

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Perkembangan *generative artificial intelligence (AI)* telah mengubah cara konten digital diproduksi, disebar, dan dikonsumsi oleh masyarakat. Teknologi ini memungkinkan pembuatan teks, gambar, audio, dan video sintetis dengan tingkat realisme yang semakin tinggi. Akibatnya, batas antara konten autentik dan konten hasil manipulasi digital menjadi semakin kabur. Dalam konteks platform digital dan media sosial, kondisi ini menjadi persoalan penting karena konten berbasis AI dapat diproduksi secara cepat, disebar secara luas, dan berpotensi dimanfaatkan untuk manipulasi informasi, penipuan, maupun penyebaran misinformasi (Wessel, 2025).

Perkembangan tersebut juga sejalan dengan meningkatnya penggunaan AI secara global. Laporan McKinsey & Company (2025) menunjukkan bahwa *generative AI* semakin banyak digunakan dalam berbagai fungsi organisasi dan proses bisnis. Di satu sisi, teknologi ini mampu meningkatkan produktivitas dan mempercepat produksi konten. Namun di sisi lain, meningkatnya kemampuan AI dalam menghasilkan konten multimodal juga menimbulkan tantangan baru terkait akurasi, keamanan data, kepercayaan terhadap informasi, serta kemampuan manusia dalam memverifikasi konten yang dihasilkan oleh AI.

Dalam konteks Indonesia, persoalan ini menjadi semakin relevan karena masyarakat memiliki tingkat eksposur digital yang tinggi. APJII (2024) mencatat bahwa jumlah pengguna internet Indonesia telah mencapai lebih dari 221 juta orang, sementara DataReportal (2025) mencatat lebih dari 143 juta pengguna aktif media sosial di Indonesia. Data ini menunjukkan bahwa masyarakat Indonesia berada dalam ekosistem digital yang sangat aktif dan berpotensi tinggi terpapar berbagai bentuk konten berbasis AI. Selain itu, laporan Microsoft dan LinkedIn (2024) serta pemberitaan Kompas.id (2025) juga menunjukkan bahwa

penggunaan AI di Indonesia semakin meningkat, baik dalam konteks pekerjaan, pembelajaran, maupun aktivitas sehari-hari. Dengan demikian, masyarakat tidak hanya semakin sering berinteraksi dengan konten digital, tetapi juga semakin dekat dengan teknologi AI sebagai bagian dari aktivitas digitalnya.

AI Experience menjadi faktor penting untuk dikaji karena pengalaman individu dalam menggunakan AI dapat membentuk pemahaman awal mengenai karakteristik, manfaat, keterbatasan, dan pola output yang dihasilkan oleh teknologi tersebut (Long & Magerko, 2020; Ng et al., 2021). Dalam kehidupan sehari-hari, masyarakat tidak hanya terpapar konten berbasis AI, tetapi juga semakin sering menggunakan berbagai aplikasi *AI generatif*, seperti *chatbot*, *image generator*, *text generator*, dan fitur AI yang terintegrasi dalam *platform digital* (Dwivedi et al., 2023). Pengalaman tersebut dapat membuat individu lebih *familiar* dengan cara kerja AI, jenis respons yang dihasilkan, serta kemungkinan adanya kesalahan, bias, atau ketidakwajaran pada *output* AI (Bender et al., 2021).

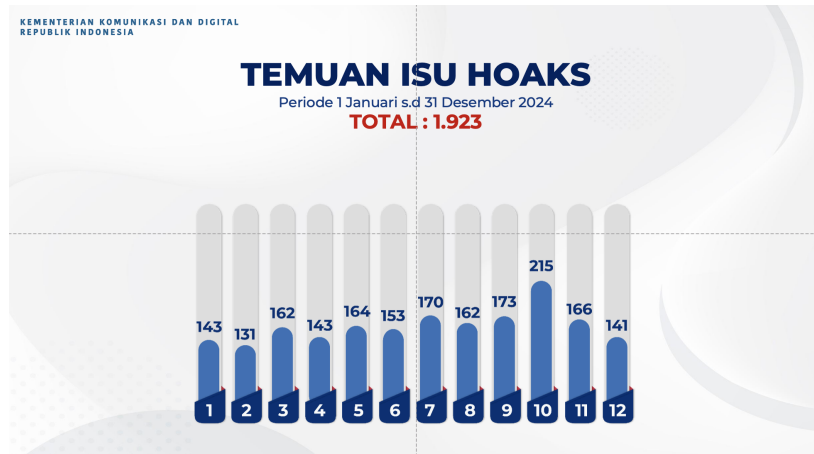
Namun, pengalaman menggunakan AI tidak selalu berarti individu memiliki kemampuan yang memadai untuk mengidentifikasi konten AI secara akurat. Seseorang dapat sering menggunakan AI untuk mencari informasi, menulis teks, membuat ide, atau membantu pekerjaan, tetapi belum tentu memiliki kemampuan evaluatif untuk membedakan konten autentik dan konten sintetis (Guess et al., 2020; Vaccari & Chadwick, 2020). Oleh karena itu, *AI Experience* perlu dikaji secara khusus dalam penelitian ini untuk mengetahui apakah pengalaman penggunaan dan familiaritas terhadap AI dapat berkontribusi terhadap pembentukan *AI Literacy* dan Kemampuan Identifikasi Konten AI.

