PÉNDALAS
Jurnal Penelitian Tindakan Kelas dan Pengabdian Masyarakat
Publisher: Islamic and Multicultural Education Foundation
e-ISSN: 2807-7334

Vol. 6 No. 2, Mei – Agustus 2026	Copyright © 2026, is licensed under a CC-BY-NO-SA
Halaman: 666-686	*Publisher: Islamic and Multicultural Education Foundation*
https://journal.yaspim.org/index.php/pendalas/index	e-ISSN: 2807-7334

KEBIJAKAN KOMUNIKASI DALAM MENANGGULANGI DISINFORMASI BERBASIS KECERDASAN BUATAN DI MEDIA SOSIAL: STUDI LITERASI DIGITAL MAHASISWA

Ahmad Syukri Siregar
Universitas Islam Negeri Sumatera Utara Medan, Indonesia
ahmad4004253015@uinsu.ac.id

Hasan Sazali
Universitas Islam Negeri Sumatera Utara Medan, Indonesia
hasansazali@uinsu.ac.id

Received: 30 Juni 2026 Accepted: 31 Juli 2026 Published: 10 Agustus 2026

Abstrak

Disinformasi berbasis kecerdasan buatan meningkatkan skala, kecepatan, dan daya persuasi informasi menyesatkan di media sosial melalui produksi otomatis, manipulasi visual dan audiovisual, identitas sintetis, personalisasi pesan, serta pengulangan algoritmik, sedangkan kedekatan mahasiswa dengan teknologi digital belum selalu diikuti kemampuan verifikasi informasi yang memadai. Penelitian ini bertujuan menganalisis karakter disinformasi berbasis kecerdasan buatan, mengidentifikasi kerentanan dan kapasitas literasi digital mahasiswa, serta merumuskan model kebijakan komunikasi publik yang adaptif. Penelitian menggunakan kajian literatur integratif terhadap 34 publikasi ilmiah, laporan kebijakan, regulasi, dan dokumen resmi yang diterbitkan pada 2020–2026, kemudian dianalisis melalui ekstraksi informasi, pengodean tematik, perbandingan antarsumber, dan triangulasi jenis sumber. Hasil penelitian menunjukkan bahwa kebijakan yang berorientasi pada penghapusan konten dan klarifikasi setelah informasi menjadi viral tidak memadai sehingga kebijakan yang efektif harus bersifat preventif, adaptif, partisipatif, transparan, dan berbasis bukti dengan mengintegrasikan pemetaan risiko, respons cepat, pendidikan literasi media dan kecerdasan buatan, verifikasi informasi, partisipasi mahasiswa sebagai pemeriksa fakta sebaya, transparansi moderasi, serta kolaborasi antara perguruan tinggi, pemerintah, media, platform digital, dan masyarakat sipil.

Kata Kunci: Disinformasi, Kecerdasan Buatan, Komunikasi Publik, Literasi Digital, Mahasiswa**.**

Abstract

Artificial intelligence-based disinformation increases the scale, speed, and persuasive power of misleading information on social media through automated production, visual and audiovisual manipulation, synthetic identities, message personalization, and algorithmic amplification, while university students' familiarity with digital technologies is not always accompanied by adequate information-verification skills. This study aims to analyze the characteristics of artificial intelligence-based disinformation, identify students' vulnerabilities and digital literacy capacities, and formulate an adaptive public communication policy model. It employed an integrative literature review of 34 scholarly publications, policy reports, regulations, and official documents published between 2020 and 2026, which were analyzed through information extraction, thematic coding, cross-source comparison, and source triangulation. The findings indicate that policies focusing primarily on content removal and clarification after information becomes viral are insufficient so that effective policies must be preventive, adaptive, participatory, transparent, and evidence-based by integrating risk mapping, rapid response, media and artificial intelligence literacy education, information verification, student participation as peer fact-checkers, moderation transparency, and collaboration among universities, governments, media organizations, digital platforms, and civil society.

Keywords: Disinformation, Artificial Intelligence, Public Communication, Digital Literacy, University Students.

PENDAHULUAN

Ekosistem informasi digital memasuki fase yang semakin krusial ketika kecerdasan buatan generatif mampu memproduksi teks, gambar, audio, dan video sintetis dalam jumlah besar, cepat, murah, serta tampak meyakinkan. Di media sosial, kemampuan tersebut bertemu dengan mekanisme rekomendasi yang mengutamakan keterlibatan dan respons emosional, sehingga informasi menyesatkan dapat dipersonalisasi, diperbanyak, dan disebarkan melalui akun maupun identitas sintetis (Metzler & Garcia, 2024). Fenomena ini tidak hanya membuat konten palsu terlihat benar, tetapi juga memungkinkan bukti autentik disangkal sebagai hasil rekayasa. World Economic Forum (2025) menempatkan misinformasi dan disinformasi sebagai risiko global utama jangka pendek, sedangkan OECD menegaskan bahwa integritas informasi membutuhkan transparansi platform, komunikasi publik yang tepercaya, media yang plural, dan literasi warga (OECD, 2024). Mahasiswa menjadi kelompok penting karena intensif menggunakan platform digital dan kerap berperan sebagai penghubung informasi dalam keluarga maupun komunitas. Secara akademik, fenomena ini menuntut perluasan literasi digital menuju literasi kecerdasan buatan; secara praktis, diperlukan kebijakan komunikasi yang memperkuat verifikasi, transparansi, kebebasan berekspresi, dan kepercayaan publik.

Literatur terdahulu telah merespons masalah tersebut melalui sedikitnya empat jalur. Pertama, penelitian literasi media menunjukkan bahwa pelatihan evaluasi sumber dapat meningkatkan kemampuan membedakan berita arus utama dan informasi palsu (Guess et al., 2020). Kedua, penelitian tentang pengingat akurasi dan inokulasi psikologis

menunjukkan bahwa perhatian terhadap kebenaran serta pengenalan dini terhadap teknik manipulasi dapat mengurangi kerentanan pengguna (Pennycook et al., 2020). Ketiga, kajian literasi kecerdasan buatan merumuskan kompetensi memahami, menggunakan, mengevaluasi, dan mempertimbangkan dimensi etis teknologi (Long & Magerko, 2020). Keempat, studi kebijakan menekankan transparansi platform, mitigasi risiko sistemik, pelabelan konten sintetis, dan akuntabilitas kelembagaan (UNESCO, 2024). Penelitian lain menegaskan bahwa kedekatan mahasiswa dengan platform tidak otomatis menjamin kemampuan evaluasi sumber dan ketahanan terhadap manipulasi (Kozyreva et al., 2020). Namun, kajian tersebut masih berjalan terpisah dan belum mengintegrasikan kerentanan mahasiswa, komunikasi perguruan tinggi, teknologi verifikasi, regulasi, serta partisipasi mahasiswa dalam satu model kebijakan yang operasional.

