

Analisis Penyalahgunaan AI Dalam Memanipulasi Opini Publik di Media Sosial

Elsa Omega Napitupulu, Fawwaz Aozadel, Nadia Inggrida Prashayu,
Samuel Hotdin*, Winson

Fakultas Hukum, Universitas Tarumanagara Jakarta

*Penulis korespondensi: samuel.205250205@stu.untar.ac.id

Dikirim : 20 November 2025

Direvisi : 2 Desember 2025

Diterima : 3 Desember 2025

Abstrak: Akses yang semakin mudah dan terbuka terhadap teknologi kecerdasan buatan (Artificial Intelligence/AI) telah mendorong lahirnya inovasi baru dalam produksi konten digital. Namun, di sisi lain, perkembangan ini juga memunculkan tantangan serius terhadap keandalan dan objektivitas informasi. Penyalahgunaan AI di media sosial semakin meningkat, ditandai dengan munculnya konten manipulatif yang tampak sangat realistis dan sulit dikenali, seperti deepfake pada video dan audio serta rekayasa visual untuk menyebarkan disinformasi dan memicu konflik. Ancaman ini melampaui hoaks berbasis teks karena AI mampu menciptakan situasi krisis kepercayaan, di mana perbedaan antara fakta dan kebohongan menjadi semakin kabur di mata publik. Konten semacam ini umumnya dirancang untuk memengaruhi emosi dan membentuk persepsi, sehingga audiens cenderung menerima informasi tanpa melakukan verifikasi. Artikel ini mengkaji pola penyalahgunaan AI oleh pengguna serta implikasinya terhadap ekosistem media sosial. Hasil kajian menunjukkan bahwa fenomena tersebut berpotensi menurunkan kepercayaan publik secara luas, tidak hanya terhadap media, tetapi juga terhadap lembaga penegak hukum, pemerintah, dan pihak-pihak terkait lainnya. Dalam jangka panjang, kondisi ini dapat memperparah polarisasi sosial dan melemahkan kredibilitas publik. Oleh karena itu, diperlukan upaya kolaboratif melalui penegakan aturan penggunaan teknologi dan penguatan literasi digital agar masyarakat lebih kritis dalam menyikapi informasi.

Kata kunci: deepfake, kebebasan berekspresi, kecerdasan buatan, literasi digital, media sosial

Abstract: Increasingly easy and open access to artificial intelligence (AI) technology has fostered new innovations in digital content creation. However, this development also gives rise to serious challenges to the reliability and objectivity of information. The misuse of AI on social media is becoming more prevalent, as evidenced by the emergence of highly realistic and difficult-to-detect manipulative content, such as video and audio deepfakes and visual manipulation aimed at spreading disinformation and inciting division. This threat goes beyond text-based hoaxes, as AI has the potential to generate a crisis of trust in which the distinction between truth and falsehood becomes increasingly blurred in the public eye. Such content is typically designed to influence emotions and shape perceptions, leading audiences to accept information without verification. This article examines patterns of AI misuse by users and their implications for the social media ecosystem. The findings indicate that this phenomenon can significantly erode public trust, not only in the media but also in law enforcement institutions, government bodies, and other relevant stakeholders. In the long term, this condition may exacerbate social polarization and undermine public credibility. Therefore, collaborative efforts are required through the enforcement of regulations governing the use of technology and the strengthening of digital literacy to encourage more critical engagement with information.

Keywords: artificial intelligence, deepfake, digital literacy social media

1. Pendahuluan

Media sosial telah menjadi pusat aktivitas publik di era digital. Berbagai media ini menjadi tempat utama bagi masyarakat untuk berbagi informasi, menyuarakan pendapat, dan melakukan berbagai diskusi. Namun, kebebasan ini juga dapat mengancam dan memicu masalah serius, yaitu maraknya berita bohong atau hoaks yang terbukti bisa memecah belah masyarakat dan merusak kepercayaan publik. Selama ini, tantangan utama dalam media sosial adalah melawan penyebaran informasi salah yang dahulu hanya berbentuk teks (Azzahra dkk., 2025; Hamama, 2024).

