

Deepfake dan Disinformasi: Krisis Literasi Media dan Tantangan Etika Komunikasi di Era Kecerdasan Buatan

Fathimah Nadia Qurrota A'yun*

Fitri Ariana Putri

Nabila Zahrani Putri

UIN Walisongo Semarang, UIN Walisongo Semarang, Universitas Diponegoro

ABSTRAK

Kemajuan teknologi kecerdasan buatan (AI) telah membawa perubahan besar dalam ekosistem komunikasi publik, khususnya dalam proses produksi dan distribusi informasi. Penelitian ini bertujuan untuk menganalisis bagaimana konten deepfake dan disinformasi berbasis AI membentuk persepsi dan kepercayaan publik, serta bagaimana nilai-nilai komunikasi Islam dapat dijadikan dasar etika dalam menghadapi fenomena tersebut. Hasilnya bahwa fenomena deepfake dan disinformasi AI memperkuat krisis literasi media dan kepercayaan publik melalui tiga mekanisme utama: *Pertama*, ilusi objektivitas, di mana rekayasa digital tampak sangat meyakinkan secara visual dan emosional; *Kedua*, rendahnya literasi digital, yang membuat publik tidak mampu membedakan fakta dari manipulasi; *Ketiga*, penyalahgunaan teknologi, yang digunakan untuk propaganda dan pembingkaiian kepentingan tertentu. Deepfake dan disinformasi bukan hanya tantangan teknologi, tetapi juga krisis etika komunikasi yang menuntut sinergi antara teori komunikasi modern dan nilai-nilai Islam. Dengan menjadikan prinsip tabayyun sebagai landasan verifikasi informasi, komunikasi publik di era AI dapat diarahkan menuju praktik yang lebih etis, manusiawi, dan berkeadilan.

Kata Kunci : deepfake; disinformasi; literasi media; framing; etika komunikasi Islam

Pendahuluan

Perkembangan pesat kecerdasan buatan (*Artificial Intelligence/AI*), khususnya generatif AI, telah membawa transformasi besar dalam dunia produksi dan distribusi informasi. Salah satu inovasi yang paling kontroversial adalah teknologi *deepfake* yaitu rekayasa digital berbasis *deep learning* yang mampu menghasilkan konten audio, video, maupun gambar yang sangat realistis meskipun sepenuhnya palsu. Berkat kemajuan model generatif, teknologi ini kini semakin mudah diakses oleh publik, bahkan dengan sumber daya yang relatif terbatas. Hal ini memungkinkan siapa pun untuk menciptakan rekaman yang meniru wajah, suara, ekspresi, atau perilaku individu lain dengan tingkat keaslian yang dapat mengecoh mata dan telinga manusia.

Teknologi deepfake tidak hanya membuka peluang positif di bidang kreatif, hiburan, atau simulasi pendidikan, tetapi juga menghadirkan risiko serius ketika digunakan sebagai alat penyebaran disinformasi. Menurut Chesney (2019), Deepfake menciptakan ancaman berupa "*reputational harm at scale*," yaitu kerusakan reputasi yang dapat terjadi secara cepat, luas, dan sulit dipulihkan. Kemampuan deepfake untuk

* Corresponding author: Fathimah Nadia Qurrota A'yun

meniru tindakan atau pernyataan seseorang secara meyakinkan mengancam ketahanan ruang publik, integritas komunikasi, dan kredibilitas institusi demokratis.

Sejumlah penelitian menekankan bahaya yang ditimbulkan deepfake terhadap kepercayaan publik dan ekosistem media. Boediman (2025) menjelaskan bahwa deepfake “mengaburkan batas antara kebenaran dan rekayasa,” sehingga mengikis kredibilitas media dan memunculkan ketidakpastian epistemik. Pada tingkat global, para peneliti memperingatkan fenomena *liar’s dividend* yaitu sebuah kondisi ketika pelaku manipulasi dengan mudah menyangkal rekaman asli sebagai deepfake untuk menghindari akuntabilitas (Fallis, 2021). Dengan demikian, keberadaan deepfake tidak hanya menyebabkan munculnya informasi palsu, tetapi juga melemahkan validitas bukti visual yang selama ini dianggap paling kuat dalam komunikasi.