Penelitian ini bertujuan merespons kesenjangan tersebut melalui kajian literatur integratif yang mempertemukan temuan ilmu komunikasi, pendidikan, psikologi, ilmu informasi, kebijakan publik, dan tata kelola teknologi. Secara khusus, penelitian menganalisis tiga masalah. Pertama, bagaimana kecerdasan buatan mengubah produksi dan penyebaran disinformasi melalui otomatisasi, manipulasi multimodal, personalisasi pesan, identitas sintetis, dan amplifikasi algoritmik. Kedua, bagaimana kerentanan dan kapasitas mahasiswa terbentuk melalui bias konfirmasi, tekanan kecepatan, heuristik sosial, kemampuan membaca lateral, pemahaman keterbatasan model, serta keterampilan komunikasi korektif. Ketiga, bagaimana kebijakan komunikasi publik dapat menghubungkan regulasi, protokol respons cepat, transparansi platform, teknologi verifikasi, pendidikan literasi media dan kecerdasan buatan, serta jejaring kolaborasi antarlembaga. Untuk menjawabnya, penelitian menyintesiskan 34 publikasi periode 2020–2026 melalui ekstraksi informasi, pengodean tematik, perbandingan antarsumber, dan triangulasi jenis dokumen. Pendekatan ini menempatkan mahasiswa bukan sekadar penerima edukasi, melainkan mitra yang dapat memeriksa fakta, memproduksi klarifikasi, dan memperkuat ketahanan informasi kampus. Sintesis tersebut diarahkan untuk menghasilkan kerangka kebijakan yang dapat diterapkan dan dievaluasi oleh perguruan tinggi.

Argumen utama penelitian ini adalah bahwa kebijakan komunikasi yang hanya berorientasi pada penghapusan konten dan klarifikasi setelah informasi menjadi viral tidak cukup untuk menanggulangi disinformasi berbasis kecerdasan buatan. Secara konseptual, semakin tinggi kemampuan teknologi menghasilkan konten sintetis dan semakin kuat amplifikasi algoritmik, semakin besar kerentanan mahasiswa apabila kemampuan verifikasi, pemahaman kecerdasan buatan, dan dukungan komunikasi institusi tidak berkembang secara seimbang. Sebaliknya, integrasi antara pemetaan risiko, respons cepat berbasis bukti, *prebunking*, transparansi moderasi, literasi media dan kecerdasan buatan,

partisipasi pemeriksa fakta sebaya, serta kolaborasi lintas lembaga diperkirakan meningkatkan ketahanan informasi mahasiswa. Hubungan sebab-akibat tersebut bekerja karena pendidikan membangun kompetensi, komunikasi publik mengurangi kekosongan informasi, teknologi verifikasi mempercepat pemeriksaan, transparansi memperkuat kepercayaan, dan partisipasi mahasiswa meningkatkan relevansi pesan bagi komunitas sebaya. Dengan demikian, jawaban sementara penelitian menyatakan bahwa kebijakan preventif, adaptif, partisipatif, transparan, dan berbasis bukti lebih efektif daripada kebijakan reaktif dan terfragmentasi. Karena menggunakan kajian literatur, argumen tersebut diuji melalui konsistensi pola temuan antarsumber, bukan melalui pengujian statistik pada satu perguruan tinggi tertentu.

LANDASAN TEORI

Landasan teori mengenai kebijakan komunikasi, disinformasi berbasis kecerdasan buatan, dan literasi digital mahasiswa berkembang melalui empat kecenderungan utama. Pertama, studi tentang karakter disinformasi menelaah bagaimana otomatisasi, manipulasi multimodal, identitas sintetis, dan amplifikasi algoritmik meningkatkan skala serta ketidakpastian informasi. *Deepfake*, misalnya, tidak selalu langsung menipu, tetapi dapat menimbulkan keraguan terhadap kebenaran dan menurunkan kepercayaan pada berita digital (Vaccari & Chadwick, 2020a). Kedua, penelitian literasi media menguji hubungan kemampuan evaluasi sumber dengan ketahanan terhadap informasi palsu melalui membaca lateral, pemeriksaan fakta, pengingat akurasi, dan *prebunking* (Roozenbeek et al., 2022). Ketiga, kajian literasi kecerdasan buatan memusatkan perhatian pada pemahaman cara kerja sistem, evaluasi keluaran, bias, keterbatasan, dan tanggung jawab etis (Ng et al., 2021). Keempat, literatur kebijakan menekankan integritas informasi, transparansi platform, komunikasi publik tepercaya, serta tata kelola kolaboratif (OECD, 2024). Peta ini menunjukkan bahwa aspek teknologi, kompetensi mahasiswa, dan kelembagaan telah dibahas, tetapi masih sering dipisahkan.

Kecenderungan pertama berorientasi pada karakter teknis, komunikatif, dan psikologis disinformasi berbasis kecerdasan buatan. Penelitian dalam pola ini memandang teknologi generatif sebagai pengganda kapasitas yang menurunkan biaya produksi pesan, memperbanyak variasi narasi, memalsukan identitas, serta mengubah disinformasi dari dominasi teks menuju gambar, audio, dan video sintetis. Vaccari & Chadwick (2020), melalui eksperimen terhadap video politik sintetis, menunjukkan bahwa dampak *deepfake* tidak hanya berupa penipuan, tetapi juga peningkatan ketidakpastian dan penurunan kepercayaan terhadap berita di media sosial. Temuan tersebut memperluas kajian tentang penyebaran informasi palsu yang sebelumnya menunjukkan bahwa kebaruan dan muatan emosional membantu informasi salah menyebar lebih cepat daripada informasi benar

(Vosoughi et al., 2018). Kajian mutakhir juga menempatkan sistem rekomendasi sebagai mekanisme yang dapat memperbesar visibilitas konten yang memicu keterlibatan, meskipun platform bukan pencipta langsung niat manipulatif (Metzler & Garcia, 2024). Secara metodologis, pola ini didominasi eksperimen, analisis platform, studi komputasional, dan kajian konseptual. Kekuatan utamanya ialah menjelaskan bentuk ancaman, tetapi fokusnya sering berhenti pada karakter konten dan respons pengguna.

Kecenderungan kedua memusatkan perhatian pada literasi media, perilaku verifikasi, dan intervensi pendidikan untuk meningkatkan ketahanan mahasiswa. Orientasi tematisnya adalah hubungan antara kemampuan mengevaluasi sumber dengan keputusan mempercayai atau membagikan informasi. Guess et al (2020) menggunakan eksperimen lintas negara dan menemukan bahwa pelatihan literasi media digital meningkatkan kemampuan membedakan berita arus utama dan berita palsu. Pennycook et al (2020) menunjukkan bahwa pengingat sederhana untuk mempertimbangkan akurasi dapat memperbaiki kualitas informasi yang dipilih untuk dibagikan. Sementara itu, Brodsky et al (2021) memperlihatkan bahwa pengajaran membaca lateral membantu mahasiswa meninggalkan halaman awal, memeriksa reputasi sumber, mencari verifikasi eksternal, dan membandingkan beberapa dokumen. Pendekatan *prebunking* juga berkembang melalui inokulasi psikologis yang memperkenalkan teknik manipulasi sebelum pengguna berhadapan dengan pesan menyesatkan; intervensi video berskala besar terbukti meningkatkan ketahanan terhadap misinformasi (Roozenbeek et al., 2022). Pola ini menggunakan eksperimen, kuasi-eksperimen, tes pra-pasca, dan pengukuran perilaku. Kontribusinya kuat pada bukti efektivitas, tetapi intervensinya sering bersifat singkat, individual, dan belum dilembagakan dalam kebijakan komunikasi perguruan tinggi.