Namun, kini situasinya dapat berubah total dengan munculnya kecerdasan buatan (AI) yang bisa dipakai oleh siapa saja dan dimana saja. Teknologi yang dulunya rumit, terbatas, dan mahal, sekarang tersedia dalam aplikasi sederhana yang bahkan memungkinkan orang awam dalam membuat konten palsu yang terlihat sangat nyata. Contoh paling umum yang sedang marak adalah teknologi *deepfake*, yang dapat memanipulasi video, dan suara seseorang, bahkan rekayasa gambar yang sangat sulit dibedakan dari foto aslinya. Kemudahan ini menjadi jalan bagi penyalahgunaan AI dalam skala yang sangat besar dan belum pernah ada sebelumnya.

Ancaman baru ini jauh lebih berbahaya, karena hoaks dapat dilawan dengan data dan fakta. Namun, sekarang yang diserang adalah rasa kepercayaan para pengguna media sosial dan kemampuan untuk percaya kepada apa yang dilihat dan didengar. Konten manipulatif dan palsu sengaja dirancang agar dengan mudah menyentuh emosi yang dapat membuat kesulitan berpikir rasional dan dengan mudah diadu domba. Akibatnya, batas antara mana yang asli dan mana yang direkayasa menjadi semakin sulit untuk diketahui oleh publik.

Kondisi dan situasi seperti ini tentu sangat membahayakan nilai demokrasi di negeri ini. Kepercayaan pada pihak penting seperti proses pemilu, media, tokoh publik, wakil rakyat, dan hukum dapat terguncang ketika publik tidak lagi bisa membedakan antara fakta dan tidak nyata. Reputasi seseorang dapat dirusak dengan mudah dan perpecahan dapat dipicu oleh konten rekayasa negatif yang dibuat dengan sengaja. Oleh karena itu, penting bagi masyarakat untuk memahami bagaimana penyalahgunaan AI ini dapat bekerja dan apa saja dampaknya bagi masyarakat, serta perilaku selektif seperti apa yang sebaiknya diterapkan oleh masyarakat agar dapat membedakan mana yang benar dan mana yang sudah direkayasa. Artikel ini disusun untuk menganalisis ancaman baru ini secara mendalam, sebagai langkah awal dalam membangun lingkungan sosial yang positif di tengah gempuran disinformasi era modern.

2. Metode

Metode penelitian yang digunakan menggunakan metode kualitatif deskriptif yang bertujuan untuk menganalisis penyalahgunaan teknologi AI dalam memanipulasi opini publik di media sosial. Analisis konten dan studi kasus digunakan untuk mengamati berbagai konten di media sosial yang diasumsikan telah dimanipulasi oleh sistem AI seperti *bot* dan *deepfake*. Metode pengumpulan data meliputi observasi digital dengan menelusuri tren, *hashtag*, atau akun bot, pola komentar yang dihasilkan oleh AI, serta sumber pustaka pendukung dan situs internet, baik berupa tulisan, gambar, maupun dokumentasi dari platform media sosial tersebut (Facebook, Instagram, Twitter).

3. Hasil dan Diskusi

3.1 Bentuk Penyalahgunaan AI di Media Sosial

Contoh pertama dan mungkin yang paling umum adalah dalam pembuatan konten. Teknologi *deepfake* mampu memalsukan wajah, suara, dan ekspresi dengan alat yang paling mungkin sekaligus paling banyak digunakan untuk menyebarkan hoaks dalam bentuk visual. Teknologi ini menciptakan video atau gambar menggunakan teknik pembelajaran mendalam menggunakan kerangka kerja *Generative Adversarial Networks* (GANs) yang tampak nyata padahal sebenarnya sepenuhnya palsu. Kebanyakan *deepfake* yang menargetkan opini publik biasanya digunakan untuk membuat video palsu tampak seperti orang penting, misalnya politisi, pejabat publik, dan selebritas yang sedang berbicara. Video-video ini dapat membangun atau menghancurkan reputasi, menciptakan sentimen massa, atau mengubah persepsi tentang suatu isu karena kebanyakan orang tidak dapat membedakan video *deepfake* dari yang asli. Ketika kebenaran akhirnya terungkap, dampak persepsi publik yang sudah terbentuk sering kali sulit diubah. Fenomena ini dikenal sebagai *illusory truth effect*. Lebih jauh lagi, setiap orang tidak bisa mengabaikan bagaimana *deepfake* yang seolah-olah nyata itu semakin menyebar. Algoritma media sosial dengan segala kecanggihannya seakan-akan mengangkat konten viral tanpa memeriksa keasliannya terlebih dahulu. Hal ini memperparah penyebaran disinformasi dan mempercepat pembentukan opini publik yang keliru sebagaimana disimpulkan dari hasil wawancara dengan narasumber (Gambar 1).