Penelitian terdahulu juga menunjukkan bahwa pengalaman penggunaan AI berkaitan dengan persepsi, kepercayaan, dan penerimaan individu terhadap teknologi AI (Shin, 2021; Choung et al., 2022). Namun, pengalaman tersebut belum selalu dikaitkan secara langsung dengan kemampuan individu dalam mengevaluasi dan mengidentifikasi konten visual berbasis AI (Groh et al., 2022). Dengan demikian, memasukkan *AI Experience* dalam model penelitian ini menjadi relevan karena dapat menjelaskan apakah pengalaman praktis

menggunakan AI mampu berkembang menjadi literasi AI yang lebih kritis, serta apakah pengalaman tersebut dapat membantu individu mengenali konten yang dihasilkan atau dimanipulasi oleh AI.

Namun, tingginya paparan terhadap konten digital dan teknologi AI tidak secara otomatis menunjukkan bahwa masyarakat memiliki kemampuan yang memadai untuk mengenali konten yang dihasilkan atau dimanipulasi oleh AI. Di sinilah muncul problem utama penelitian ini, yaitu adanya kesenjangan antara tingginya eksposur masyarakat terhadap AI dan media sosial dengan kemampuan individu dalam mengidentifikasi konten AI. Semakin banyak masyarakat terpapar konten digital, semakin besar pula risiko mereka menghadapi konten sintesis yang tampak autentik, terutama apabila kemampuan evaluatif dan literasi digital tidak berkembang sejalan dengan peningkatan paparan tersebut.

Risiko tersebut mulai terlihat dalam berbagai kasus di Indonesia. Antara News (2025) melaporkan adanya peringatan kepada masyarakat terkait penipuan berbasis AI, termasuk penggunaan *deepfake*. Kasus ini menunjukkan bahwa teknologi AI mampu menciptakan konten visual yang tampak meyakinkan, tetapi dapat memuat informasi palsu dan berpotensi menyesatkan publik. Dampak penyebaran konten berbasis AI juga tidak hanya terjadi pada ranah sosial dan politik, tetapi juga pada ekosistem ekonomi digital. Otoritas Jasa Keuangan melalui Satgas PASTI/IASC melaporkan bahwa sejak 22 November 2024 sampai 11 November 2025, IASC telah menerima 343.402 laporan penipuan dengan memanfaatkan teknologi berbasis AI. Total rekening terkait penipuan yang dilaporkan mencapai 563.558 rekening, dengan 106.222 rekening telah diblokir. Total kerugian dana yang dilaporkan korban mencapai Rp7,8 triliun, sedangkan dana yang berhasil diblokir sebesar Rp386,5 miliar (Otoritas Jasa Keuangan, 2025).



Gambar 1.1 Temuan Isu Hoaks
Sumber: Komdigi(2025)

Selain itu, seperti yang dipaparkan pada Gambar 1.1. Kementerian Komunikasi dan Digital mencatat 1.923 temuan hoaks sepanjang tahun 2024, yang sebagian di antaranya berkaitan dengan konten berbasis AI. Data ini menunjukkan bahwa risiko manipulasi digital berbasis AI telah menjadi tantangan serius bagi ekosistem informasi di Indonesia (Kementerian Komunikasi dan Digital Republik Indonesia, 2025). Dalam konteks ini, kemampuan masyarakat untuk mengidentifikasi konten AI menjadi penting karena berkaitan langsung dengan kemampuan menjaga kepercayaan terhadap informasi digital, mencegah penyebaran misinformasi, dan mengurangi risiko penipuan berbasis teknologi.