Kecenderungan ketiga berkembang pada persilangan literasi kecerdasan buatan, pendidikan tinggi, dan tata kelola komunikasi publik. Long & Magerko (2020) merumuskan kompetensi literasi kecerdasan buatan yang mencakup kemampuan mengenali sistem kecerdasan buatan, memahami mekanisme dasarnya, menilai kekuatan dan keterbatasan, serta mempertimbangkan implikasi sosial dan etis. Ng et al (2021) kemudian mengelompokkan literasi kecerdasan buatan ke dalam dimensi mengetahui dan memahami, menggunakan dan menerapkan, mengevaluasi dan mencipta, serta etika. Dalam pendidikan tinggi, perhatian bergeser pada kemampuan mahasiswa memeriksa halusinasi, bias, referensi palsu, dan probabilitas keluaran model, bukan sekadar keterampilan menggunakan aplikasi. Kasneci et al (2023) dan Dwivedi et al (2023) menegaskan bahwa model bahasa besar menawarkan manfaat pendidikan sekaligus risiko terhadap integritas, akurasi, dan ketergantungan kognitif. Pada tingkat kebijakan, OECD (2024), UNESCO (2021), dan *Artificial Intelligence Act* Uni Eropa menekankan transparansi, pemberdayaan pengguna, akuntabilitas, serta mitigasi risiko. Pola ini didominasi telaah

konseptual, tinjauan integratif, analisis kebijakan, dan dokumen normatif, tetapi belum banyak menghasilkan model operasional khusus bagi komunikasi institusi perguruan tinggi.

Ketiga kecenderungan tersebut telah memberikan dasar ilmiah penting, tetapi masih "melupakan" hubungan sistemik antara produksi disinformasi, kapasitas literasi mahasiswa, komunikasi institusi, dan tata kelola platform. Studi karakter ancaman menjelaskan bagaimana konten sintetis diproduksi serta diperkuat, tetapi belum selalu menunjukkan cara perguruan tinggi merespons secara terkoordinasi. Studi intervensi membuktikan efektivitas membaca lateral, pengingat akurasi, pemeriksaan fakta, dan *prebunking*, tetapi cenderung menempatkan mahasiswa sebagai individu yang harus memperbaiki dirinya sendiri. Kajian literasi kecerdasan buatan dan regulasi telah menawarkan kompetensi serta prinsip normatif, tetapi belum cukup menerjemahkannya menjadi pembagian peran, protokol respons, kanal klarifikasi, indikator evaluasi, dan mekanisme partisipasi mahasiswa. Dengan demikian, kebaruan penelitian ini terletak pada integrasi empat komponen yang selama ini terfragmentasi, yaitu regulasi komunikasi, tata kelola platform, literasi digital dan kecerdasan buatan mahasiswa, serta kapasitas komunikasi institusi pendidikan tinggi. Sintesis tersebut menggeser fokus dari koreksi konten menuju pembangunan ketahanan informasi.

Berdasarkan kesenjangan tersebut, penelitian ini menawarkan arah baru melalui perspektif kebijakan komunikasi publik adaptif berbasis institusi dan partisipasi mahasiswa. Fokus penelitian tidak hanya menilai kemampuan mahasiswa mengenali informasi palsu, tetapi menganalisis bagaimana perguruan tinggi dapat menghubungkan pencegahan, deteksi, respons, transparansi, dan kolaborasi dalam satu tata kelola. Orientasi preventif diwujudkan melalui integrasi literasi media dan kecerdasan buatan ke dalam kurikulum, orientasi mahasiswa, perpustakaan, serta organisasi kemahasiswaan. Orientasi adaptif menuntut pemetaan risiko, tim verifikasi lintas unit, komunikasi ketidakpastian, pembaruan berkala, dan arsip klarifikasi yang mudah diakses. Orientasi partisipatif menempatkan mahasiswa sebagai duta literasi, pemeriksa fakta sebaya, dan produsen klarifikasi, bukan semata-mata penerima kampanye. Sementara itu, orientasi transparan mengharuskan institusi menjelaskan sumber data, kriteria pelabelan, prosedur koreksi, serta mekanisme keberatan. Arah baru ini menghasilkan proposisi bahwa integrasi edukasi, regulasi, teknologi verifikasi, komunikasi institusi, dan jejaring eksternal akan memperkuat kemampuan mahasiswa menilai sumber, menunda pembagian informasi, serta membangun kepercayaan. Model tersebut selanjutnya dapat diuji melalui survei lintas kampus, eksperimen, dan evaluasi longitudinal.

METODE PENELITIAN

Unit analisis penelitian ini adalah artefak dokumenter yang membahas kebijakan komunikasi dalam menanggulangi disinformasi berbasis kecerdasan buatan di media sosial dan penguatan literasi digital mahasiswa. Artefak tersebut meliputi artikel penelitian empiris, artikel konseptual, laporan kebijakan lembaga internasional, regulasi, dan dokumen resmi yang diterbitkan pada 2020–2026. Fokus penelitian tidak diarahkan pada mahasiswa sebagai individu yang diwawancarai atau diukur secara langsung, tetapi pada konsep, data, temuan, argumentasi, dan rekomendasi kebijakan yang terdapat dalam setiap dokumen. Setiap publikasi diperlakukan sebagai unit sumber, sedangkan pernyataan yang membahas produksi dan amplifikasi disinformasi, kerentanan mahasiswa, intervensi literasi, komunikasi publik, serta tata kelola kolaboratif diperlakukan sebagai unit makna. Objek substantif penelitian adalah institusi perguruan tinggi sebagai lingkungan akademik dan komunikasi yang memiliki tanggung jawab membangun ketahanan informasi mahasiswa. Penetapan unit analisis tersebut memungkinkan penelitian menghubungkan aspek teknologi, perilaku pengguna, pendidikan literasi, kapasitas komunikasi institusi, dan kebijakan platform dalam satu kerangka analitis, tanpa mengklaim menggambarkan tingkat literasi mahasiswa pada kampus tertentu.

Penelitian ini menggunakan orientasi kualitatif dengan desain kajian literatur integratif. Pendekatan kualitatif dipilih karena penelitian bertujuan memahami, membandingkan, dan menyintesiskan pola konseptual serta rekomendasi kebijakan, bukan menguji hubungan antarvariabel menggunakan perhitungan statistik. Kajian literatur integratif memungkinkan penggabungan bukti yang berasal dari pendekatan berbeda, termasuk eksperimen tentang misinformasi, penelitian literasi media, kajian konseptual kecerdasan buatan, analisis kebijakan, dan dokumen regulasi. Desain tersebut sesuai dengan karakter persoalan yang bersifat multidisipliner karena disinformasi berbasis kecerdasan buatan melibatkan ilmu komunikasi, pendidikan, psikologi, ilmu informasi, kebijakan publik, dan tata kelola teknologi. Torraco (2016) menjelaskan bahwa kajian integratif dapat menggunakan literatur terdahulu untuk membangun pengetahuan dan kerangka konseptual baru, sedangkan Snyder (2019) menempatkan kajian literatur sebagai metodologi penelitian yang memerlukan prosedur pencarian, pemilihan, analisis, dan sintesis secara sistematis. Penelitian ini bukan metode campuran karena tidak mengumpulkan data kuantitatif primer, meskipun temuan penelitian kuantitatif terdahulu tetap dianalisis sebagai bagian dari korpus bukti.

Data penelitian seluruhnya berupa data sekunder berbentuk teks yang berasal dari 34 publikasi yang diterbitkan pada 2020–2026. Sumber tersebut mencakup artikel jurnal bereputasi, laporan kebijakan lembaga internasional, regulasi, dan dokumen resmi yang menyediakan konsep, data empiris, hasil eksperimen, atau rekomendasi operasional.