3.2 Respons Publik dan Efektivitas Manipulasi

Berdasarkan hasil pengamatan digital, analisis komentar di media sosial, dan data dari beberapa survei yang bisa diandalkan, menunjukkan bahwa mayoritas pengguna media sosial

pada umumnya masih kesulitan membedakan antara konten yang asli dan yang telah dipoles dan dimanipulasi oleh AI. AI kini berperan sebagai aktor utama dalam panggung informasi, membentuk persepsi dan opini publik. Hal ini menunjukkan bahwa kemampuan literasi digital masyarakat masih rendah, terutama dalam menghadapi konten yang dihasilkan oleh AI.

Kebanyakan pengguna media sosial tidak memiliki kesadaran atau keterampilan teknis untuk mengenali tanda-tanda konten yang telah dimanipulasi secara otomatis. Akibatnya, opini publik sangat mudah terbentuk berdasarkan informasi yang salah atau menyesatkan, yang tersebar secara masif dan cepat melalui media sosial. Selain itu, algoritma media sosial sendiri justru memperkuat konten yang bersifat kontroversial, emosional, atau viral yang seringkali merupakan hasil dari manipulasi AI. Hal ini memperbesar efek dimana pengguna hanya terpapar pada informasi yang mendukung keyakinannya dan menolak informasi yang berseberangan tanpa memeriksa keasliannya terlebih dahulu.

3.3 Tantangan Regulasi dan Upaya Mitigasi

Di Indonesia, aspek hukum terkait *deepfake* masih menghadapi tantangan serius. Regulasi seperti Undang-Undang Informasi dan Transaksi Elektronik (UU ITE) memang dapat dijadikan dasar hukum untuk menindak pelanggaran, namun belum mengatur secara spesifik mengenai konten manipulatif berbasis AI (Rangkuti dkk., 2025; Respati dkk., 2024). Kekosongan hukum ini menciptakan area abu-abu, terutama pada penggunaan *deepfake* untuk hiburan yang tidak merugikan pihak tertentu. Namun, ketika konten *deepfake* dipakai untuk merusak reputasi, menyebarkan kebencian politik, atau tujuan kriminal seperti pemerasan, maka permasalahannya jelas masuk ranah pidana (Fitri dkk., 2025). Oleh karena itu, regulasi yang lebih adaptif diperlukan agar hukum mampu menjawab tantangan perkembangan teknologi.

Selain pendekatan hukum, strategi mitigasi perlu dijalankan melalui tiga jalur. Pertama, pendekatan teknologi dengan mengembangkan sistem deteksi otomatis di platform digital. Kedua, pendekatan sosial berupa peningkatan literasi digital secara masif, agar masyarakat mampu mengenali konten manipulatif (Rahmawati, 2023). Ketiga, pendekatan etika yang menekankan pentingnya kesadaran moral dalam menggunakan teknologi, sehingga masyarakat tidak sekadar menghindari sanksi hukum, tetapi juga memahami dampak sosial dari penyalahgunaan *deepfake* (Cahyono dkk., 2023; Bijaksono dkk., 2022).

3.4. Implikasi untuk Masa Depan

Keberadaan *deepfake* memiliki implikasi jangka panjang yang signifikan. Dalam ranah politik, teknologi ini diperkirakan akan semakin digunakan untuk manipulasi elektoral dan propaganda, yang dapat menurunkan kualitas demokrasi (Vaccari & Chadwick, 2020). Dalam ranah sosial, *deepfake* berpotensi mempercepat munculnya fenomena *infopocalypse*, yakni kondisi ketika publik kehilangan kemampuan membedakan realitas dengan rekayasa (Zahro dkk., 2024). Dari sisi ekonomi, *deepfake* membuka peluang munculnya modus penipuan baru, seperti pencurian identitas, pemalsuan suara, hingga serangan reputasi terhadap perusahaan (Fallis, 2021). Hal ini dapat mengancam kepercayaan masyarakat terhadap transaksi digital maupun dunia bisnis secara umum.