Ironisnya, peningkatan risiko konten manipulatif berbasis AI tidak selalu diikuti oleh peningkatan literasi digital masyarakat. Indeks literasi digital Indonesia mengalami penurunan dari 58,25 menjadi 49,28 (IMDI, 2025). Penurunan ini menunjukkan bahwa kemampuan evaluatif masyarakat dalam menghadapi konten digital berisiko masih menjadi tantangan. Dengan demikian, masalah utama dalam konteks ini bukan hanya terletak pada semakin canggihnya teknologi AI, tetapi juga pada kesiapan manusia dalam mengevaluasi, memverifikasi, dan mengidentifikasi konten yang beredar di ruang digital.

Berbagai penelitian global juga menunjukkan bahwa manusia masih memiliki keterbatasan dalam membedakan konten AI dan konten autentik. Diel et

al. (2024) menemukan bahwa akurasi manusia dalam mendeteksi *deepfake* hanya sekitar 55,54%, yang mendekati peluang tebakan acak. Studi lain oleh Cooke et al. (2025) dan Fraser et al. (2025) menunjukkan bahwa kemampuan manusia dalam membedakan teks yang dihasilkan AI dan teks manusia sering kali mendekati peluang 50%. Cunningham (2025) juga menemukan bahwa masyarakat umum mengalami kesulitan dalam membedakan karya visual AI dan karya manusia. Temuan-temuan tersebut memperkuat asumsi bahwa kemampuan identifikasi konten AI masih relatif terbatas, bahkan ketika individu telah terbiasa dengan teknologi digital.

Dalam konteks konten visual, tantangan identifikasi menjadi lebih penting karena gambar dan video memiliki tingkat persuasi dan kepercayaan yang tinggi dibandingkan teks. Vaccari dan Chadwick (2020) menunjukkan bahwa *video deepfake* dapat mempengaruhi persepsi publik secara signifikan, sedangkan Nightingale dan Farid (2022) menemukan bahwa mayoritas individu mengalami kesulitan membedakan wajah sintesis dan wajah asli. Kemajuan teknologi seperti *Generative Adversarial Networks (GANs)* juga semakin memperkuat realisme konten visual, sehingga menurunkan akurasi deteksi manusia (Karras et al., 2019; Tolosana et al., 2020). Oleh karena itu, penelitian ini memfokuskan kemampuan identifikasi konten AI pada konten visual berupa gambar dan video.

Kemampuan identifikasi konten AI tidak hanya berkaitan dengan aspek teknis, tetapi juga dipengaruhi oleh faktor sosial, pengalaman, dan literasi. Dalam lingkungan digital, individu dapat belajar melalui paparan informasi, pengalaman menggunakan teknologi, serta interaksi sosial dengan orang lain. Hal ini sejalan dengan *Social Learning Theory* yang menjelaskan bahwa proses pembelajaran individu dapat terbentuk melalui observasi, interaksi sosial, dan pengalaman langsung (Bandura, 1977). Dalam konteks media sosial, interaksi dengan konten, diskusi dengan pengguna lain, serta pengalaman menggunakan AI dapat membentuk cara individu memahami dan mengevaluasi informasi digital (Guess et al., 2020; Pennycook & Rand, 2021).

Salah satu faktor yang relevan dalam konteks tersebut adalah *AI Social Interaction*. Interaksi sosial terkait AI dapat terjadi melalui diskusi, berbagi informasi, bertanya, atau meminta pendapat mengenai penggunaan dan konten berbasis AI. Melalui interaksi tersebut, individu dapat memperoleh paparan tambahan mengenai ciri-ciri konten AI, risiko manipulasi, dan pengalaman orang lain dalam menghadapi konten digital. Namun, interaksi sosial tidak selalu menghasilkan pemahaman konseptual yang mendalam. Oleh karena itu, perlu diuji apakah interaksi sosial terkait AI benar-benar berkontribusi terhadap kemampuan individu dalam mengidentifikasi konten AI.

Selain interaksi sosial, *AI Experience* juga menjadi faktor penting. Pengalaman menggunakan AI dapat meningkatkan familiaritas individu terhadap cara kerja teknologi AI, jenis *output* yang dihasilkan, serta keterbatasan teknologi tersebut. Individu yang lebih sering menggunakan AI berpotensi memiliki pemahaman yang lebih baik mengenai pola, karakteristik, dan kemungkinan kesalahan pada *output* AI. Namun, pengalaman penggunaan saja belum tentu cukup untuk menghasilkan kemampuan identifikasi yang akurat apabila tidak disertai pemahaman kritis mengenai cara mengevaluasi konten AI.