Publikasi dipilih apabila berhubungan langsung dengan sedikitnya satu fokus penelitian, yaitu karakter disinformasi berbasis kecerdasan buatan, perilaku verifikasi, literasi digital mahasiswa, intervensi pendidikan, kebijakan komunikasi publik, atau tata kelola platform. Artikel empiris digunakan untuk memahami kerentanan pengguna dan efektivitas intervensi, sedangkan laporan kebijakan serta regulasi digunakan untuk mengidentifikasi prinsip tata kelola, transparansi, akuntabilitas, dan pembagian tanggung jawab kelembagaan. Publikasi sebelum 2020 tidak dimasukkan agar korpus kajian merepresentasikan perkembangan media sosial dan kecerdasan buatan generatif yang lebih mutakhir. Penelitian ini tidak menggunakan informan, responden, hasil wawancara, rekaman dialog, maupun data audiovisual sebagai data primer. Meskipun beberapa publikasi membahas gambar, audio sintetis, video, dan *deepfake*, data penelitian tetap berupa penjelasan dan temuan tertulis mengenai fenomena tersebut.

Pengumpulan data dilakukan melalui penelusuran, penyaringan, pembacaan, dan pencatatan dokumen secara bertahap. Tahap pertama adalah menetapkan fokus penelitian dan kombinasi kata kunci, yaitu *artificial intelligence disinformation, generative AI misinformation, deepfake literacy, media and information literacy, university students, public communication policy, information integrity, prebunking, fact-checking*, dan *platform governance*. Tahap kedua adalah menelusuri publikasi melalui basis data ilmiah dan situs lembaga resmi. Tahap ketiga dilakukan dengan memeriksa judul, abstrak, tahun penerbitan, relevansi pembahasan, kualitas akademik, serta ketersediaan konsep, data, atau rekomendasi yang dapat mendukung analisis. Tahap keempat adalah membaca teks lengkap dan menerapkan kriteria inklusi serta eksklusi sampai diperoleh 34 sumber akhir. Informasi dari setiap sumber kemudian dicatat dalam lembar ekstraksi yang memuat identitas publikasi, tujuan, pendekatan, metode, populasi atau konteks, temuan utama, keterbatasan, dan rekomendasi. Karena merupakan penelitian kepustakaan, pengumpulan data tidak dilakukan melalui observasi lapangan, survei, wawancara, kuesioner, diskusi kelompok terpumpun, dialog, atau *talkshow*. Prosedur yang transparan diperlukan agar kajian dapat ditelusuri dan dipertanggungjawabkan secara akademik (Snyder, 2019).

Analisis data dilakukan melalui tiga tahap utama yang dikembangkan menjadi proses interpretatif berurutan. Pertama, peneliti mengekstraksi informasi dari setiap publikasi mengenai tujuan, konsep, metode, populasi, konteks, temuan, dan rekomendasi. Kedua, data tekstual dikodekan secara tematik ke dalam lima kategori awal, yaitu produksi dan amplifikasi disinformasi, kerentanan mahasiswa, intervensi literasi, komunikasi publik, dan tata kelola kolaboratif. Pengodean dilakukan dengan mengidentifikasi pernyataan yang berulang, hubungan antarkonsep, perbedaan hasil, serta penjelasan mengenai faktor yang mendukung atau menghambat kebijakan. Ketiga, temuan antarsumber dibandingkan untuk menemukan pola konsisten, kontradiksi, kesenjangan penelitian, dan implikasi

kelembagaan. Hasil perbandingan kemudian disintesiskan menjadi model kebijakan komunikasi publik adaptif yang mencakup pencegahan, deteksi, respons, transparansi, dan kolaborasi. Analisis tematik tidak hanya menghitung frekuensi istilah, tetapi menafsirkan makna serta hubungan antartema secara reflektif (Braun & Clarke, 2021). Keandalan analisis dijaga melalui triangulasi jenis sumber dengan membandingkan penelitian eksperimen, tinjauan ilmiah, laporan kebijakan, dan regulasi, serta memastikan setiap kesimpulan memiliki dukungan dokumenter yang relevan.

RESULTS AND DISCUSSION

Result

Transformasi Ancaman Disinformasi Berbasis Kecerdasan Buatan

Hasil sintesis terhadap 34 publikasi menunjukkan bahwa disinformasi berbasis kecerdasan buatan telah mengubah gangguan informasi dari produksi terbatas menjadi proses yang murah, cepat, multimodal, dan mudah dipersonalisasi. Data tekstual yang dibaca memperlihatkan empat bentuk utama: pembuatan narasi dalam banyak versi, penggunaan gambar atau video sintetis, pembentukan identitas dan akun buatan, serta distribusi melalui sistem rekomendasi yang mengutamakan keterlibatan. Literatur juga menunjukkan bahwa ancaman tidak berhenti pada keberhasilan membuat informasi palsu tampak benar. Konten sintetis dapat menimbulkan ketidakpastian terhadap bukti autentik karena pihak yang berkepentingan dapat menyangkal rekaman asli sebagai rekayasa. Vaccari & Chadwick (2020) menemukan bahwa video politik sintetis terutama meningkatkan ketidakpastian dan menurunkan kepercayaan pada berita, sedangkan penelitian mengenai persuasi menunjukkan bahwa pesan yang dibuat oleh model generatif dapat disesuaikan dengan karakter penerima sehingga lebih berpengaruh daripada pesan umum. Bukti tersebut menegaskan adanya transformasi dan keadaan mendesak sehingga kecerdasan buatan bukan sekadar alat produksi konten, melainkan pengganda kapasitas persuasi, penyamaran, dan amplifikasi disinformasi di media sosial.

Tabel 1. Transformasi Disinformasi Berbasis Kecerdasan Buatan

Dimensi	Data yang Teridentifikasi	Perubahan yang Terjadi	Risiko Utama
Produksi	Teks, gambar, audio, dan video sintetis	Produksi cepat dan berskala besar	Banjir konten menyesatkan
Personalisasi	Pesan disesuaikan dengan profil sasaran	Persuasi menjadi lebih terarah	Manipulasi sikap dan keputusan
Identitas	Akun, figur, suara, dan dokumen sintetis	Identitas palsu semakin meyakinkan	Penipuan dan legitimasi semu
Distribusi	Bot dan rekomendasi algoritmik	Penyebaran berulang dan lintas platform	Amplifikasi konten emosional
Kepercayaan	Bukti autentik dapat dituduh sebagai rekayasa	Meningkatnya ketidakpastian epistemik	Sinisme dan hilangnya konsensus

Secara sederhana, data tersebut berarti bahwa masalah disinformasi berbasis kecerdasan buatan tidak dapat dinilai hanya dari benar atau salahnya satu unggahan. Sebuah pesan menyesatkan dapat diproduksi dalam bentuk teks, gambar, audio, atau video, diterjemahkan ke berbagai bahasa, disesuaikan dengan kelompok tertentu, lalu disebarkan kembali melalui banyak akun. Konten yang tampil rapi, lancar, dan menyerupai produk jurnalistik dapat memperoleh legitimasi semu meskipun sumber serta buktinya lemah. Mahasiswa kemudian menghadapi beban verifikasi yang lebih besar karena penilaian berdasarkan tata bahasa, kualitas desain, atau kelancaran penyampaian tidak lagi cukup. Studi terhadap mahasiswa jurnalistik di Portugal dan Spanyol bahkan menunjukkan bahwa peserta cenderung memberi penilaian kualitas lebih tinggi kepada berita buatan kecerdasan buatan daripada berita yang ditulis jurnalis, sekaligus mengalami kesulitan membedakan keduanya. Dengan demikian, inti persoalan bukan hanya keberadaan konten palsu, tetapi perubahan lingkungan informasi yang membuat tanda-tanda tradisional tentang kredibilitas semakin tidak dapat diandalkan. Kebijakan komunikasi perlu menempatkan asal-usul, konteks, bukti, dan pola distribusi sebagai objek pemeriksaan utama.

