

EVALUASI PROGRAM AI CHATBOT ANSIA DI PERGURUAN TINGGI: ANALISIS MODEL CIPP (CONTEXT, INPUT, PROCESS, PRODUCT)**Rafi Ardy Rizqullah¹, Meylia Elizabeth Ranu²**

Program Studi S1 Pendidikan Administrasi Perkantoran, Fakultas Ekonomika dan Bisnis,
Universitas Negeri Surabaya^{1,2}
e-mail: rafi.22149@mhs.unesa.ac.id

Diterima: 22/09/2026; Direvisi: 21/10/2026; Diterbitkan: 30/10/2026

ABSTRAK

Penelitian ini bertujuan mengevaluasi program chatbot berbasis *artificial intelligence*, yaitu ANSIA (Asisten SIDIA), di Universitas Negeri Surabaya menggunakan model evaluasi CIPP (*Context, Input, Process, Product*). Penelitian menggunakan pendekatan kualitatif dengan desain deskriptif evaluatif. Data dikumpulkan melalui wawancara mendalam dan dokumentasi yang melibatkan pengembang dari Direktorat Inovasi Pembelajaran Digital (DIPD) UNESA sebagai informan kunci serta enam dosen dan 19 mahasiswa Program Studi S1 Pendidikan Administrasi Perkantoran angkatan 2024 dan 2025 sebagai informan utama. Keabsahan data diperiksa melalui triangulasi sumber, sedangkan analisis data menggunakan model interaktif Miles dan Huberman. Hasil evaluasi menunjukkan empat kesenjangan utama. Pada komponen *context*, pengembangan ANSIA tidak didahului asesmen kebutuhan formal sehingga ditemukan kesenjangan kesadaran (*awareness gap*) pengguna. Pada komponen *input*, dukungan infrastruktur server dan integrasi WhatsApp belum diimbangi anggaran khusus, sosialisasi formal, serta panduan operasional yang memadai. Pada komponen *process*, peningkatan jumlah akses berdasarkan log server agregat tidak sepenuhnya mencerminkan pemanfaatan ANSIA oleh mahasiswa di tingkat program studi. Temuan wawancara menunjukkan bahwa kecenderungan menggunakan *artificial intelligence* komersial dan keterbatasan basis pengetahuan internal menjadi faktor yang berkaitan dengan rendahnya pemanfaatan tersebut. Pada komponen *product*, terdapat perbedaan antara klaim pengembang mengenai ketersediaan layanan selama 24 jam dan pengalaman pengguna dalam mengakses serta memanfaatkan sistem. Penelitian ini merekomendasikan perluasan basis pengetahuan ilmiah, penyediaan panduan visual, pengalokasian anggaran khusus, dan integrasi ANSIA dalam kegiatan pembelajaran.

Kata kunci: ANSIA, evaluasi program, *artificial intelligence*, model CIPP, perguruan tinggi.

ABSTRACT

This study aims to evaluate ANSIA (SIDIA Assistant), an artificial intelligence -based chatbot program at Universitas Negeri Surabaya, using the CIPP (*Context, Input, Process, Product*) evaluation model. A descriptive-evaluative qualitative approach was employed. Data were collected through in-depth interviews and document analysis involving developers from the Directorate of Digital Learning Innovation (DIPD) at UNESA as key informants, as well as six lecturers and 19 students from the Office Administration Education undergraduate study program, cohorts 2024 and 2025, as primary informants. Data credibility was examined through source triangulation, and the data were analyzed using the Miles and Huberman interactive model. The evaluation revealed four major gaps. In the *context* component, ANSIA was developed without a formal needs assessment, resulting in a gap in user awareness. In the *input* component, server infrastructure and WhatsApp integration were not adequately supported by a dedicated budget, formal socialization, or sufficient operational guidelines. In the *process*

component, an increase in aggregate server-log access did not fully reflect student use of ANSIA at the study-program level. Interview findings indicated that preferences for commercial *artificial intelligence* tools and limitations in the internal knowledge base were factors associated with low utilization. In the *product* component, a discrepancy emerged between developers' claims of 24-hour service availability and users' experiences in accessing and using the system. The study recommends expanding the scientific knowledge base, providing visual user guides, allocating a dedicated budget, and integrating ANSIA into teaching and learning activities.