Dalam hal ini, *AI Literacy* menjadi variabel penting karena berhubungan dengan kemampuan individu untuk memahami, mengevaluasi, dan menggunakan AI secara kritis serta bertanggung jawab. Ng et al. (2021) menjelaskan bahwa literasi AI mencakup kemampuan memahami, menggunakan, mengevaluasi, dan mempertimbangkan aspek etis dalam penggunaan AI. Long dan Magerko (2020) juga menekankan bahwa literasi AI bukan hanya berkaitan dengan pemahaman teknis, tetapi juga kemampuan individu untuk mengevaluasi teknologi AI secara kritis. Dalam penelitian ini, *AI Literacy* diposisikan sebagai mekanisme kognitif yang dapat menjembatani pengalaman dan interaksi individu dengan kemampuan identifikasi konten AI.

Fokus wilayah penelitian ini diarahkan pada masyarakat Jabodetabek, yaitu Jakarta, Bogor, Depok, Tangerang, dan Bekasi. Pemilihan wilayah ini didasarkan pada beberapa pertimbangan. Pertama, mengacu pada publikasi Badan Pusat

Statistik yang digunakan dalam penelitian ini, Jabodetabek merupakan kawasan metropolitan terbesar di Indonesia dengan jumlah penduduk lebih dari 30 juta jiwa serta tingkat urbanisasi dan kepadatan yang tinggi (BPS, 2024). Kedua, berdasarkan laporan APJII yang digunakan dalam penelitian ini, wilayah DKI Jakarta dan daerah penyangganya memiliki tingkat penetrasi internet yang relatif tinggi dibandingkan rata-rata nasional (APJII, 2024). Ketiga, mengacu pada laporan Bank Indonesia yang digunakan dalam penelitian ini, wilayah Jabodetabek memiliki peran penting dalam aktivitas ekonomi digital, *e-commerce*, dan *fintech* nasional. Kondisi tersebut menunjukkan bahwa masyarakat Jabodetabek memiliki intensitas interaksi digital yang tinggi serta potensi paparan yang besar terhadap konten berbasis AI (Bank Indonesia, 2024).

Sebagai wilayah dengan akses teknologi dan aktivitas digital yang tinggi, Jabodetabek menyediakan konteks yang relevan untuk menguji apakah interaksi sosial, pengalaman penggunaan AI, dan literasi AI benar-benar berkontribusi terhadap kemampuan identifikasi konten AI. Jika pada wilayah dengan tingkat akses digital tinggi saja kemampuan identifikasi konten AI masih menjadi tantangan, maka persoalan ini dapat memiliki implikasi yang lebih luas bagi masyarakat Indonesia secara keseluruhan.

Berdasarkan uraian tersebut, penelitian ini diarahkan untuk menguji pengaruh *AI Social Interaction*, *AI Experience*, dan *AI Literacy* terhadap Kemampuan Identifikasi Konten AI pada pengguna media sosial di Jabodetabek. Penelitian ini juga menempatkan *AI Literacy* sebagai variabel mediasi untuk menjelaskan apakah pengalaman dan interaksi sosial terkait AI dapat membentuk kemampuan identifikasi melalui proses literasi yang lebih kritis dan evaluatif. Berbeda dari penelitian yang hanya mengandalkan persepsi responden, penelitian ini menggunakan pendekatan *performance-based measurement* untuk mengukur kemampuan identifikasi konten AI secara lebih objektif melalui tes performa pada konten visual berupa gambar dan video.

Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi empiris dalam memahami faktor-faktor yang mempengaruhi kemampuan

masyarakat dalam mengidentifikasi konten berbasis AI. Selain itu, penelitian ini juga diharapkan dapat memperkuat pengembangan literasi AI yang tidak hanya berfokus pada penggunaan teknologi, tetapi juga pada kemampuan mengevaluasi, memverifikasi, dan mengenali risiko manipulasi konten digital di era generative AI.