Deskripsi lintas sumber menghasilkan empat pola pada bukti pertama. Pertama, terjadi industrialisasi disinformasi karena model generatif memungkinkan produksi pesan dalam jumlah besar dengan biaya dan waktu yang lebih rendah. Kedua, berkembang

manipulasi multimodal yang menggabungkan teks, gambar, audio, video, tangkapan layar, dan identitas sintetis sehingga pemeriksaan tidak dapat bergantung pada satu jenis keterampilan. Ketiga, personalisasi meningkatkan daya persuasi karena bahasa dan argumen dapat disesuaikan dengan profil psikologis, preferensi, atau identitas sasaran. Penelitian Metzler & Garcia (2024) menunjukkan pesan yang dipersonalisasi oleh model bahasa lebih berpengaruh daripada pesan yang tidak dipersonalisasi. Keempat, amplifikasi algoritmik memberi keunggulan pada konten sederhana, emosional, dan mudah dibagikan dibandingkan klarifikasi yang membutuhkan konteks. Pola tersebut menjelaskan mengapa penghapusan satu konten tidak menyelesaikan masalah sehingga salinan, variasi, dan akun baru dapat segera muncul. Kesimpulan data ini adalah bahwa objek kebijakan harus bergeser dari moderasi unggahan individual menuju pengelolaan risiko sistemik yang mencakup produksi, distribusi, personalisasi, dan dampak terhadap kepercayaan publik.

Kerentanan dan Kapasitas Literasi Digital Mahasiswa

Bukti kedua berkaitan dengan kerentanan dan kapasitas literasi digital mahasiswa. Dokumen yang dianalisis menunjukkan bahwa kemampuan menggunakan perangkat, aplikasi, media sosial, dan sistem kecerdasan buatan tidak selalu sejalan dengan kemampuan mengevaluasi sumber. Data literatur mengidentifikasi bias konfirmasi, tekanan untuk membagikan informasi dengan cepat, ketergantungan pada jumlah suka atau dukungan figur populer, serta kecenderungan menganggap keluaran yang lancar sebagai informasi akurat. Pada saat yang sama, bukti eksperimen menunjukkan bahwa kemampuan tersebut dapat diperbaiki. Guess et al (2020) menemukan bahwa intervensi literasi media digital meningkatkan kemampuan membedakan berita arus utama dan berita palsu. Brodsky et al (2021) menunjukkan bahwa pengajaran membaca lateral membantu mahasiswa memeriksa reputasi sumber melalui situs lain. Pennycook et al (2020) menemukan bahwa pengingat untuk mempertimbangkan akurasi dapat memperbaiki kualitas informasi yang dibagikan, sedangkan Roozenbeek et al (2022) membuktikan bahwa inokulasi terhadap teknik manipulasi meningkatkan ketahanan. Data tersebut memperlihatkan solusi yang konsisten sehingga kompetensi verifikasi perlu diajarkan melalui latihan, pengulangan, dan konteks yang menyerupai konsumsi informasi nyata.

Tabel 2. Kerentanan dan Intervensi Literasi Mahasiswa

Kerentanan	Bentuk Perilaku	Intervensi	Kemampuan yang Dikembangkan
Bias konfirmasi	Menerima pesan yang sesuai keyakinan	Analisis kasus dan sumber berlawanan	Penalaran reflektif
Tekanan kecepatan	Membagikan sebelum memeriksa	Pengingat akurasi	Penundaan berbagi
Heuristik sosial	Percaya karena banyak suka atau komentar	Membaca lateral	Evaluasi reputasi sumber
Manipulasi visual	Percaya pada gambar dan video yang tampak nyata	Penelusuran gambar dan konteks	Verifikasi multimodal
Kepercayaan pada keluaran AI	Menganggap bahasa lancar sebagai fakta	Literasi kecerdasan buatan	Evaluasi bias dan halusinasi
Resistensi terhadap koreksi	Menolak klarifikasi yang menghakimi	Komunikasi korektif	Koreksi berbasis bukti dan empati

Penyajian ulang bukti kedua menunjukkan bahwa mahasiswa tidak cukup hanya diberi penjelasan mengenai bahaya hoaks atau daftar ciri konten palsu. Mereka perlu mempraktikkan cara membuka sumber pembanding, memeriksa reputasi penerbit, menelusuri gambar, membandingkan tanggal dan konteks, memeriksa kutipan, serta menilai kemungkinan halusinasi dan bias dalam keluaran kecerdasan buatan. Pengingat akurasi dapat ditempatkan sebelum mahasiswa membagikan konten, sedangkan *prebunking* dapat digunakan untuk memperkenalkan pola manipulasi seperti otoritas semu, penggunaan emosi, dikotomi palsu, identitas tiruan, dan bukti yang dipotong. Literasi juga harus mencakup komunikasi korektif karena klarifikasi yang mempermalukan pengguna dapat menimbulkan resistensi. Koreksi yang lebih baik menyajikan informasi pengganti, menjelaskan alasan suatu klaim keliru, menggunakan bahasa yang tidak menghakimi, dan menyediakan sumber yang dapat diperiksa. Dengan demikian, mahasiswa diposisikan sebagai penerima, evaluator, produsen, dan penyebar informasi yang bertanggung jawab. Pembelajaran perlu terhubung dengan mata kuliah, orientasi mahasiswa, perpustakaan, organisasi kemahasiswaan, dan kebijakan penggunaan kecerdasan buatan agar kompetensi tidak berhenti setelah satu seminar.

Deskripsi data menghasilkan empat kecenderungan pada bukti kedua. Pertama, intervensi aktif lebih efektif daripada penyampaian pasif karena mahasiswa menjalankan sendiri proses mencari, membandingkan, dan menilai bukti. Kedua, ketahanan informasi dibentuk oleh kombinasi keterampilan kognitif dan kebiasaan perilaku; kemampuan mengetahui sumber tepercaya belum tentu digunakan ketika pengguna berada di bawah tekanan waktu atau emosi. Ketiga, literasi kecerdasan buatan memperluas literasi media dengan pemahaman mengenai data pelatihan, probabilitas keluaran, bias, halusinasi, serta tanggung jawab manusia atas keputusan akhir. Keempat, pembelajaran sebaya berpotensi memperluas jangkauan karena mahasiswa menggunakan bahasa dan saluran yang dekat dengan komunitasnya. UNESCO (2024) menempatkan pemberdayaan pengguna dan integrasi literasi media serta informasi dalam pendidikan sebagai unsur penting respons terhadap kecerdasan buatan generatif. Kesimpulan data ini ialah bahwa literasi tidak boleh diperlakukan sebagai atribut individual yang tetap, tetapi sebagai kapasitas yang dibentuk melalui latihan, lingkungan institusi, dan kesempatan berpartisipasi. Karena itu, indikator keberhasilan perlu mengukur evaluasi sumber, penundaan berbagi, kualitas koreksi, dan penggunaan bukti, bukan hanya jumlah peserta atau sertifikat.