Keywords: *ANSIA, artificial intelligence, CIPP model, higher education, program evaluation.*

PENDAHULUAN

Akselerasi teknologi *Artificial Intelligence*, khususnya *Large Language Model* (LLM), telah mengubah lanskap pendidikan tinggi. Perkembangan ini mendorong institusi pendidikan tinggi untuk merumuskan strategi integrasi *artificial intelligence* guna mendukung personalisasi pembelajaran dan meningkatkan efisiensi layanan administrasi. Di Indonesia, upaya pengembangan ekosistem *artificial intelligence* di perguruan tinggi turut memperoleh perhatian melalui inisiatif pemerintah, termasuk pengembangan pusat dan fasilitas pendukung pemanfaatan *artificial intelligence*. Dalam konteks tersebut, *artificial intelligence* akademik menjadi salah satu bentuk penerapan *artificial intelligence* yang berpotensi memperluas akses informasi sekaligus mendukung pelayanan akademik di perguruan tinggi.

Menindaklanjuti perkembangan tersebut, Universitas Negeri Surabaya (UNESA) melalui Direktorat Inovasi Pembelajaran Digital (DIPD) mengembangkan ANSIA (Asisten SIDIA), yaitu asisten akademik digital berbasis *Natural Language Processing* (NLP). ANSIA dirancang untuk beroperasi secara terintegrasi melalui arsitektur *Single Sign-On* (SSO) dan terhubung dengan ekosistem digital kampus, seperti SIAKADU, LMS, dan SIDIA Mobile. Pengembangan ANSIA ditujukan untuk menyediakan layanan informasi akademik yang dapat diakses setiap saat oleh sivitas akademika serta membantu mengurangi beban operator administrasi dalam menangani pertanyaan yang berulang.

Meskipun demikian, implementasi teknologi *artificial intelligence* di tingkat institusi tidak selalu diikuti oleh tingkat pemanfaatan yang sesuai dengan tujuan pengembangannya. Hasil studi pendahuluan terhadap 42 mahasiswa Program Studi S1 Pendidikan Administrasi Perkantoran (PAP) UNESA angkatan 2024 dan 2025 menunjukkan adanya kesenjangan antara penggunaan *artificial intelligence* secara umum dan pemanfaatan *artificial intelligence* yang disediakan institusi. Seluruh responden menyatakan menggunakan bantuan *artificial intelligence* dalam kegiatan akademik sehari-hari, dengan platform yang paling banyak digunakan adalah ChatGPT (73,8%) dan Claude *artificial intelligence* (16,7%), diikuti oleh Gemini *artificial intelligence*. Namun, meskipun 61,9% responden mengetahui keberadaan ANSIA, sebanyak 73,8% mengaku tidak pernah menggunakannya, sedangkan 19% lainnya menyatakan jarang menggunakannya. Alasan yang paling banyak diungkapkan meliputi kebiasaan menggunakan platform *artificial intelligence* eksternal (50%) dan keterbatasan pemahaman mengenai fitur khusus ANSIA (40,5%). Temuan awal ini memperlihatkan bahwa pengetahuan tentang keberadaan platform belum tentu diikuti oleh pemanfaatan secara nyata.

Kondisi tersebut menunjukkan adanya kesenjangan antara tujuan pengembangan teknologi *artificial intelligence* institusional dan pola penggunaannya oleh mahasiswa. Rendahnya pemanfaatan ANSIA perlu dikaji dengan mempertimbangkan berbagai aspek, termasuk kesesuaian layanan dengan kebutuhan pengguna, kesiapan sumber daya pendukung, pelaksanaan sosialisasi, serta pengalaman pengguna dalam mengakses layanan. Tanpa evaluasi yang menyeluruh, institusi akan mengalami kesulitan dalam mengidentifikasi faktor-faktor

yang menghambat pemanfaatan program dan menentukan perbaikan yang diperlukan. Oleh karena itu, evaluasi ANSIA perlu dilakukan tidak hanya untuk menilai hasil pelaksanaannya, tetapi juga untuk memahami kondisi yang melatarbelakangi pengembangan program dan proses implementasinya.

