelatihan bagi pengajar serta pembaruan penilaian (Lo, 2023). Institusi memerlukan pedoman yang jelas yang memperlakukan keluaran chatbot sebagai informasi yang wajib diverifikasi sebelum dipakai (Sutabri, 2012b), disertai pelatihan dan langkah-langkah untuk menangkal ketidakjujuran akademik (Cotton dkk., 2024; Rahman & Watanobe, 2023). Peneliti sebaiknya memadukan ukuran persepsi dengan hasil belajar objektif pada sampel yang lebih besar dan tugas yang lebih kompleks (Sun dkk., 2024).

Tabel 6. Implikasi sintesis untuk masing-masing pemangku kepentingan

Pemangku kepentingan	Tindakan yang direkomendasikan	Dasar
Mahasiswa	Memakai ChatGPT untuk menjelaskan konsep, menelusuri kesalahan, dan membandingkan pendekatan; berupaya menyelesaikan masalah sendiri lebih dulu; memverifikasi keluaran melalui dokumentasi atau sumber lain	Penjelasan ulang dan *debugging* menjadi tujuan utama (Makna dkk., 2026); literasi dan kesadaran akan keterbatasan dibutuhkan (Kasneci dkk., 2023)
Dosen	Merancang tugas yang menuntut penjelasan penalaran, pengujian, dan *code review*; memperbarui metode penilaian serta mengikuti pelatihan	Kualitas tidak berbeda signifikan pada tugas dasar (Utami dkk., 2026); pelatihan serta penilaian baru dipandang penting (Lo, 2023)
Institusi	Menerbitkan pedoman yang jelas mengenai pemakaian kecerdasan buatan dan etika akademik, meliputi kebijakan, pelatihan, dan pencegahan ketidakjujuran	Pemakaian yang proporsional dan kritis direkomendasikan oleh studi yang ditinjau; strategi integritas (Cotton dkk., 2024)
Peneliti	Memakai sampel yang lebih besar, tugas yang lebih kompleks, dan metrik yang lebih banyak; memadukan persepsi dengan hasil belajar objektif	Sampel dan tugas pada eksperimen terbatas (Utami dkk., 2026); rancangan gabungan (Sun dkk., 2024)

Keterbatasan dan Penelitian Lanjutan

Kajian ini hanya menganalisis tiga studi, sehingga simpulannya bersifat indikatif dan tidak dapat digeneralisasi. Ketiga studi berfokus pada ChatGPT, sehingga chatbot lain tidak

tercakup. Perbedaan populasi, termasuk keterlibatan program studi non-informatika pada dua studi (Hendrik dkk., 2025; Makna dkk., 2026), serta perbedaan instrumen membuat angka-angka antarstudi hanya dapat dibandingkan secara kualitatif. Keterbatasan lainnya berkaitan dengan prosedur kajian itu sendiri: pencarian bertumpu pada sejumlah kecil kata kunci dan dibatasi pada artikel berbahasa Indonesia terbitan 2025–2026, sehingga studi relevan yang memakai istilah, bahasa, atau periode yang lain mungkin tidak terjaring.

Penelitian lanjutan sebaiknya melibatkan sampel yang lebih besar dan lebih beragam, tugas pemrograman yang lebih kompleks, metrik kualitas kode yang lebih banyak, serta beberapa chatbot kecerdasan buatan. Rancangan longitudinal dan metode campuran dianjurkan (Baig dkk., 2026), dengan memadukan ukuran persepsi dan luaran objektif, termasuk pemikiran kritis dan integritas akademik. Kajian mendatang juga dapat memperluas pencarian ke lebih banyak bahasa dan basis data, serta menerapkan instrumen penilaian kualitas yang formal.

SIMPULAN

Kajian komparatif atas ketiga studi ini memperlihatkan bahwa ChatGPT telah dipakai secara luas oleh mahasiswa informatika (96,7% pada satu studi (Hendrik dkk., 2025)) dan dipersepsikan mudah digunakan (81,31%), bermanfaat (78,63%), serta cukup berhasil (77,30%), dengan ketergantungan pada kategori netral (54,79%) (Makna dkk., 2026). Mahasiswa memakainya terutama untuk memperoleh penjelasan ulang materi, memahami konsep, dan *debugging*. Secara objektif, bantuan ChatGPT tidak menghasilkan perbedaan kualitas kode yang signifikan pada tugas pemrograman dasar ($p > 0,05$ pada semua metrik) (Utami dkk., 2026). Dengan demikian, ChatGPT berdaya guna sebagai pendamping pemahaman dan pemercepat pekerjaan, namun belum terbukti mampu mendongkrak kualitas luaran.

Sumbangan penelitian ini berupa perbandingan terstruktur yang menjajarkan bukti berbasis persepsi dan bukti objektif, sekaligus memperlihatkan bahwa ketiga studi saling melengkapi: kepercayaan diri dan adopsi (Hendrik dkk., 2025), kerangka multi-indikator dengan pengujian instrumen (Makna dkk., 2026), dan pengukuran objektif atas kualitas luaran (Utami dkk., 2026). Implikasinya, pemakaian chatbot hendaknya proporsional, kritis, dan berintegritas akademik, dan dosen maupun institusi dapat menopangnya melalui rancangan tugas serta pedoman yang jelas. Keterbatasan kajian ini, yang bertumpu pada tiga studi dan satu chatbot, beserta arah penelitian lanjutan diuraikan pada subbab Keterbatasan dan Penelitian Lanjutan.