Kebutuhan Model Kebijakan Komunikasi Publik Adaptif

Bukti ketiga menunjukkan kelemahan kebijakan komunikasi yang bersifat reaktif dan terfragmentasi. Data yang dibaca memperlihatkan bahwa klarifikasi sering diterbitkan setelah konten menjadi viral, ketika narasi menyesatkan telah memperoleh dukungan emosional dan sosial. Penghapusan unggahan juga tidak menghapus tangkapan layar, salinan lintas platform, atau versi baru yang dapat diproduksi secara otomatis. Pada tingkat institusi, unit hubungan masyarakat dapat menyusun bantahan tanpa dukungan analisis digital, unit teknologi informasi menangani keamanan tanpa memahami struktur narasi, sedangkan dosen membahas literasi secara terpisah dari protokol komunikasi kampus. Sebagai respons, sintesis menghasilkan model kebijakan komunikasi publik adaptif yang mencakup lima lapisan yaitu pemantauan dan pemetaan risiko, verifikasi serta respons cepat, *prebunking* dan pendidikan, transparansi serta partisipasi, dan kolaborasi eksternal. Kerangka pada Tabel 3 menghubungkan setiap masalah dengan instrumen serta indikator kebijakan. OECD (2024) mendukung pendekatan sistemik melalui tiga dimensi: transparansi dan pluralitas sumber, ketahanan masyarakat, serta peningkatan kapasitas tata kelola dan institusi publik untuk menjaga integritas ruang informasi.

Tabel 3. Kerangka Kebijakan Komunikasi Publik Adaptif

Dimensi	Masalah Utama	Instrumen Kebijakan	Indikator
Pencegahan	Mahasiswa belum memahami manipulasi AI	Kurikulum literasi AI, *prebunking*, simulasi verifikasi	Skor evaluasi sumber dan penundaan berbagi
Deteksi	Konten sintetis menyebar cepat	Pemantauan isu, kanal pelaporan, alat penelusuran	Waktu deteksi dan laporan terverifikasi
Respons	Klarifikasi terlambat dan tidak konsisten	Protokol lintas unit, pusat informasi, pembaruan berkala	Waktu respons dan jangkauan klarifikasi
Transparansi	Ketidakpercayaan terhadap moderasi	Kriteria pelabelan, arsip koreksi, mekanisme keberatan	Kepercayaan dan penyelesaian pengaduan
Partisipasi	Mahasiswa hanya menjadi objek edukasi	Duta literasi dan pemeriksa fakta sebaya	Keterlibatan dan produksi konten edukatif
Kolaborasi	Kapasitas lembaga terfragmentasi	Kemitraan pemerintah, kampus, media, dan platform	Pertukaran data dan kegiatan bersama

Dalam bentuk yang lebih mudah dipahami, bukti ketiga menyatakan bahwa perguruan tinggi harus bertindak sebelum, selama, dan setelah munculnya disinformasi. Sebelum krisis, institusi perlu memetakan isu berisiko, mengajarkan teknik manipulasi, membentuk pedoman penggunaan kecerdasan buatan, dan melatih tim lintas unit. Ketika isu berkembang, tim perlu memeriksa klaim, menghubungi sumber internal, menjelaskan fakta yang telah diketahui, menyatakan bagian yang masih diverifikasi, serta menetapkan waktu pembaruan berikutnya. Setelah respons, kampus perlu mengarsipkan klarifikasi, menyediakan mekanisme keberatan, mengevaluasi waktu respons, dan memperbaiki protokol. Mahasiswa dapat dilibatkan sebagai duta literasi, pemeriksa fakta sebaya, atau anggota laboratorium komunikasi digital, tetapi kegiatan tersebut harus memiliki supervisi akademik serta batas yang jelas agar tidak berubah menjadi pengawasan pendapat pribadi. Kolaborasi dengan pemerintah, media, platform, organisasi pemeriksa fakta, dan masyarakat sipil diperlukan karena tidak satu pun lembaga memiliki seluruh data, teknologi, serta keahlian. Restatement ini menegaskan bahwa kebijakan komunikasi bukan

hanya produk aturan, melainkan rangkaian kapasitas, layanan, relasi, dan prosedur yang bekerja secara berkelanjutan.

Deskripsi terhadap bukti ketiga menghasilkan empat pola. Pertama, pencegahan lebih kuat ketika literasi media dan kecerdasan buatan diintegrasikan ke kurikulum lintas disiplin, sebab bentuk disinformasi berbeda pada kesehatan, ekonomi, pendidikan, politik, dan bidang lainnya. Kedua, respons yang cepat tetap harus transparan sehingga institusi dapat mengakui ketidakpastian dan memberi jadwal pembaruan daripada membiarkan kekosongan informasi diisi spekulasi. Ketiga, partisipasi mahasiswa meningkatkan relevansi dan jangkauan komunikasi, tetapi memerlukan perlindungan privasi, etika, dan mekanisme pertanggungjawaban. Keempat, kolaborasi lintas unit dan lintas lembaga mengurangi fragmentasi sumber daya serta mempercepat pertukaran data dan verifikasi. Pola tersebut menunjukkan bahwa ketahanan informasi tidak dihasilkan oleh satu aplikasi pendeteksi, satu mata kuliah, atau satu tim hubungan masyarakat. Ketahanan muncul dari kesesuaian antara regulasi, pendidikan, komunikasi institusi, teknologi verifikasi, dan jejaring eksternal. Kesimpulan data ini menegaskan solusi utama penelitian yaitu kebijakan harus bergerak dari koreksi konten menuju tata kelola preventif, adaptif, partisipatif, transparan, dan berbasis bukti, dengan evaluasi yang mengukur perilaku serta kualitas respons.

Discussion

Penelitian ini menganalisis bagaimana kebijakan komunikasi dapat menanggulangi disinformasi berbasis kecerdasan buatan di media sosial melalui penguatan literasi digital mahasiswa. Sintesis terhadap 34 publikasi menghasilkan tiga temuan utama. Pertama, kecerdasan buatan mentransformasikan disinformasi melalui produksi otomatis, manipulasi multimodal, identitas sintetis, personalisasi pesan, dan amplifikasi algoritmik, sehingga ancaman tidak lagi terbatas pada satu konten palsu. Kedua, mahasiswa memiliki akses dan keterampilan teknis, tetapi tetap rentan terhadap bias konfirmasi, tekanan kecepatan, heuristik sosial, serta kecenderungan menganggap keluaran yang lancar sebagai fakta sehingga kerentanan tersebut dapat dikurangi melalui membaca lateral, pemeriksaan fakta, pengingat akurasi, *prebunking*, dan literasi kecerdasan buatan. Ketiga, kebijakan reaktif yang bertumpu pada penghapusan konten dan klarifikasi terlambat tidak memadai. Penelitian menghasilkan model komunikasi publik adaptif yang menggabungkan pencegahan, deteksi, respons, transparansi, partisipasi, dan kolaborasi. Model tersebut menempatkan perguruan tinggi sebagai institusi akademik sekaligus aktor komunikasi yang bertanggung jawab membangun ketahanan informasi mahasiswa. Kerangka ini sekaligus mempertemukan dimensi teknologi, psikologi pengguna, pendidikan, dan tata kelola kelembagaan dalam satu penjelasan yang utuh.