Penelitian ini menggunakan model evaluasi CIPP (*Context, Input, Process, Product*) yang dikembangkan oleh Daniel L. Stufflebeam. Model ini menyediakan kerangka evaluasi yang komprehensif melalui empat dimensi. Dimensi *context* digunakan untuk mengidentifikasi kebutuhan pengguna dan menilai kesesuaian tujuan pengembangan ANSIA dengan permasalahan layanan akademik. Dimensi *input* diarahkan untuk mengkaji kesiapan sumber daya manusia, infrastruktur teknis, dukungan anggaran, serta ketersediaan panduan operasional. Dimensi *process* digunakan untuk menelaah pelaksanaan sosialisasi, keterlibatan dosen, pemanfaatan platform, dan hambatan implementasi. Sementara itu, dimensi *product* berfokus pada pencapaian tujuan program, pengalaman dan kepuasan pengguna, serta manfaat yang dirasakan dari keberadaan ANSIA. Keempat dimensi tersebut memungkinkan evaluasi dilakukan secara terpadu, mulai dari dasar kebutuhan program hingga hasil pelaksanaannya.

Penelitian ini memfokuskan kajian pada implementasi ANSIA di Program Studi S1 Pendidikan Administrasi Perkantoran UNESA, khususnya pada mahasiswa angkatan 2024 dan 2025 serta dosen yang terlibat dalam kegiatan akademik. Pemilihan subjek tersebut didasarkan pada relevansi bidang administrasi perkantoran dengan pemanfaatan teknologi informasi serta keterkaitan pengalaman akademik kedua angkatan dengan periode implementasi ANSIA. Kebaruan penelitian terletak pada penggunaan model CIPP untuk mengevaluasi ANSIA melalui triangulasi perspektif pengembang, dosen, dan mahasiswa, sehingga evaluasi tidak hanya menggambarkan pelaksanaan program, tetapi juga mengidentifikasi faktor-faktor yang berkaitan dengan pemanfaatan *artificial intelligence* institusional. Dengan demikian, penelitian ini bertujuan untuk mengevaluasi implementasi ANSIA di Program Studi S1 Pendidikan Administrasi Perkantoran UNESA berdasarkan dimensi *context, input, process, dan product*, serta mengidentifikasi faktor pendukung dan penghambat pemanfaatannya berdasarkan perspektif pengembang, dosen, dan mahasiswa.

METODE PENELITIAN

Penelitian ini mengadopsi pendekatan kualitatif dengan jenis deskriptif evaluatif. Desain evaluatif dipilih untuk menghasilkan deskripsi faktual dan mendalam mengenai kondisi program ANSIA tanpa melakukan manipulasi variabel atau pengujian hipotesis statistik.

Lokasi dan Waktu Penelitian

Penelitian dilaksanakan di Program Studi S1 Pendidikan Administrasi Perkantoran (PAP), Fakultas Ekonomika dan Bisnis, Universitas Negeri Surabaya (UNESA), Kampus Ketintang, Kota Surabaya, Jawa Timur. Lokasi dipilih karena dosen, mahasiswa, dan pengembang ANSIA berada di lingkungan UNESA sehingga data dapat dikumpulkan secara langsung. Penelitian berlangsung pada Juni sampai Agustus 2026 yang mencakup penyusunan instrumen, pengumpulan data, analisis data, dan penyusunan laporan. Wawancara dengan dosen dilaksanakan pada 1-2 Juli 2026 di lingkungan Fakultas Ekonomika dan Bisnis UNESA, wawancara dengan pengembang pada 16 Juli 2026 di ruang Direktorat Inovasi Pembelajaran Digital (DIPD), Gedung Perpustakaan Lantai 2, Kampus UNESA Lidah Wetan, dan pengisian wawancara tertulis oleh mahasiswa pada 14-17 Agustus 2026.

Subjek dan Teknik Sampling

Subjek penelitian ditentukan menggunakan teknik *purposive sampling*, yaitu pemilihan informan berdasarkan kriteria relevansi dan kapasitas pengetahuan terhadap objek evaluasi. Informan dikelompokkan menjadi dua kategori:

Informan Kunci (Key Informant): Pengembang/admin program ANSIA dari Direktorat Inovasi Pembelajaran Digital (DIPD) UNESA (1 orang).

Informan Utama (Main Informants): Dosen tetap Prodi S1 Pendidikan Administrasi Perkantoran UNESA yang aktif mengajar angkatan 2024/2025 (6 orang) dan mahasiswa aktif Prodi S1 PAP angkatan 2024 dan 2025 (19 orang).