DAFTAR PUSTAKA

Alanazi, M., Soh, B., Samra, H., & Li, A. (2025). The influence of artificial intelligence tools on learning outcomes in computer programming: A systematic review and meta-analysis. *Computers*, *14*(5), Article 185. https://doi.org/10.3390/computers14050185

Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education*, *3*(1), 1–26. https://doi.org/10.1007/s44217-024-00138-2

Baig, M. I., Yadegaridehkordi, E., & Bokani, A. (2026). ChatGPT in computer programming education: A review of current literature and applications. *Australasian Journal of Educational Technology*, *42*(2). https://doi.org/10.14742/ajet.10588

Bazelais, P., Lemay, D. J., & Doleck, T. (2024). User acceptance and adoption dynamics of ChatGPT in educational settings. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(2), Article em2393. https://doi.org/10.29333/ejmste/14151

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

da Silva, C. A. G., Ramos, F. N., de Moraes, R. V., & dos Santos, E. L. (2024). ChatGPT: Challenges and benefits in software programming for higher education. *Sustainability*, *16*(3), Article 1245. https://doi.org/10.3390/su16031245

Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. *Communications of the ACM*, *67*(2), 56–67. https://doi.org/10.1145/3624720

Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence*, *5*, Article 100190. https://doi.org/10.1016/j.caeai.2023.100190

Hendrik, F. N., Sogen, M. M. B., Baidenggan, W. M., & Do'o, F. (2025). Analisis pemanfaatan AI ChatGPT dalam membantu proses belajar mahasiswa Pendidikan Informatika di Universitas Citra Bangsa. *JUPE: Jurnal Pendidikan Mandala*, *10*(2), 556–558. https://doi.org/10.58258/jupe.v10i2.8860

Hou, I., Mettille, S., Man, O., Li, Z., Zastudil, C., & MacNeil, S. (2024). The effects of generative AI on computing students' help-seeking preferences. In *Proceedings of the 26th Australasian Computing Education Conference* (ACE 2024). ACM. https://doi.org/10.1145/3636243.3636248

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., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kazemitabaar, M., Chow, J., Ma, C. K. T., Ericson, B. J., Weintrop, D., & Grossman, T. (2023). Studying the effect of AI code generators on supporting novice learners in introductory programming. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23). ACM. https://doi.org/10.1145/3544548.3580919

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

Maheshwari, G. (2024). Factors influencing students' intention to adopt and use ChatGPT in higher education: A study in the Vietnamese context. *Education and Information Technologies*, *29*(10), 12167–12195. https://doi.org/10.1007/s10639-023-12333-z

Makna, L., Siswanto, V. A., & Kurniawan, I. (2026). Persepsi penggunaan ChatGPT sebagai alat bantu belajar mata kuliah pemrograman di Institut Widya Pratama. *Jurnal Teknologi Terpadu*, *12*(1), 66–76.

Manorat, P., Tuarob, S., & Pongpaichet, S. (2025). Artificial intelligence in computer programming education: A systematic literature review. *Computers and Education: Artificial Intelligence*, *8*, Article 100403. https://doi.org/10.1016/j.caeai.2025.100403

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers*, *12*(8), Article 153. https://doi.org/10.3390/computers12080153

Moradi Dakhel, A., Majdinasab, V., Nikanjam, A., Khomh, F., Desmarais, M. C., & Jiang, Z. M. (2023). GitHub Copilot AI pair programmer: Asset or liability? *Journal of Systems and Software*, *203*, Article 111734. https://doi.org/10.1016/j.jss.2023.111734

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, Article n71. https://doi.org/10.1136/bmj.n71

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

Shoufan, A. (2023). Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey. *IEEE Access*, *11*, 38805–38818. https://doi.org/10.1109/ACCESS.2023.3268224

Singh, H., Tayarani-Najaran, M.-H., & Yaqoob, M. (2023). Exploring computer science students' perception of ChatGPT in higher education: A descriptive and correlation study. *Education Sciences*, *13*(9), Article 924. https://doi.org/10.3390/educsci13090924

Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, *49*, 223–245. https://doi.org/10.1007/s10755-023-09686-1

Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, *21*(1), Article 14. https://doi.org/10.1186/s41239-024-00446-5

Sutabri, T. (2012a). *Analisis sistem informasi*. Andi Offset.

Sutabri, T. (2012b). *Konsep sistem informasi*. Andi Offset.

Sutabri, T. (2014). *Pengantar teknologi informasi*. Andi Offset.

Utami, S. S., Oktavia, S. A., & Habib, M. A. (2026). Analisis dampak ChatGPT sebagai code assistant terhadap kualitas kode mahasiswa Informatika UTS. *Management of Information System Journal*, *4*(3), 301–308. https://journal.fkpt.org/index.php/mis/article/view/3043

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, *4*, Article 100147. https://doi.org/10.1016/j.caeai.2023.100147

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*, Article 28. https://doi.org/10.1186/s40561-024-00316-7