Dalam konteks literasi media, fenomena ini memperlihatkan tantangan yang semakin mendesak. Penelitian empiris menemukan bahwa intervensi literasi media, seperti pelatihan dan pendidikan formal, dapat meningkatkan kemampuan publik dalam mendeteksi konten yang direkayasa serta mengurangi kecenderungan membagikan konten menyesatkan (Jones, 2023). Berbagai studi juga menunjukkan bahwa tingkat literasi digital di banyak negara, termasuk Indonesia, masih belum memadai. Meskipun akses teknologi dan konsumsi media digital cukup tinggi, pemahaman mendalam tentang cara kerja AI dan manipulasi audio-visual masih rendah. Di Indonesia, misalnya, penelitian menunjukkan bahwa kemampuan generasi X dalam mengenali video deepfake masih berada pada level yang memprihatinkan (Fahrudin, 2025).

Dari sudut pandang etika komunikasi, deepfake menimbulkan persoalan mendasar mengenai consent, identitas, privasi, dan akuntabilitas. Seseorang dapat “diciptakan ulang” dalam situasi tertentu tanpa sepengetahuan atau persetujuannya, menimbulkan implikasi serius bagi martabat dan keamanan individu. Selain itu, deepfake memiliki konsekuensi sosial-politik yang signifikan karena dapat memperkuat polarisasi, menyebarkan rumor politik, memanipulasi opini publik, dan merusak kepercayaan terhadap lembaga publik. Penelitian (Nurdin, 2025) menunjukkan bahwa penyebaran deepfake dapat mengganggu stabilitas sosial dan keamanan informasi nasional apabila tidak ditangani dengan regulasi dan teknologi deteksi yang memadai.

Dengan demikian, krisis deepfake tidak dapat dipahami semata-mata sebagai persoalan teknologi, tetapi juga sebagai persoalan literasi media, regulasi, dan etika komunikasi digital. Diperlukan strategi respons yang holistik: edukasi literasi media yang berorientasi pada pemahaman AI, pengembangan teknologi deteksi deepfake regulasi yang adaptif terhadap perkembangan teknologi, serta pembentukan norma etika baru bagi pembuat konten, platform digital, dan pembuat kebijakan.

Metode

Penelitian ini terfokus pada pada fenomena deepfake dan disinformasi di media sosial dengan melihat aspek krisis literasi media dan tantangan etika komunikasi di era Kecerdasan Buatan (AI). Selanjutnya penelitian ini melakukan pemetaan dalam pemilihan data terkait deepfake pada media sosial intagram dan tiktok. Adapun korban penyebaran deepfake saat ini, salah satunya pejabat pemerintahan. Mulai dari keplaa daerah, gubernur, menteri bahkan mantan presiden juga menjadi korban deepfake.

Penelitian ini menggunakan pendekatan kualitatif deskriptif yang bertujuan untuk menggali makna, pola, serta hubungan antar fenomena yang kompleks dan tidak dapat diukur secara kuantitatif (Fitri et al., 2025). Penelitian ini mendasarkan diri pada analisis wacana kritis atau Critical Discourse Analysis (CDA). Norman Fairclough (Ni & Sartini, 2017) dapat berkontribusi dalam analisis sosial dan budaya. Model analisis ini sebagai kerangka yang paling komprehensif dan relevan dalam mengkaji berita dan video deepfake yang beredar di media sosial. Analisis didasarkan pada Teori Framing (Entman, 1993) untuk mengidentifikasi bagaimana pesan dan makna dimanipulasi melalui struktur bahasa dan visual, Teori Konstruksi Sosial atas Realitas (Berger, 1966)

untuk memahami bagaimana realitas palsu terbentuk dan diterima publik, serta Etika Komunikasi Islam sebagai kerangka normatif yang menekankan prinsip amanah, sidq, dan tabayyun dalam proses penyebaran informasi (Anggraini et al., 2023). Selanjutnya, data primer merupakan data yang diperoleh melalui kajian terhadap platform media social Instagram dan tiktok. Sedangkan data sekunder diperoleh dari berbagai websites yang memiliki relevansi dengan tema penelitian, seperti <https://www.cnnindonesia.com/>, Kompas.id, KBR.id, dan sebagainya. Selain itu, data sekunder digali melalui buku, artikel, jurnal sebagai bahan yang mendukung dalam penelitian ini.

Hasil

Hasil sintesis berbagai penelitian menunjukkan bahwa teknologi deepfake telah menjadi faktor signifikan dalam meningkatnya disinformasi serta memburuknya kualitas literasi media di era kecerdasan buatan. Deepfake mampu merekayasa wajah, suara, dan konteks secara sangat realistis sehingga publik semakin sulit membedakan konten asli dan palsu (Chesney, 2019). Kondisi ini memperkuat fenomena post-truth, di mana emosi dan persepsi publik lebih dominan daripada fakta objektif (Widjaja, 2025).