Kriteria pemilihan ditetapkan sebelum pengumpulan data. Pengembang dipilih sebagai informan kunci karena merancang, membangun, dan mengelola ANSIA sejak awal sehingga paling mengetahui komponen context dan input. Dosen dipilih dengan kriteria: (1) dosen tetap Prodi S1 PAP UNESA, (2) aktif mengajar mahasiswa angkatan 2024/2025, dan (3) bersedia menjadi informan. Mahasiswa dipilih dengan kriteria: (1) mahasiswa aktif Prodi S1 PAP UNESA, (2) angkatan 2024/2025, dan (3) bersedia menjadi informan. Pembatasan pada angkatan 2024 dan 2025 didasarkan pada kronologi program: ANSIA mulai digunakan oleh sivitas akademika pada 2024, sehingga kedua angkatan tersebut menempuh perkuliahan sejak ANSIA tersedia dan memiliki peluang yang setara untuk mengenal serta memanfaatkannya. Mahasiswa yang belum pernah menggunakan ANSIA tetap dilibatkan karena kondisi itu merupakan bagian dari objek evaluasi, yaitu tingkat kesadaran, jangkauan sosialisasi, dan hambatan adopsi. Jumlah informan tidak ditentukan secara statistik, tetapi berdasarkan relevansi terhadap tujuan evaluasi dan kecukupan variasi informasi. Setiap informan diberi kode (DIPD.INFORMAN.01, DSN.2026.01–06, MHS.2024.xx, dan MHS.2025.xx).

Teknik Pengumpulan Data dan Keabsahan Data

Pengumpulan data dilakukan melalui dua teknik utama: wawancara mendalam (*in-depth interview*) dan dokumentasi. Wawancara dengan pengembang dan dosen dilakukan secara tatap muka langsung, sedangkan wawancara dengan mahasiswa dilakukan melalui instrumen tertulis terstruktur guna efisiensi pengumpulan data tanpa mengurangi kedalaman respon. Untuk menjamin kredibilitas dan keabsahan data, penelitian ini menggunakan teknik **triangulasi sumber**. Data yang diperoleh dari pengembang disandingkan dan diverifikasi silang dengan data dari kelompok dosen dan mahasiswa pada setiap dimensi CIPP.

Analisis Data

Analisis data menerapkan model interaktif Miles dan Huberman yang terdiri dari tiga tahapan yang berlangsung secara siklis: **Reduksi Data:** Memilah, memfokuskan, merangkum, dan mengkategorikan data wawancara mentah ke dalam empat matriks komponen CIPP. **Penyajian Data (Data Display):** Menyusun temuan kualitatif dalam bentuk narasi analitis dan tabel triangulasi antar-sumber. **Verifikasi dan Penarikan Kesimpulan:** Menginterpretasikan pola data, menghubungkannya dengan teori adopsi teknologi, dan merumuskan rekomendasi perbaikan

HASIL DAN PEMBAHASAN

Hasil

Hasil penelitian disajikan dan dibahas secara terpadu berdasarkan empat dimensi model CIPP. Sintesis triangulasi dari ketiga kelompok informan dirangkum pada Tabel 1.

Tabel 1. Matriks Triangulasi Sumber Evaluasi CIPP Program ANSIA

Komponen Evaluasi	Informan Kunci (DIPD UNESA)	Informan Utama (Dosen PAP)	Informan Utama (Mahasiswa PAP)	Temuan Analitis (Kesenjangan Utama)
Context	Dikembangkan berdasarkan	Menilai relevan secara teoretis;	5 dari 19 mahasiswa baru	Awareness Gap: ketiadaan need