Hubungan antara kecerdasan buatan, kerentanan mahasiswa, dan kebijakan komunikasi dapat dijelaskan melalui interaksi teknologi, psikologi, dan kelembagaan. Kecerdasan buatan menurunkan biaya produksi pesan serta memungkinkan variasi dan personalisasi dalam skala besar. Metzler & Garcia (2024) menunjukkan bahwa pesan yang disesuaikan dengan profil penerima lebih berpengaruh daripada pesan umum, sedangkan Salvi et al (2025) menemukan bahwa GPT-4 dengan akses pada informasi personal dapat memiliki daya persuasi lebih tinggi dalam debat daring. Pada sisi pengguna, bias konfirmasi, respons emosional, dan tekanan kecepatan mengurangi kemungkinan verifikasi. Pada sisi platform, sistem rekomendasi dapat memperkuat konten yang menghasilkan keterlibatan. Karena itu, peningkatan kemampuan individual saja tidak cukup apabila lingkungan komunikasi tetap mendorong penyebaran cepat dan institusi terlambat menyediakan informasi. Pendidikan membentuk keterampilan, pengingat akurasi mengaktifkan perhatian, *prebunking* menyiapkan pola pengenalan manipulasi, respons cepat mengurangi kekosongan informasi, dan transparansi menjaga kepercayaan. Hubungan tersebut menjelaskan mengapa kebijakan terintegrasi lebih berpeluang meningkatkan ketahanan informasi daripada intervensi yang berdiri sendiri.

Temuan penelitian memiliki kesamaan dengan studi terdahulu, tetapi juga memperlihatkan perbedaan yang menjadi kebaruannya. Kesamaan pertama ialah pengakuan bahwa kedekatan dengan teknologi tidak otomatis menghasilkan evaluasi sumber yang baik, sebagaimana ditunjukkan penelitian literasi digital dan membaca lateral. Kesamaan kedua ialah bukti bahwa pengingat akurasi, pemeriksaan fakta, dan inokulasi psikologis dapat meningkatkan ketelitian pengguna (Guess et al., 2020). Kesamaan ketiga terdapat pada rekomendasi OECD (2024) dan UNESCO (2024) mengenai transparansi platform, ketahanan masyarakat, pemberdayaan pengguna, dan integrasi literasi media serta informasi dalam pendidikan. Perbedaannya, penelitian terdahulu cenderung memisahkan karakter teknologi, intervensi individual, literasi kecerdasan buatan, dan tata kelola institusi. Penelitian ini menyatukan keempat bidang tersebut dalam model operasional perguruan tinggi yang menjelaskan tahap, pelaksana, instrumen, dan indikator. Kebaruan lainnya adalah penempatan mahasiswa sebagai pemeriksa fakta sebaya dan mitra komunikasi, bukan hanya sasaran edukasi. Dengan demikian, fokus bergeser dari kemampuan mengenali konten palsu menuju kapasitas institusional untuk menjaga integritas informasi.

Makna temuan penelitian dapat ditafsirkan melalui perubahan historis dari kelangkaan informasi menuju kelimpahan informasi sintetis. Pada masa komunikasi massa, produksi pesan publik terutama dikendalikan oleh institusi media, sedangkan media sosial mengubah pengguna menjadi produsen dan distributor. Kecerdasan buatan generatif memperluas perubahan tersebut karena pengguna tidak lagi harus memiliki keterampilan

tinggi untuk menghasilkan teks, gambar, audio, atau video yang meyakinkan. Secara sosial, kondisi ini memperbesar partisipasi sekaligus mengaburkan batas antara kesaksian, dokumentasi, opini, dan rekayasa. Secara ideologis, ekonomi perhatian mendorong platform menilai informasi berdasarkan keterlibatan, sementara nilai akademik menuntut ketelitian, bukti, transparansi, dan keterbukaan terhadap koreksi. OECD (2024) menegaskan bahwa digitalisasi memperbesar jangkauan disinformasi dan bahwa integritas informasi memerlukan sumber yang plural, tata kelola, dan ketahanan masyarakat. Oleh karena itu, model adaptif bermakna sebagai upaya mempertahankan kebebasan berekspresi tanpa menyerahkan ruang publik kepada manipulasi. Tujuannya bukan menyeragamkan pendapat, melainkan memastikan klaim faktual dapat diperiksa dan keputusan komunikasi dapat dipertanggungjawabkan. Dalam konteks perguruan tinggi, verifikasi menjadi praktik kewargaan akademik, bukan sekadar keterampilan teknis tambahan.

Secara reflektif, penerapan model kebijakan adaptif memiliki fungsi dan potensi disfungsi. Fungsinya ialah memperkuat kemampuan mahasiswa menilai sumber, meningkatkan kecepatan klarifikasi, membangun koordinasi lintas unit, serta menjadikan literasi digital bagian dari budaya akademik. Partisipasi mahasiswa dapat memperluas jangkauan pesan dan menumbuhkan rasa memiliki terhadap integritas informasi. Transparansi prosedur juga dapat meningkatkan kepercayaan karena pengguna mengetahui alasan pelabelan, koreksi, atau moderasi. Namun, pemantauan isu dapat berubah menjadi pengawasan terhadap pendapat pribadi apabila batasnya tidak jelas. Kebijakan yang terlalu menekankan sanksi dapat mengkriminalisasi kesalahan, membatasi kritik, dan menghasilkan kepatuhan semu. Alat deteksi otomatis juga dapat menimbulkan rasa aman palsu karena hasilnya tetap memiliki kesalahan dan memerlukan penilaian manusia. Selain itu, mahasiswa dengan akses teknologi atau kemampuan yang lebih rendah dapat menghadapi beban lebih besar. Karena itu, evaluasi kebijakan harus mencakup efektivitas, keadilan, privasi, aksesibilitas, kebebasan akademik, dan mekanisme keberatan. Ketahanan informasi harus dibangun melalui pemberdayaan, bukan ketakutan. Prinsip proporsionalitas, transparansi, dan pemulihan harus menjadi pagar agar tujuan perlindungan tidak berubah menjadi kontrol yang berlebihan.

Implikasi hasil penelitian diwujudkan melalui rencana aksi enam tahap. Pertama, perguruan tinggi melakukan audit risiko informasi dengan memetakan isu, saluran, kelompok sasaran, kapasitas unit, dan kebutuhan mahasiswa. Kedua, institusi menetapkan standar kompetensi yang mencakup membaca lateral, verifikasi visual, penilaian sumber, pemahaman algoritma, evaluasi keluaran kecerdasan buatan, dan komunikasi korektif. Ketiga, kompetensi tersebut diintegrasikan ke dalam mata kuliah, orientasi mahasiswa, perpustakaan, pelatihan dosen, dan organisasi kemahasiswaan. Keempat, kampus