Studi-studi internasional, termasuk (Vaccari, 2020) menunjukkan bahwa paparan deepfake politik meningkatkan kebingungan informasi dan menurunkan kepercayaan terhadap media dan institusi, bahkan setelah video tersebut dibuktikan palsu. Rendahnya literasi digital memperparah kerentanan masyarakat, sebagaimana ditemukan oleh (Fahrudin, 2025), yang mencatat bahwa publik cenderung menyebarkan konten emosional tanpa proses verifikasi.

Selain mengganggu pemahaman publik, deepfake juga memicu tantangan etika komunikasi dan merusak kredibilitas media. Deepfake berpotensi menciptakan infopocalypse berupa banjir informasi palsu yang mengacaukan acuan realitas masyarakat (Zahro, 2024). Hal ini diperkuat oleh temuan (Irsyad, 2025) yang menyebut bahwa jurnalisme menghadapi hambatan besar dalam proses verifikasi ketika konten manipulatif semakin sulit dikenali.

Secara keseluruhan, penelitian ini menegaskan bahwa deepfake tidak hanya persoalan teknologi, tetapi merupakan ancaman serius terhadap kepercayaan publik, integritas komunikasi, dan stabilitas demokrasi. Upaya edukasi literasi media, regulasi etika AI, dan peningkatan kapasitas verifikasi digital menjadi kebutuhan mendesak untuk mengurangi dampak destruktifnya.

Kasus Sri Mulyani di Indonesia

Ancaman deepfake dalam ekosistem komunikasi digital Indonesia tampak jelas melalui penyebaran video manipulatif yang menarget Menteri Keuangan Sri Mulyani. Dalam video tersebut, ia tampak seolah-olah mengatakan bahwa “guru adalah beban negara.” Padahal, pernyataan aslinya dalam forum KSTI 2025 hanya menjelaskan tantangan APBN dalam menjaga keberlanjutan pembiayaan pendidikan. Video yang beredar merupakan kombinasi pemotongan konteks, rekayasa audio-visual, dan penyesuaian sinkronisasi bibir, sehingga tampilannya tampak autentik di mata masyarakat awam.



Sumber : <https://www.tempo.co/ekonomi/polemik-deepfake-guru-beban-negara-usai-pernyataan-sri-mulyani-2061974>

Penyebaran video tersebut menimbulkan efek berantai yang signifikan: kemarahan publik, polarisasi opini, dan lahirnya narasi negatif terhadap pemerintah. Dalam konteks ini, temuan (Widjaja, 2025) menunjukkan relevansinya bahwa deepfake menciptakan kondisi post-truth, yaitu situasi ketika persepsi dan emosi publik lebih kuat daripada fakta empiris. Publik dengan cepat mempercayai video tersebut tanpa melakukan verifikasi, mempertegas rendahnya literasi media digital di Indonesia.

Masyarakat Indonesia cenderung membagikan informasi yang memicu emosi moral tanpa pemeriksaan sumber (Fahrudin, 2025). Pola ini dapat memperkuat penyebaran disinformasi berbasis AI dan menunjukkan betapa mudahnya manipulasi digital menembus ekosistem media sosial.

Deepfake Presiden Ukraina Volodymyr Zelensky (Kasus Internasional yang Terverifikasi)

Kasus deepfake paling terkenal di tingkat global adalah video Presiden Ukraina Volodymyr Zelensky yang menampilkan dirinya sedang menyampaikan pernyataan bahwa militer Ukraina "harus menyerah." Video ini muncul pada Maret 2022, saat awal invasi Rusia ke Ukraina. Para peretas mengunggah video tersebut ke televisi Ukraina dan berbagai kanal media sosial, sehingga menimbulkan kebingungan dan kekhawatiran publik sebelum pemerintah berhasil mengklarifikasi bahwa video itu palsu.



Sumber : <https://news.virginia.edu/content/qa-zelenskyy-surrender-hoax-feared-future-deepfakes-here>

Analisis forensik digital menunjukkan bahwa video tersebut merupakan hasil face-swap deepfake, ditambah manipulasi text-to-speech untuk meniru suara Zelensky. Studi (Diakopoulos, 2022) menyebut kasus ini sebagai bukti nyata penggunaan deepfake dalam perang informasi geopolitik, terutama untuk menurunkan semangat juang pasukan dan mempengaruhi persepsi publik internasional.