	tren LLM dan pertanyaan berulang; tanpa need assessment formal; sosialisasi melalui banner SIDIA.	1 dari 6 dosen sama sekali tidak tahu ANSIA.	tahu ANSIA saat wawancara; relevan untuk informasi jadwal/KRS.	assessment menyebabkan pengembangan berpatokan pada asumsi sistem, bukan kebutuhan riil pengguna.
Input	Tim teknis 3 orang; migrasi ke WhatsApp API melalui SSO; tanpa anggaran khusus dan tanpa user manual formal.	Menyoroti ketiadaan user manual resmi; 4 dari 6 dosen tidak pernah menerima sosialisasi formal.	Keluhan ketiadaan tutorial resmi; sistem SSO/SIDIA kadang mengalami gangguan kelambatan (lag).	Structural Gap: kapasitas teknis tim tidak diimbangi alokasi anggaran, buku panduan, dan sosialisasi terstruktur.
Process	Mengklaim efektif berdasarkan kenaikan data log server agregat kampus.	Hampir tidak pernah menggunakan dalam pengajaran; ragu dengan akurasi dan cakupan data.	14 dari 19 mahasiswa tidak pernah mencoba ANSIA; ketergantungan tinggi pada ChatGPT/Claude.	Usage Gap dan Status Quo Bias: data log agregat menyembunyikan adopsi riil prodi yang sangat rendah akibat kebiasaan memakai <i>artificial intelligence</i> eksternal.
Product	Menilai tujuan 24 jam tercapai penuh dalam melayani informasi akademik.	Meminta perluasan basis data ke referensi ilmiah bereputasi (Scopus/Sinta).	Mengharapkan tutorial visual, respons lebih cepat, dan kemampuan menjawab tugas ilmiah.	System-User Gap: keberhasilan fungsional sistem pengembang tidak berkorelasi dengan dampak riil pada pengguna akhir.

Evaluasi Context: Need Assessment dan Awareness Gap

Evaluasi konteks menilai kesesuaian antara latar belakang pengembangan ANSIA dengan kebutuhan nyata civitas akademika. DIPD UNESA mengembangkan ANSIA untuk menyediakan layanan informasi akademik 24 jam sekaligus mengurangi beban kerja operasional tenaga kependidikan. Namun, pengembangan platform ini tidak didahului oleh penelitian atau *need assessment* formal, melainkan hanya berpatokan pada pengamatan praktis tim teknis terhadap pertanyaan yang sering muncul.

Ketiadaan *need assessment* sistematis bertentangan dengan prinsip evaluasi konteks Stufflebeam yang menegaskan bahwa diagnosis kebutuhan lapangan merupakan fondasi utama penetapan tujuan program. Akibatnya, terjadi kesenjangan kesadaran (*awareness gap*) yang tajam. Strategi pengembang yang hanya mengandalkan *banner* pasif pada aplikasi SIDIA

terbukti tidak efektif. Di tingkat fakultas, 1 dari 6 dosen tetap sama sekali tidak mengetahui keberadaan ANSIA, sementara di tingkat mahasiswa, 5 dari 19 informan baru mengetahui keberadaan ANSIA saat sesi wawancara penelitian berlangsung.

Pengembang menyatakan, “Pengembangan ANSIA tidak diawali dengan penelitian formal mengenai need assessment, melainkan berangkat dari pengamatan terhadap kebutuhan nyata di lapangan” (DIPD.INFORMAN.01). Dari sisi pengguna, seorang dosen menyampaikan, “Saya tidak tau bahwa UNESA memiliki platform *artificial intelligence* bernama ANSIA” (DSN.2026.03), dan seorang mahasiswa menyatakan, “Saya baru mengetahui ANSIA dari wawancara ini” (MHS.2025.01).

Fenomena ini selaras dengan tahap *knowledge/awareness* dalam teori Difusi Inovasi Rogers, yang menyatakan bahwa adopsi teknologi tidak akan pernah terjadi tanpa adanya kesadaran awal yang merata. Rendahnya kesadaran pengguna mengonfirmasi terjadinya kesenjangan digital tingkat kedua (*second-level digital divide*) sebagaimana dirumuskan oleh Van Dijk, bukan lagi terkait akses fisik terhadap teknologi, melainkan kesenjangan dalam kesadaran dan pemanfaatan fungsionalnya.

Evaluasi Input: Kapasitas Teknologis vs Ketiadaan Panduan dan Anggaran

Evaluasi masukan memeriksa kesiapan sumber daya pendukung program. Dari aspek teknis, ANSIA didukung oleh tim pengembang yang kompeten (3 personel) dan infrastruktur server kampus yang terintegrasi dengan arsitektur *Single Sign-On* (SSO). Keputusan tim pengembang mentransformasi antarmuka ANSIA dari aplikasi berbasis *web* menjadi layanan berbasis WhatsApp API merupakan langkah adaptif yang meningkatkan kompatibilitas (*compatibility*) dan menurunkan kompleksitas (*complexity*) aksesibilitas. Namun, keberhasilan teknis ini mengalami kesenjangan struktural (*structural gap*) akibat ketiadaan alokasi anggaran khusus untuk pengembangan dan sosialisasi. Ketiadaan anggaran berimplikasi langsung pada dua kendala utama:

Ketiadaan Buku Panduan Resmi: Pengembang berasumsi bahwa antarmuka WhatsApp yang sederhana tidak memerlukan buku panduan (user manual). Asumsi ini kurang tepat karena pengguna memerlukan panduan mengenai batasan domain, sintaks pertanyaan, dan fitur-fitur akademik yang mampu dilayani ANSIA. **Minimnya Sosialisasi Terstruktur:** Tanpa anggaran khusus, kegiatan sosialisasi interaktif atau pelatihan bagi dosen dan mahasiswa baru tidak dapat dilaksanakan.

Pengembang mengakui keterbatasan tersebut, “Tidak ada anggaran khusus yang dialokasikan untuk pengembangan ANSIA” (DIPD.INFORMAN.01). Seorang dosen juga menjelaskan bahwa sosialisasi yang diterima belum bersifat praktik, “Sosialisasi yang saya dapatkan hanya berupa pemaparan materi dari pihak akademik, tidak ada praktik secara langsung” (DSN.2026.01).

Kondisi ini mencerminkan lemahnya dukungan kelembagaan (*organizational support*). Sebagaimana dijelaskan oleh Almatrodi et al. dalam analisis resistensi organisasi, ketiadaan dukungan struktural dan finansial memperkuat kecenderungan pengguna untuk bertahan pada metode atau sistem lama.

Evaluasi Process: Paradoks Pemanfaatan, Keterlibatan Dosen, dan Status Quo Bias

Evaluasi proses mengukur pelaksanaan implementasi ANSIA di lapangan. Pada dimensi ini ditemukan kesenjangan penggunaan (*usage gap*) atau paradoks pemanfaatan. Pengembang mengklaim implementasi berjalan sangat efektif berdasarkan data agregat *log server* yang menunjukkan peningkatan grafik percakapan di tingkat universitas. Namun, analisis kualitatif di tingkat program studi menunjukkan bahwa 14 dari 19 mahasiswa belum pernah mengakses atau mencoba ANSIA sama sekali. Rendahnya pemanfaatan riil ini dipicu oleh dua faktor utama: **Status Quo Bias dan Switching Cost:** Dosen dan mahasiswa telah

memiliki keterikatan kebiasaan yang kuat (habit) terhadap platform *artificial intelligence* eksternal seperti ChatGPT, Claude, dan Gemini. Peralihan ke ANSIA menimbulkan switching cost psikologis, di mana pengguna enggan berpindah ke sistem baru yang keunggulannya belum terbukti. **Keterbatasan Basis Pengetahuan (Knowledge Base):** Ruang lingkup jawaban ANSIA terbatas pada informasi administrasi internal UNESA (seperti jadwal kuliah, KRS, dan lokasi gedung). ANSIA belum dibekali kemampuan untuk menjawab kueri ilmiah eksternal atau membantu pengerjaan tugas akademik yang membutuhkan referensi luas.

Selain itu, keterlibatan dosen sebagai agen perubahan (*opinion leader*) sangat minim. Hanya 1 dari 6 dosen yang pernah mengenalkan ANSIA secara sepintas pada awal semester. Mayoritas dosen enggan merekomendasikan ANSIA karena tidak mengetahui cakupan fiturnya atau ragu terhadap akurasi jawaban sistem.

Alasan rendahnya pemanfaatan tampak dalam wawancara. Seorang dosen menyebut kendala merekomendasikan ANSIA terjadi “karena beberapa dosen/mahasiswa sudah terbiasa menggunakan *artificial intelligence* lainnya” (DSN.2026.01), sedangkan seorang mahasiswa menyatakan, “Belum pernah mencoba menggunakan ANSIA, karena sebelumnya saya memang belum tahu tentang ANSIA” (MHS.2025.01).

Evaluasi Product: Kesenjangan Kinerja Sistem dan Pengalaman Pengguna

Evaluasi produk menilai ketercapaian tujuan program dan dampak riil pada pengguna akhir. Pihak pengembang menilai tujuan ANSIA telah tercapai 100% karena sistem *chatbot* otomatis 24 jam beroperasi dengan lancar untuk menjawab pertanyaan dasar tanpa tergantung jam kerja operator.