membentuk tim verifikasi lintas fungsi dengan protokol yang menetapkan alur laporan, batas waktu respons, sumber otoritatif, format pembaruan, dan arsip klarifikasi. Kelima, institusi melibatkan mahasiswa sebagai duta literasi atau pemeriksa fakta sebaya serta membangun kemitraan dengan pemerintah, media, platform, dan organisasi masyarakat sipil. Keenam, penjaminan mutu mengevaluasi skor verifikasi, perilaku berbagi, waktu deteksi, jangkauan klarifikasi, penyelesaian pengaduan, dan tingkat kepercayaan. Data evaluasi digunakan untuk memperbaiki kebijakan secara berkala dan mencegah program berhenti pada seminar seremonial. Pelaksanaan bertahap memungkinkan kampus memulai dari uji coba terbatas, menilai hasil, lalu memperluas program berdasarkan bukti.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan, kesimpulan terpenting dari penelitian ini menunjukkan bahwa disinformasi berbasis kecerdasan buatan tidak dapat ditanggulangi hanya dengan penghapusan konten atau klarifikasi setelah informasi menjadi viral. Hikmah utama yang dapat diambil ialah bahwa kedekatan mahasiswa dengan teknologi tidak otomatis menghasilkan ketahanan informasi, sebab keterampilan menggunakan aplikasi berbeda dari kemampuan memeriksa sumber, memahami konteks, mengenali manipulasi multimodal, dan menilai keterbatasan keluaran kecerdasan buatan. Disinformasi kini bekerja melalui produksi otomatis, identitas sintetis, personalisasi pesan, amplifikasi algoritmik, serta penciptaan keraguan terhadap bukti autentik. Karena itu, respons yang lebih tepat harus bersifat preventif, adaptif, partisipatif, transparan, dan berbasis bukti. Perguruan tinggi perlu menggabungkan pemetaan risiko, respons cepat, *prebunking*, literasi media dan kecerdasan buatan, verifikasi informasi, transparansi moderasi, serta partisipasi mahasiswa sebagai pemeriksa fakta sebaya. Pelajaran penting lainnya ialah bahwa ketahanan informasi bukan hanya tanggung jawab individu, tetapi hasil dari hubungan antara pendidikan, teknologi, komunikasi institusi, tata kelola, dan kolaborasi lintas lembaga.

Kekuatan penelitian ini terletak pada kontribusinya dalam menyatukan temuan yang sebelumnya tersebar dalam kajian disinformasi, literasi digital, literasi kecerdasan buatan, komunikasi publik, dan tata kelola platform. Secara keilmuan, penelitian ini menyumbangkan model kebijakan komunikasi publik adaptif yang menghubungkan lima lapisan utama, yaitu pemantauan dan pemetaan risiko, verifikasi dan respons cepat, *prebunking* dan pendidikan, transparansi dan partisipasi, serta kolaborasi eksternal. Model tersebut memperluas analisis dari sekadar hubungan antara konten palsu dan kerentanan pengguna menuju hubungan yang lebih sistemik antara kemampuan teknologi, perilaku mahasiswa, kapasitas institusi, dan kebijakan komunikasi. Penelitian ini juga menawarkan pendekatan baru dengan menempatkan mahasiswa bukan hanya sebagai penerima

edukasi, tetapi sebagai mitra dalam pemeriksaan fakta, produksi klarifikasi, dan penguatan integritas informasi kampus. Selain itu, kajian ini membuka pertanyaan baru mengenai bagaimana integrasi literasi digital, respons kelembagaan, dan transparansi platform memengaruhi perilaku verifikasi, penundaan berbagi, kualitas koreksi, serta tingkat kepercayaan mahasiswa terhadap komunikasi institusi.

Keterbatasan penelitian ini terutama berasal dari penggunaan desain kajian literatur integratif yang sepenuhnya bergantung pada data sekunder berupa artikel ilmiah, laporan kebijakan, regulasi, dan dokumen resmi. Penelitian belum mengumpulkan data empiris melalui survei, wawancara, observasi, eksperimen, atau diskusi kelompok terpumpun pada mahasiswa di perguruan tinggi tertentu. Karena itu, model yang dihasilkan belum dapat menunjukkan secara langsung tingkat literasi mahasiswa, variasi kerentanan antarkelompok, efektivitas kebijakan pada konteks kampus yang berbeda, atau perubahan perilaku setelah intervensi dilakukan. Korpus penelitian juga dibatasi pada 34 publikasi periode 2020–2026 sehingga hasilnya lebih merepresentasikan perkembangan mutakhir kecerdasan buatan generatif, tetapi belum mencakup seluruh konteks sosial, budaya, disiplin ilmu, dan sistem pendidikan tinggi. Penelitian selanjutnya perlu menguji model melalui eksperimen pendidikan, survei lintas kampus, studi kasus multisitus, dan evaluasi longitudinal. Kajian lanjutan juga perlu menilai dampak privasi, potensi pengawasan berlebihan, efektivitas pemeriksa fakta sebaya, serta keseimbangan antara penanggulangan disinformasi, kebebasan akademik, transparansi, dan perlindungan hak mahasiswa.

DAFTAR PUSTAKA

1. Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238
2. Brodsky, J. E., Brooks, P. J., Scimeca, D., Todorova, R., Galati, P., Batson, M., Grosso, R., Matthews, M., Miller, V., & Caulfield, M. (2021). Improving college students' fact-checking strategies through lateral reading instruction in a general education civics course. *Cognitive Research: Principles and Implications*, *6*(1), 23. https://doi.org/10.1186/s41235-021-00291-4
3. Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). "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

4. Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(27), 15536–15545. https://doi.org/10.1073/pnas.1920498117
5. 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*, 102274. https://doi.org/10.1016/j.lindif.2023.102274
6. Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools. *Psychological Science in the Public Interest*, *21*(3), 103–156. https://doi.org/10.1177/1529100620946707
7. Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Conference on Human Factors in Computing Systems - Proceedings*, 1–16. https://doi.org/10.1145/3313831.3376727
8. Metzler, H., & Garcia, D. (2024). Social Drivers and Algorithmic Mechanisms on Digital Media. *Perspectives on Psychological Science*, *19*(5), 735–748. https://doi.org/10.1177/17456916231185057
9. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041
10. OECD. (2024). Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity. In *Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity*. OECD Publishing. https://doi.org/10.1787/d909ff7a-en
11. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science*, *31*(7), 770–780. https://doi.org/10.1177/0956797620939054
12. Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, *8*(34), 6254. https://doi.org/10.1126/sciadv.abo6254
13. Salvi, F., Horta Ribeiro, M., Gallotti, R., & West, R. (2025). On the conversational persuasiveness of GPT-4. *Nature Human Behaviour*, *9*(8), 1645–1653. https://doi.org/10.1038/s41562-025-02194-6
14. Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339.

https://doi.org/10.1016/j.jbusres.2019.07.039

15. Torraco, R. J. (2016). Writing Integrative Literature Reviews: Using the Past and Present to Explore the Future. *Human Resource Development Review*, *15*(4), 404–428. https://doi.org/10.1177/1534484316671606
16. UNESCO. (2021). Recommendation on The Ethics of Artificial Intelligence. In *Bots and Beasts* (Issue November). UNESCO.
17. UNESCO. (2024). *User empowerment through Media and Information Literacy responses to the evolution of Generative Artificial Intelligence (GAI)*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000388547
18. Vaccari, C., & Chadwick, A. (2020a). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media and Society*, *6*(1), 1–13. https://doi.org/10.1177/2056305120903408
19. Vaccari, C., & Chadwick, A. (2020b). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media and Society*, *6*(1), 2056305120903408. https://doi.org/10.1177/2056305120903408
20. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559
21. World Economic Forum. (2025). The Global Risks Report 2025 (20th ed.). In *Https://Www.Weforum.Org/Publications/Global-Risks-Report-2025/*