Lebih jauh, penelitian (Westerlund, 2019) telah mengantisipasi skenario bahwa deepfake bukan hanya ancaman pada individu, tetapi dapat digunakan sebagai instrumen propaganda, sabotase diplomatik, dan manipulasi opini publik pada skala nasional maupun global. Kasus Zelensky membuktikan prediksi tersebut secara nyata.

Kasus Deepfake Barack Obama oleh BuzzFeed (Eksperimen Publik)

Pada tahun 2018, BuzzFeed merilis video deepfake Barack Obama, yang dalam video tersebut tampak berbicara hal-hal yang tidak pernah ia ucapkan. Video ini merupakan kolaborasi antara BuzzFeed dan sutradara Jordan Peele sebagai eksperimen publik untuk menunjukkan betapa mudahnya manipulasi wajah dan suara dilakukan menggunakan AI.



Sumber : <https://nccnews.newhouse.syr.edu/deepfake-technology-can-make-convincing-fake-videos-of-real-people/>

Tujuan eksperimen tersebut adalah meningkatkan kesadaran publik bahwa siapa pun dapat menjadi korban deepfake, dan bahwa teknologi ini berkembang lebih cepat daripada mekanisme deteksinya. (Chesney, 2019) menilai video ini sebagai salah satu momen terpenting dalam edukasi publik tentang ancaman etis dan sosial dari teknologi deepfake.

Pembahasan

Fenomena Deepfake dan Disinformasi

Terjadinya deepfake dan disinformasi pada saat ini masih banyak orang yang belum menyadarinya. Deepfake berasal dari kata deep learning (pembelajaran mendalam) dan fake (palsu). Deepfake merupakan teknologi manipulasi video dan audio yang memanfaatkan kecerdasan buatan AI untuk menciptakan konten yang seolah-olah nyata, padahal sebenarnya palsu. Deepfake memanfaatkan algoritma deep learning untuk meniru wajah seseorang dengan sangat akurat, termasuk sudut pandang dan ekspresi. Namun, hal demikian juga menghadirkan sisi positif di bidang hiburan dan efek visual, dampaknya jauh lebih mengkhawatirkan (Fitri et al., 2025). Kemudahan dalam pembuatan salah satunya melalui <https://gemini.google.com/> secara gratis yang menimbulkan resiko penyebaran informasi palsu, manipulasi opini publik dan merusak reputasi individu maupun orang lain.

Teknologi saat ini terus berkembang, sejak 2017 deepfake telah muncul dengan menghasilkan video dan audio yang semakin sulit dibedakan dengan yang asli. Melalui kecerdasan buatan AI, kemampuan untuk mengganti wajah, suara dalam video menimbulkan kekhawatiran yang nyata terkait keamanan privasi. Pada dimensi wacana, deepfake berhasil membangun kohesi dan koherensi visual yang meyakinkan melalui teknologi AI, memalsukan mimik wajah dan intonasi suara tokoh publik. Konten ini merupakan bentuk penipuan digital yang canggih (Ariana Putri & Musyafak, 2022). Hal ini menjadi semakin mengkhawatirkan, mengingat siapa saja yang memiliki jejak digital di media sosial bisa menjadi sasaran. Pola penyebarannya (praktik wacana) sangat bergantung pada algoritma media sosial yang cenderung mempromosikan konten sensasional dan provokatif, menyebabkan deepfake tertentu mencapai viralitas tinggi dengan cepat.

Di sisi lain, disinformasi merupakan penyebaran informasi secara sengaja yang dibuat tidak akurat dan menyesatkan. Tujuannya untuk memengaruhi opini publik, menciptakan kesan yang keliru, merugikan reputasi orang lain secara individu maupun kelompok. Hal ini berbeda dengan misinformasi, yaitu informasi yang salah yang tersebar tanpa kesengajaan untuk menipu. Maka, disinformasi dan misinformasi merupakan dua hal yang berbeda dan maknanya keterbalikan antara satu sama lain. Disinformasi memanfaatkan kecepatan algoritma media sosial untuk menyebarluaskan pesan palsu secara masif. Algoritma cenderung memprioritaskan konten yang memicu emosi kuat dan kontroversi, membuat deepfake lebih mudah mencapai viralitas, sehingga mengeliminasi peluang bagi fact-checker untuk melakukan verifikasi secara tepat waktu.