Sebaliknya, dari perspektif pengguna akhir, terjadi kesenjangan kualitas sistem-pengguna (*system-user gap*). Dosen dan mahasiswa belum dapat merasakan dampak fungsional yang signifikan dari keberadaan ANSIA. Mahasiswa menganggap ANSIA belum mampu berfungsi sebagai asisten akademik sejati selama basis pengetahuannya belum dihubungkan dengan repositori jurnal bereputasi (seperti Sinta atau Scopus).

Perbedaan perspektif ini tampak pada kutipan wawancara. Pengembang menyatakan, “Menurut saya tujuan awal sudah tercapai” (DIPD.INFORMAN.01), sedangkan seorang dosen mengamati bahwa “mahasiswa tidak menggunakan *artificial intelligence* ANSIA untuk kegiatan akademik mereka, sepertinya mereka menggunakan *artificial intelligence* lainnya” (DSN.2026.01). Seorang mahasiswa menambahkan, “Karena saya belum pernah menggunakan ANSIA, saya belum tahu manfaatnya secara langsung” (MHS.2025.01).

Kesenjangan ini menegaskan filosofi evaluasi CIPP Stufflebeam bahwa evaluasi produk bertujuan untuk memperbaiki program (*to improve*), bukan sekadar membuktikan keberhasilan teknis (*to prove*). Metrik keberhasilan pada tingkat server tidak dapat menggantikan indikator adopsi riil di tingkat pengguna akhir.

KESIMPULAN

Evaluasi program *artificial intelligence* ANSIA di Program Studi S1 Pendidikan Administrasi Perkantoran Universitas Negeri Surabaya menggunakan model CIPP menunjukkan bahwa implementasi program masih menghadapi kendala pada aspek perencanaan, dukungan sumber daya, pelaksanaan, dan manfaat bagi pengguna. Pada dimensi *context*, pengembangan ANSIA belum didukung *need assessment* formal sehingga kesesuaiannya dengan kebutuhan pengguna belum teridentifikasi secara menyeluruh. Pada dimensi *input*, kesiapan infrastruktur dan akses layanan belum diimbangi dengan dukungan anggaran khusus, sosialisasi formal, dan panduan operasional yang memadai. Pada dimensi *process*, pemanfaatan ANSIA belum optimal karena pengguna cenderung memilih *artificial intelligence* komersial, basis pengetahuan internal masih terbatas, dan keterlibatan dosen dalam

sosialisasi belum maksimal. Pada dimensi *product*, keberfungsian layanan selama 24 jam belum sepenuhnya mencerminkan manfaat yang dirasakan pengguna. Dengan demikian, peningkatan pemanfaatan ANSIA memerlukan penyelarasan antara kebutuhan pengguna, kesiapan dukungan institusi, strategi implementasi, dan kualitas layanan.

Pengembang melalui Direktorat Inovasi Pembelajaran Digital (DIPD) UNESA disarankan memperluas basis pengetahuan ANSIA, menyediakan panduan penggunaan yang mudah diakses, serta mengevaluasi stabilitas integrasi sistem. Pimpinan universitas perlu memperkuat dukungan anggaran, sosialisasi, dan pelatihan bagi sivitas akademika. Dosen diharapkan memperkenalkan ANSIA sebagai salah satu sumber informasi akademik yang relevan dalam kegiatan perkuliahan, sedangkan mahasiswa perlu memanfaatkannya secara kritis dengan memverifikasi informasi yang diperoleh. Evaluasi lanjutan juga diperlukan untuk menilai kesesuaian layanan dengan kebutuhan pengguna dan mengukur pemanfaatan ANSIA pada tingkat program studi secara lebih akurat.