Untuk menghadapi tantangan tersebut, masa depan literasi digital dan regulasi hukum harus diarahkan pada pembangunan ekosistem digital yang lebih sehat. Pemerintah, akademisi, dan sektor swasta perlu berkolaborasi dalam menciptakan kebijakan adaptif, mengembangkan teknologi deteksi, serta memperluas program edukasi publik yang berkelanjutan. Dengan cara ini, masyarakat tidak hanya menjadi pengguna pasif informasi, tetapi juga memiliki kapasitas kritis untuk menghadapi tantangan era digital.



Gambar 1. Wawancara dengan Rektor Universitas Buddhi Dharma terkait risiko AI dalam pembentukan opini masyarakat

4. Kesimpulan

Hasil penyalahgunaan teknologi AI di media sosial, seperti membuat dan menyebarkan konten palsu, dapat memanipulasi opini publik dan menyebarkan disinformasi yang berbahaya. Penyalahgunaan informasi dapat berdampak negatif pada kepercayaan masyarakat,

meningkatkan kemungkinan konflik sosial, dan mengancam stabilitas politik dan keamanan. Oleh karena itu, pemerintah harus memastikan bahwa penggunaan AI tetap berada di jalur yang menguntungkan dan bertanggung jawab melalui edukasi, peraturan, dan pengawasan. Suara Pemerintah diharapkan membuat regulasi yang jelas dan tegas tentang penggunaan AI, serta program pendidikan berkelanjutan untuk mencegah penyalahgunaan, agar masyarakat mampu membedakan antara informasi asli dan palsu dan juga untuk menekan peredaran informasi menyesatkan, platform media sosial juga perlu meningkatkan sistem keamanan dan deteksi konten palsu berbasis AI. Sementara itu, akademisi dan peneliti dapat membantu dengan mengembangkan inovasi yang menekankan keamanan dan etika penggunaan AI.

Hasil penelitian menunjukkan bahwa penyalahgunaan kecerdasan buatan (AI) di media sosial adalah fenomena yang semakin kompleks yang memiliki dampak yang signifikan terhadap kehidupan masyarakat. Penelitian menunjukkan bahwa karena sulit membedakan teknologi seperti *deepfake*, *bot* otomatis, dan manipulasi visual dapat mempengaruhi persepsi publik secara signifikan. Tingkat literasi digital masyarakat semakin rendah, sehingga sebagian besar pengguna media sosial tidak dapat membedakan informasi penting. Hal ini berdampak negatif pada kepercayaan masyarakat terhadap media digital, munculnya polarisasi sosial, dan meningkatnya kerentanan terhadap penipuan, ujaran kebencian, dan kemungkinan manipulasi politik.

Dari segi hukum, undang-undang saat ini, seperti Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), pada dasarnya dapat digunakan sebagai dasar untuk penindakan, tetapi undang-undang ini tidak secara khusus mengatur konten manipulatif berbasis kecerdasan buatan, yang menyebabkan celah hukum yang memungkinkan penyalahgunaan teknologi berlanjut tanpa pengawasan yang memadai. Dalam jangka panjang, kondisi ini mengancam stabilitas sosial dan demokrasi serta kepercayaan masyarakat terhadap pemerintah dan sistem hukum dalam menghadapi disrupsi teknologi.