1.2 Rumusan Masalah

Banyak penelitian telah membuktikan bahwa secara umum individu masih mengalami kesulitan dalam membedakan konten yang dihasilkan oleh AI dan konten autentik. Studi oleh Nightingale dan Farid (2022) menunjukkan bahwa manusia sering kali tidak mampu membedakan wajah sintetis dengan wajah asli, bahkan pada tingkat akurasi yang mendekati tebakan acak. Diel et al. (2024) dalam studinya juga mengemukakan temuan serupa yang menyatakan bahwa akurasi manusia dalam mendeteksi konten *deepfake* relatif rendah, serta oleh Vaccari dan Chadwick (2020) yang menunjukkan bahwa konten berbasis AI, khususnya video, memiliki tingkat persuasi yang tinggi sehingga sulit diidentifikasi sebagai manipulasi. Ketika ditelusuri lebih lanjut, berbagai argumen berpendapat bahwa peningkatan kemampuan identifikasi konten AI dapat dipengaruhi oleh faktor seperti literasi AI, pengalaman penggunaan teknologi, serta interaksi sosial dalam lingkungan digital (Guess et al., 2020; Pennycook & Rand, 2021; Ng et al., 2021). Namun sebagian besar hasil temuan dari penelitian tersebut masih berupa argumen, survey berbasis persepsi atau *self-reported measures*, sehingga bukti empiris yang secara langsung menguji hubungan antara faktor-faktor tersebut dan kemampuan identifikasi konten AI secara objektif masih relatif terbatas.

Karena itu, permasalahan utama yang ingin diangkat dalam penelitian ini adalah apakah *AI social interaction*, *AI experience*, dan *AI literacy* mampu meningkatkan kemampuan identifikasi konten AI?

1.3 Tujuan Penelitian

Mengacu pada pemaparan rumusan masalah diatas, penelitian ini bertujuan untuk melakukan analisis pengaruh *AI social interaction* dan *AI experience* terhadap kemampuan identifikasi konten berbasis kecerdasan buatan dengan *AI literacy* sebagai variabel mediasi pada masyarakat di wilayah Jabodetabek.

Untuk mencapai tujuan tersebut, penelitian ini mencoba menjawab beberapa pertanyaan sebagai berikut:

1. Bagaimana pengaruh *AI social interaction* dan *AI experience* terhadap *AI literacy*?
2. Bagaimana pengaruh *AI social interaction*, *AI experience*, dan *AI literacy* terhadap kemampuan identifikasi konten AI?
3. Apakah *AI literacy* memediasi pengaruh *AI social interaction* dan *AI experience* terhadap kemampuan identifikasi konten AI?

1.4 Manfaat Penelitian

Melalui penelitian ini, diharapkan mampu memberikan kontribusi yang signifikan terhadap pengembangan dalam aspek teori dan praktik di bidang literasi digital, perilaku evaluasi informasi, dan kemampuan identifikasi konten berbasis kecerdasan buatan (AI). Berikut ini adalah beberapa manfaat dari penelitian ini:

a. Manfaat Teoritis

Secara teoritis, penelitian ini berkontribusi pada literatur manajemen teknologi, khususnya terkait kemampuan individu dalam mengidentifikasi konten AI. Penelitian ini mengembangkan model konseptual yang mengintegrasikan *AI Social Interaction*, *AI Experience*, dan *AI Literacy*, serta memperkuat relevansi *Social Learning Theory* dalam menjelaskan

pembentukan literasi AI sebagai mekanisme kognitif yang mempengaruhi kemampuan evaluatif individu di era kecerdasan buatan.

b. Manfaat Praktis

Secara praktis, penelitian ini memberikan wawasan bagi berbagai pemangku kepentingan tentang pentingnya interaksi sosial, pengalaman penggunaan AI, dan literasi AI dalam menghadapi konten sintetis. Bagi masyarakat, penelitian ini menekankan bahwa penggunaan AI perlu diimbangi dengan kemampuan evaluatif yang kritis, karena pengalaman menggunakan AI saja belum tentu cukup untuk membedakan konten autentik dan berbasis AI tanpa literasi yang memadai.

Bagi pemerintah dan regulator, hasil penelitian ini dapat menjadi dasar dalam merancang program literasi AI yang tidak hanya mengenalkan teknologi, tetapi juga mendorong penggunaan AI secara kritis, aman, dan bertanggung jawab, termasuk kemampuan memahami cara kerja, keterbatasan, dan verifikasi konten AI.

Bagi industri digital dan *platform* teknologi, penelitian ini dapat menjadi acuan dalam mengembangkan fitur edukasi, verifikasi konten, serta strategi komunikasi yang membantu pengguna memahami dan mengidentifikasi konten berbasis AI secara lebih kritis.

Bagi institusi pendidikan dan organisasi, penelitian ini dapat digunakan untuk merancang pelatihan literasi AI berbasis pengalaman, yang tidak hanya bersifat teoretis tetapi juga melibatkan praktik langsung dalam mengevaluasi dan membandingkan *output* AI dengan konten autentik.