DAFTAR PUSTAKA

- Aldhafeeri, L., Aljumah, F., Thabyan, F., Alabbad, M., AlShahrani, S., Alanazi, F., & Al-Nafjan, A. (2025). Generative AI chatbots across domains: A systematic review. *Applied Sciences*, 15(20), Article 11220. <https://doi.org/10.3390/app152011220>
- Alfred, P. A. P., Mabotha, P., & Ngcamu, B. S. (2026). Digital transformation in the higher education sector: A systematic literature review. *Administrative Sciences*, 16(1), Article 1. <https://doi.org/10.3390/admsci16010001>
- Almatrodi, I., Li, F., & Alojail, M. (2023). Organizational resistance to automation success: How status quo bias influences organizational resistance to an automated workflow system in a public organization. *Systems*, 11(4), Article 191.
- Andrews, T. C., Conaway, E. P., Zhao, J., & Dolan, E. L. (2016). Colleagues as change agents: How department networks and opinion leaders influence teaching at a single institution. *CBE—Life Sciences Education*, 15(2). [DOI atau URL artikel perlu dilengkapi].
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21, Article 4. <https://doi.org/10.1186/s41239-023-00436-z>
- 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, Article 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- El-Sobkey, S. B., Kelini, K. I., ElKholy, M., Abdeldayem, T., Abdallah, M., Al-Amir Mohamed, D., Fawzy, A., Ahmed, Y. F., El Khatib, A., Khalid, H., Shaik, B. B., Anjos, A., Alharbi, M. D., Fathy, K., & Takey, K. (2025). Acceptance of AI-powered chatbots among physiotherapy students: International cross-sectional study. *JMIR Medical Education*, 11, Article e76574.
- Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on*

- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anchors: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1), 193–206. <https://doi.org/10.1257/jep.5.1.193>
- 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>
- Kotter, J. P. (1996). *Leading change*. Harvard Business School Press.
- Liang, T. (2012). Matrix analysis of the digital divide in eHealth services using awareness, want, and adoption gap. *Journal of Medical Internet Research*, 14(1), Article e11. <https://doi.org/10.2196/jmir.1670>
- Menzli, L. J., Smirani, L. K., Boulahia, J. A., & Hadjouni, M. (2022). Investigation of open educational resources adoption in higher education using Rogers' diffusion of innovation theory. *Heliyon*, 8(7), Article e09885. <https://doi.org/10.1016/j.heliyon.2022.e09885>
- Nugawela, S., & Sedera, D. (2022). Status quo bias in users' information systems (IS) adoption and continuance intentions: A literature review and framework. *Australasian Conference on Information Systems (ACIS) 2022 Proceedings*, Article 65. <https://aisel.aisnet.org/acis2022/65>
- Rana, M., Safaet, M., Sakib, N., & Ahamed, R. (2024). Assessing AI adoption in developing country academia: A trust and privacy-augmented UTAUT framework. *Heliyon*, 10(18), Article e37569. <https://doi.org/10.1016/j.heliyon.2024.e37569>
- Ranu, M. E. (2026). *Evaluasi pendidikan: Teori, praktik, dan kontribusi terhadap pencapaian pendidikan berkualitas*. Pustaka Aksara.
- Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). The Free Press.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7–59. <https://doi.org/10.1007/BF00055564>
- Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, 6(3), 341–362. <https://doi.org/10.1177/1461444804042519>
- Strzelecki, A. (2024). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 32(9), 5142–5155. <https://doi.org/10.1080/10494820.2023.2209881>
- Stufflebeam, D. L. (1971). The relevance of the CIPP evaluation model for educational accountability. *Journal of Research and Development in Education*, 5(1), 19–25.
- Stufflebeam, D. L. (2002). *CIPP evaluation model checklist*. The Evaluation Center, Western Michigan University. [URL publikasi perlu dilengkapi].
- Stufflebeam, D. L., & Zhang, G. (2017). *The CIPP evaluation model: How to evaluate for improvement and accountability*. The Guilford Press.
- Tian, W., Ge, J., Zhao, Y., & Zheng, X. (2024). AI chatbots in Chinese higher education: Adoption, perception, and influence among graduate students—An integrated analysis utilizing UTAUT and ECM models. *Frontiers in Psychology*, 15, Article 1268549. <https://doi.org/10.3389/fpsyg.2024.1268549>



- Van Dijk, J. A. G. M. (2017). Digital divide: Impact of access. In *The international encyclopedia of media effects* (pp. 1–11). Wiley. [DOI atau URL bab perlu dilengkapi].
- Van Dijk, J. A. G. M. (2020). *The digital divide*. Polity Press.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178. <https://doi.org/10.2307/41410412>
- Warju, W. (2016). Educational program evaluation using CIPP model. *Innovation of Vocational Technology Education*, 12(1), 36–42.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C. K., & Misulis, K. (2011). Using the context, input, process, and product evaluation model (CIPP) as a comprehensive framework to guide the planning, implementation, and assessment of service-learning programs. *Journal of Higher Education Outreach and Engagement*, 15(4), 57–84.