Dengan demikian, fenomena deepfake dan disinformasi sebagai alat manipulasi yang sangat merusak dan keduanya saling terhubung antara satu sama lain. Perkembangan teknologi deepfake yang dapat menghasilkan konten audio dan visual tampak terlihat autentik tetapi realitanya diproduksi secara digital yang dapat menimbulkan masalah serius. Melalui kecerdasan buatan (AI), teknologi ini dapat menciptakan suara, wajah, atau adegan video yang sulit dibedakan dari aslinya. Kondisi tersebut menjadi keresahan publik untuk membedakan antara informasi yang akurat dan palsu yang menjadi masalah yang serius di ranah sosial dan komunikasi publik (Putra, 2025).

Pembentukan Persepsi melalui Framing (Entman)

Strategi framing yang dikemukakan oleh Robert Entman (Entman, 1993) untuk mengidentifikasi bagaimana pesan dan makna dimanipulasi melalui struktur bahasa dan visual. Dalam konteks deepfake dan disinformasi, proses framing ini bekerja pada dua level: visual (melalui teknologi AI) dan naratif (melalui teks penyerta). Konten deepfake yang efektif selalu terstruktur untuk mencapai empat fungsi framing secara sempurna: 1) Mendefinisikan Masalah (Problem Definition); 2) Mendiagnosis Penyebab (Causal Interpretation), 3) Evaluasi Moral (Moral Evaluation) dan 4) Rekomendasi Solusi (Treatment Recommendation). Strategi framing ini menjelaskan mengapa deepfake begitu kuat: ia tidak hanya menyampaikan kebohongan, tetapi juga menyediakan kerangka kognitif yang lengkap bagi audiens untuk memahami dan bertindak berdasarkan kebohongan tersebut.

Krisis Literasi Media dan Konstruksi Realitas Palsu (Analisis Konstruksi Sosial)

Tingginya tingkat sharing konten deepfake yang belum diverifikasi adalah indikasi nyata kegagalan literasi media. Literasi media merupakan suatu keterampilan yang dapat dipelajari oleh setiap orang, dengan tujuan untuk mengembangkan proses berpikir kritis dan aktif berpartisipasi dalam masyarakat dan budaya media. Sebagai upaya dalam memberikan kebebasan pada tiap orang dengan keberdayaan yang mereka miliki untuk memahami tentang media (Owon, 2024). Namun, realitanya masih dibutuhkan pemahaman secara mendalam tentang literasi media. Perkembangan

tentang literasi media di Indonesia dilihat paling lambat, hal ini disebabkan banyaknya masyarakat belum dibekali dengan keterampilan verifikasi kritis yang dibutuhkan untuk mendeteksi manipulasi AI. Krisis ini menciptakan lingkungan yang rentan terhadap penipuan digital (digital deception) dan membahayakan kohesi sosial, menuntut perubahan fokus dari literasi media tradisional ke literasi AI dan digital forensic.

Perkembangan teknologi kecerdasan buatan atau sering disebut dengan Artificial Intelligence (AI) telah menciptakan perubahan besar dalam berbagai bidang, khususnya pada media dan literasi digital. AI tidak hanya sebatas bagian dari teknologi dan industri, melainkan telah menjadi elemen dalam komunikasi, pendidikan, maupun pemberdayaan masyarakat. Dalam konteks literasi digital AI memiliki potensi yang luar biasa untuk meningkatkan kemampuan individu maupun kelompok dalam mengakses, memahami dan mengolah informasi digital. Maka, pemanfaatan AI dalam media mempunyai dampak positif dan negatif, maka perlu dikaji secara mendalam untuk memahami peluang yang dapat dioptimalkan serta tantangan yang dapat diantisipasi pada waktu mendatang (Hidayat et al., 2025).

Penggunaan AI dalam media dan literasi digital mempunyai peluang besar dalam meningkatkan kualitas pendidikan, memperoleh informasi, serta memperkuat kemampuan adaptif masyarakat pada perkembangan zaman. Harapannya dengan penggunaan AI yang bijak dapat memberikan manfaat yang berguna untuk masyarakat. Begitu pula pemerintah Indonesia telah mengambil langkah-langkah dalam mengatur teknologi AI. Adapun peraturan yang berkaitan dengan pemanfaatan teknologi AI di Indonesia tercantum di Permenkominformo Nomor 3 Tahun 2021 yang mengatur atas izin perilaku usaha dalam penggunaan AI, UU ITE dan peraturan turunannya yang mencakup terminologi agen elektronik, UU Pelindungan Data Pribadi yang mengatur pemrosesan data pribadi yang melibatkan AI, serta panduan etika pemanfaatan AI yang dikeluarkan oleh Kementerian Komunikasi dan Informatika dalam Surat Edaran Menkominformo Nomor 9 Tahun 2023 tentang Etika Kecerdasan Buatan (Juditha, 2025). Adanya regulasi tersebut sebagai upacara pencegahan dalam penyebaran hoaks AI dan memberikan sanksi bagi pelaku yang menyalahgunakan untuk menyebarkan informasi palsu melalui konten yang telah diproduksi. Dengan demikian, fenomena deepfake AI bukan hanya tantangan teknologi, melainkan panggilan mendesak untuk reorientasi literasi media menjadi literasi kritis AI dan digital citizenship, yang menekankan pada pertahanan kognitif terhadap rekayasa realitas digital.