Penelitian ini memberikan beberapa rekomendasi yang harus dipertimbangkan. Pertama, pemerintah harus segera membuat undang-undang khusus yang mengatur penggunaan dan penyalahgunaan AI. Ini akan memberikan kepastian hukum yang berubah seiring dengan kemajuan teknologi. Kedua, platform media sosial diharapkan untuk mengembangkan sistem deteksi otomatis berbasis AI yang dapat dengan cepat menemukan dan membatasi konten manipulatif. Ketiga, pendidikan formal dan non-formal harus dilakukan untuk meningkatkan literasi digital masyarakat. Ini akan membantu siswa memperoleh keterampilan kritis untuk

menilai dan memverifikasi informasi. Keempat, media harus meningkatkan kesadaran etika sehingga orang tidak hanya mematuhi hukum tetapi juga bertanggung jawab untuk menggunakan teknologi digital. Terakhir, membangun ekosistem digital yang sehat, aman, dan terpercaya membutuhkan kerja sama lintas sektor antara pemerintah, akademisi, pelaku industri, dan masyarakat sipil. Dengan aturan yang jelas, dukungan teknologi, peningkatan kapasitas masyarakat, dan penguatan nilai etika, efek negatif penyalahgunaan AI di media sosial dapat diminimalkan. Akibatnya, kemajuan teknologi dapat diarahkan untuk membangun ruang digital yang inklusif dan membantu kemajuan negara.

Ucapan Terima Kasih

Terima kasih kepada instansi yang telah memberikan data dan informasi yang diperlukan dalam penelitian ini. Penelitian ini dapat terselesaikan dengan baik berkat dukungan serta kerja sama dari semua pihak yang terlibat.

Daftar Referensi

- Azzahra, K., Triyanti, A., & Desi, Y. (2025). Media Sosial dan Tantangan di Era Digital. *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, 4(1), 1625-1629.
- Bijaksono, D., Windiyanti, N., Febrianti, S. R., & Permada, D. N. R. (2022). Etika dan Regulasi Teknologi di Era Kecerdasan Buatan. *Prosiding Seminar Nasional Manajemen*, 1(2), 146–151.
- Cahyono, N.F., 'Uyun, K., & Mukaromah, S. (2023). Etika Penggunaan Kecerdasan Buatan Pada Teknologi Informasi. *Prosiding Seminar Nasional Teknologi Dan Sistem Informasi*, 3(1), 482–491. <https://doi.org/10.33005/sitasi.v3i1.334>
- Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, 34(4), 623–643. <https://doi.org/10.1007/s13347-020-00419-2>
- Fitri, D., Hidayah, A.N., Putri, A., Tanjung, N.H., Ramadhani, S.I., Akila, D., Manurung, R.A., Mufidah, N., Akbar, S., & Zikri, M. (2025). Deepfake Dan Krisis Kepercayaan: Analisis Hukum Terhadap Penyebaran Konten Palsu di Media Sosial. *JIIIC: Jurnal Intelek Insan Cendikia*, 2(6), 11556-11568.
- Hamama, S. (2024). Etika Komunikasi dalam Media Sosial: Tantangan dan Solusinya. *Selasar KPI : Referensi Media Komunikasi dan Dakwah*, 4(2), 182-197.
- Rahmawati, A.Z. (2023). Peningkatan Literasi Digital untuk Masyarakat Berbasis Era Teknologi Informasi. *MUJAHADA: Jurnal Pengabdian Masyarakat*, 1(1), 12-20.

- Rangkuti, P.R., Hasibuan, R.F.A., Witama, V., Nasution, M.A., Kholizah, S., Andini, N. & Salmiah, S. (2025). Analisis Yuridis terhadap Penggunaan Deepfake dalam Pelanggaran UU ITE : Studi Kasus di Media Sosial Indonesia. *Depositi: Jurnal Publikasi Ilmu Hukum*, 3(2), 111-120.
- Respati, A.A., Setyarini, A.D., Parlagutan, D., Rafli, M., Mahendra, R.S., & Nugroho, A.A. (2024). Analisis Hukum Terhadap Pencegahan Kasus Deepfake Serta Perlindungan Hukum Terhadap Korban. *Media Hukum Indonesia (MHI)*, 2(2), 586-592.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- Zahro, A., Fadhilah, R.R., Hermawati, S.Z., Imaduddin, G.N. & Santoso, A.A. (2024). Dampak Penyalahgunaan Deepfake dalam Memanipulasi Visual: Menguak Potensi Infopocalypse di Era Post Truth Terhadap Asumsi Masyarakat pada Media Massa. *Jurnal Kawistara: The Journal of Social Sciences and Humanities*, 14(3), 401-415.