Etika Komunikasi dalam Islam

Etika komunikasi mempunyai standar etis yang digunakan oleh komunikator dan komunikan. Etika dapat diartikan sebagai ilmu yang mempelajari tentang apa yang baik dan buruk, serta hak dan kewajiban moral. Etika sebagai moral dalam mengatur nilai dan norma yang menjadi pegangan dalam perilaku manusia. Sedangkan komunikasi menjadi proses penyampaian pesan yang berupa pikiran, kehendak dan perasaan dari komunikator ke komunikan baik secara lisan maupun tertulis (Saili & Akbar, 2025). Adapun etika komunikasi dalam Islam didasarkan pada ajaran Al-Qur'an dan Hadis dengan menekankan nilai integritas, tanggung jawab, serta perhatian dalam komunikasi. Komunikasi dalam Islam tidak hanya dilihat dari penyampaian pesan namun kebenaran dan kemaslahatan sosial.

Prinsip-prinsip etika komunikasi dalam Islam meliputi (Ahmad Muhamad Mustain Nasoha et al., 2025); 1) Kejujuran (sidq), menuntut setiap individu untuk menyampaikan kebenaran dan menolak segala bentuk pemalsuan. Deepfake secara fundamental melanggar sidq karena bertujuan menciptakan ilusi realitas melalui kebohongan audio dan visual, 2) Tanggung jawab (amanah yang menuntut pertanggungjawaban ganda, baik dari pengembang teknologi AI agar tidak menyalahgunakan platform, maupun dari pengguna media sosial. Amanah ini menuntut individu untuk menjaga integritas informasi dan tidak menyebarkan pesan yang berpotensi merugikan atau menipu. Penerapan prinsip amanah ini sangat penting untuk mengatasi dimensi Praktik Sosial

(Fairclough) dari deepfake, yaitu mencegah penyalahgunaan kekuasaan teknologi untuk merusak tatanan sosial dan 3) verifikasi (tabayyun) dalam proses penyebaran informasi. Tabayyun menjadi mekanisme yang melawan objektivasi dan internalisasi realitas palsu yang dianalisis dalam Teori Konstruksi Sosial, menuntut audiens untuk secara aktif menanggukhkan penilaian dan meragukan otentisitas konten yang diterima, terutama jika bersifat sensasional atau mengarahkan pada tindakan tertentu. Dengan mengimplementasikan tabayyun secara ketat, masyarakat dapat meningkatkan ketahanan literasi kritis mereka, secara efektif meredam efek merusak dari deepfake dan memulihkan integritas dalam komunikasi digital.

Penguatan edukasi mengenai literasi digital pada etika komunikasi Islam dengan ketiga prinsip di atas, sebagai upaya peningkatan kesadaran akan tanggung jawab komunikasi. Penelitian ini mengungkap adanya kebutuhan mendesak untuk mengembangkan model implementasi dari ketiga prinsip etika komunikasi Islam yang sistematis dalam praktik komunikasi digital. Namun, keterbatasan studi ini terletak pada fokus normatifnya tanpa menyajikan panduan implementasi praktis. Oleh karena itu, penelitian di masa depan dianjurkan untuk mengeksplorasi strategi dan model terperinci dalam menerapkan tabayyun di konteks media sosial, sekaligus menguji dampak empirisnya terhadap perilaku pengguna dalam berbagi informasi.

Simpulan

Penelitian ini menunjukkan bahwa teknologi deepfake telah berkembang menjadi ancaman signifikan bagi ekosistem komunikasi digital, khususnya melalui kemampuannya merekayasa wajah, suara, dan konteks secara sangat realistis sehingga mengaburkan batas antara fakta dan manipulasi. Temuan penelitian menegaskan bahwa deepfake tidak hanya memperburuk arus disinformasi, tetapi juga memperlemah literasi media publik, menciptakan ketidakpastian epistemik, serta mengikis kepercayaan terhadap institusi dan media. Analisis terhadap kasus dalam konteks Indonesia, seperti manipulasi video Menteri Keuangan Sri Mulyani, memperlihatkan bagaimana deepfake dapat memicu polarisasi, kemarahan publik, dan pembentukan opini berbasis persepsi emosional alih-alih bukti empiris. Kasus internasional seperti deepfake Presiden Ukraina Zelensky semakin menegaskan bahwa ancaman ini telah melampaui level individu dan kini menjadi instrumen dalam perang informasi dan propaganda geopolitik. Melalui pendekatan kualitatif dan analisis wacana kritis, penelitian ini menemukan bahwa deepfake bekerja melalui mekanisme framing visual dan naratif yang terstruktur, sehingga mampu membangun realitas palsu yang diterima publik sebagai kebenaran. Rendahnya literasi media dan kurangnya pemahaman tentang teknologi AI memperbesar kerentanan masyarakat dalam menerima dan menyebarkan konten manipulatif. Di sisi lain, analisis berbasis Teori Konstruksi Sosial dan Etika Komunikasi Islam menunjukkan bahwa fenomena deepfake tidak dapat dipahami hanya sebagai persoalan teknologi, tetapi juga sebagai persoalan moral, sosial, dan epistemologis yang menyangkut kejujuran (sidq), tanggung jawab (amanah), dan verifikasi informasi (tabayyun).

Kontribusi penelitian ini terletak pada penyajiannya terhadap hubungan integral antara deepfake, disinformasi, literasi media, framing digital, serta etika komunikasi. Temuan tersebut menegaskan perlunya strategi respons yang holistik, meliputi penguatan literasi AI, peningkatan mekanisme verifikasi digital, penyusunan regulasi adaptif, serta internalisasi prinsip-prinsip etika dalam komunikasi digital. Pemerintah, platform digital, institusi pendidikan, dan masyarakat perlu berkolaborasi untuk membangun ketahanan kognitif terhadap manipulasi berbasis AI.

Penelitian ini memiliki keterbatasan karena masih berfokus pada analisis normatif dan belum mengembangkan model implementasi praktis dalam penerapan prinsip tabayyun maupun etika komunikasi digital lainnya. Oleh karena itu, penelitian selanjutnya disarankan untuk merancang model intervensi literasi media berbasis AI,

menguji efektivitas metode verifikasi digital pada berbagai segmen demografis, serta mengeksplorasi strategi pencegahan disinformasi yang dapat diintegrasikan dengan kebijakan nasional dan teknologi deteksi otomatis. Temuan penelitian ini diharapkan dapat memperkuat pemahaman akademik serta memberikan landasan bagi perumusan kebijakan publik dalam menghadapi eskalasi risiko deepfake di era kecerdasan buatan.

Ucapan Terima Kasih

Kami menyampaikan ucapan terima kasih kepada semua pihak yang telah memberikan dukungan moral, akademik, maupun teknis dalam proses penyusunan penelitian ini. Ucapan terima kasih khusus kami tujukan kepada rekan-rekan peneliti yang telah melengkapi dan membantu proses penyelesaian penelitian ini. Akademisi dan reviewer awal, yang turut memberikan umpan balik berharga sehingga penelitian ini menjadi lebih matang, terstruktur, dan relevan dengan konteks perkembangan teknologi kecerdasan buatan. Seluruh pihak yang berkontribusi dalam penyediaan sumber data, baik dari platform media sosial maupun portal berita daring, sehingga memungkinkan penulis melakukan pemetaan kasus deepfake secara komprehensif.

References

- Ahmad Muhamad Mustain Nasoha, Ashfiya Nur Atqiya, Hilmi Khoiri Thohir, Natasha Aurelia Ramadhani, & Rosya Ahya Sabila. (2025). Etika Komunikasi dalam Islam : Analisis terhadap Konsep Tabayyun dalam Media Sosial. *ALADALAH: Jurnal Politik, Sosial, Hukum Dan Humaniora*, 3(2), 224–232. <https://doi.org/10.59246/aladalah.v3i2.1315>
- Anggraini, N., Uin, P., Thaha, S., Jambi, S., & History, A. (2023). Etika Komunikasi Bagi Pengguna Media Sosial Menurut Al-Qur'an Lisencing. *Journal of Comprehensive Islamic Studies (JOCIS)*, 2(2), 1–9. <https://doi.org/10.56436/jocis.v2i2.242>
- Ariana Putri, F., & Musyafak, N. (2022). Pesan Komunikasi Dalam Penipuan Digital Pada Masa Pandemi Covid-19 Communication Messages In Digital Deception During The Covid-19 Pandemic. <https://www.kominfo>.
- Berger, P. L. , & L. T. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Anchor Books.
- Boediman, A. (2025). Deepfake dan keruntuhan kredibilitas media di era AI. *Jurnal Penelitian Digital*, 10(2), 45–60.
- Chesney, R., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820.
- Diakopoulos, N. (2022). Manipulated media and the Zelensky deepfake case: Implications for digital journalism. *Digital Journalism Review*, 10(4), 512–530.
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Fahrudin, R. (2025). Kemampuan generasi X dalam mendeteksi video deepfake di Indonesia. *Jurnal Literasi Digital Indonesia*, 7(1), 12–25.
- Fallis, D. (2021). The liar's dividend and the epistemic threat of deepfakes. *Philosophy & Technology*, 34, 1–20.
- Fitri, D., Hidayah, A. N., Putri, A., Tanjung, N. H., & Izzati, S. (2025). Deepfake dan Krisis Kepercayaan: Analisis Hukum terhadap Penyebaran Konten Palsu di Media Sosial. *JiIC: Jurnal Intelek Insan Cendikia*, 2 (6).
- Hidayat, T., Dian Nugraha, H., Ramzi, M. N., Bahasa, P., & Banten, H. (2025). Penggunaan Artificial Intelligence Dalam Media Dan Literasi Digital: Peluang Dan Tantangan The Use Of Artificial Intelligence In Media And Digital Literacy: Opportunities And Challenges. <https://jicnusantara.com/index.php/jiic>.

- Irsyad, F. (2025). Tantangan jurnalisme dalam mendeteksi konten manipulatif berbasis AI. *Jurnal Media dan Informasi*, 5(1), 41–58.
- Jones, M. (2023). Media literacy interventions and deepfake detection skills: An empirical study. *Journal of Digital Communication Research*, 18(1), 22–39.
- Juditha, C. (2025). Artificial Intelligence Dan Hoaks: Tantangan Disrupsi Bagi Literasi Digital Masyarakat Di Komunitas Online Dan Upaya Antisipasi Pemerintah. *Jurnal Ilmiah Manajemen Informasi Dan Komunikasi*, 9(1).
- NCC News. (2025). “Deepfake” technology can make convincing fake videos of real people. NCC News Online, S.I. Newhouse School of Public Communications, Syracuse University. <https://nccnews.newhouse.syr.edu/deepfake-technology-can-make-convincing-fake-videos-of-real-people/>
- Ni, A. S., & Sartini, W. (2017). Wacana Perlawanan Persebaya 1927 terhadap PSSI: Analisis Wacana Kritis Norman Fairclough (Persebaya 1927’s Resistance Against PSSI: A Norman Fairclough’s Critical Discourse Analysis Study) (Vol. 17, Issue 2).
- Nurdin, M. (2025). Dampak deepfake terhadap keamanan informasi nasional. *Jurnal Komunikasi dan Keamanan Siber*, 4(3), 55–70.
- Owon, R. A. S. dkk. (2024). Pengantar Ilmu Pendidikan Teori dan Invoasi Pengembangan SDM (Pertama). Widina Media Utama.
- Putra, Z. F. (2025). Deepfake Dan Disinformasi: Krisis Kebenaran Di Era Digital Kajian Filsafat Dan Etika Komunikasi.
- Saili, N. S., & Akbar, A. (2025). Analisis Evaluasi Etika Komunikasi dalam Perspektif Studi Al-Qur’an. *Jurnal Teologi Islam*, 1 (2). <https://doi.org/doi.org/10.63822/jsqe3c55>.
- Tempo. (2025, Agustus). Polemik deepfake “guru beban negara” usai pernyataan Sri Mulyani. Tempo. <https://www.tempo.co/ekonomi/polemik-deepfake-guru-beban-negara-usai-pernyataan-sri-mulyani-2061974>.
- University of Virginia. (2022, Maret). Q&A: With Zelenskyy surrender hoax, the feared future of deepfakes is here. UVA Today. <https://news.virginia.edu/content/qa-zelenskyy-surrender-hoax-feared-future-deepfakes-here>.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact on political communication. *Social Media + Society*, 6(1), 1–13.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52.
- Widjaja, S. (2025). Fenomena post-truth dalam era deepfake dan kecerdasan buatan. *Jurnal Komunikasi Kontemporer*, 9(1), 30–44.
- Zahro, L. (2024). Infocalypse dan krisis kepercayaan publik di era AI. *Jurnal Teknologi dan Masyarakat*, 3(2), 66–79.