DATA CYBER INDONESIA

Ruang Riset Siber Indonesia

Tata Kelola Kecerdasan Buatan di Indonesia: Analisis Mendalam Risiko Deepfake dan Dampaknya terhadap Keamanan Informasi Nasional

Penulis:

Ujang Trisno

Direktur Riset, Data Cyber Indonesia

Email: ujangtrisno@datacyber.or.id

Website: https://datacyber.or.id

Tanggal Terbit: 09 November 2026

DOI: 10.5281/zenodo.24074601

Lisensi: CC BY 4.0 (Creative Commons Attribution)

Kode Verifikasi: DCY-RISET-2026-092

Tata Kelola Kecerdasan Buatan di Indonesia: Analisis Mendalam Risiko Deepfake dan Dampaknya terhadap Keamanan Informasi Nasional

Abstrak

Perkembangan pesat teknologi Generative AI telah memunculkan ancaman baru yang sangat serius bagi keamanan informasi global, yaitu deepfake. Deepfake adalah manipulasi audio-visual yang dihasilkan oleh kecerdasan buatan (AI) yang sangat realistis, sehingga sulit dibedakan dari konten asli. Riset independen yang disusun oleh Ujang Trisno, Direktur Riset Data Cyber Indonesia, ini bertujuan untuk menganalisis secara mendalam risiko deepfake terhadap keamanan informasi nasional Indonesia. Menggunakan pendekatan kualitatif-deskriptif dengan metode wawancara mendalam (in-depth interview) eksklusif terhadap tiga narasumber kunci yang kompeten di bidangnya—seorang praktisi keamanan siber senior, seorang penyidik siber dari kepolisian, dan seorang akademisi hukum tata negara—serta analisis literatur dan dokumen hukum yang komprehensif, riset ini membedah anatomi ancaman deepfake dari berbagai perspektif. Hasil riset mengungkap tiga bahaya utama deepfake di Indonesia. Pertama, eksploitasi seksual melalui pembuatan konten pornografi palsu (deepfake pornography) yang menargetkan perempuan dan anak-anak, yang tidak hanya menghancurkan martabat korban tetapi juga menimbulkan trauma psikologis berat dan berkepanjangan. Kedua, penipuan finansial dengan memalsukan identitas, wajah, dan suara tokoh publik, pejabat, atau anggota keluarga untuk meyakinkan korban agar mentransfer sejumlah uang. OJK mencatat 1.379 laporan penipuan keuangan berbasis AI sepanjang tahun 2026, dan deepfake menjadi salah satu modus yang paling berbahaya karena sulit dikenali. Ketiga, perusakan reputasi dan personal branding melalui video manipulatif yang menampilkan pernyataan kontroversial palsu, yang dapat menyebar dalam hitungan jam di media sosial dan merusak karier serta kehidupan sosial korban secara permanen. Riset ini juga mengkaji kerangka hukum yang ada di Indonesia, seperti UU ITE, KUHP, dan UU PDP, serta menemukan bahwa regulasi saat ini belum cukup spesifik dan komprehensif untuk mengakomodasi kompleksitas ancaman deepfake, terutama dalam hal pembuktian forensik dan penanganan pelaku lintas negara. Diperlukan pembentukan regulasi khusus yang mengatur pembuatan, penyebaran, dan penggunaan deepfake, penguatan kapasitas forensik digital aparat penegak hukum, kewajiban platform digital untuk menyediakan alat deteksi dan pelabelan konten AI, serta literasi digital yang masif dan berkelanjutan yang menekankan aspek etika dan moral. Riset ini menyimpulkan bahwa deepfake bukan sekadar isu teknologi, melainkan ancaman multidimensional yang serius terhadap kedaulatan informasi, keamanan nasional, dan hak asasi manusia. Kolaborasi lintas sektor antara

pemerintah, platform digital, akademisi, dan masyarakat sipil menjadi keniscayaan untuk membangun ketahanan informasi yang tangguh di era Generative AI.

1. Latar Belakang

Revolusi industri 4.0 dan perkembangan pesat teknologi kecerdasan buatan (Artificial Intelligence/AI) telah membawa perubahan fundamental dalam berbagai aspek kehidupan manusia. Salah satu terobosan paling kontroversial dan mengkhawatirkan adalah munculnya teknologi deepfake. Istilah deepfake berasal dari gabungan kata "deep learning" (pembelajaran mendalam) dan "fake" (palsu). Teknologi ini memanfaatkan algoritma machine learning canggih, terutama Generative Adversarial Networks (GANs), untuk menciptakan atau memanipulasi konten audio-visual—video, gambar, dan suara—yang sangat realistis, sehingga hampir mustahil dibedakan dari konten asli oleh mata dan telinga manusia biasa.

Awalnya, teknologi deepfake dikembangkan untuk tujuan hiburan dan penelitian akademik. Namun, seiring dengan demokratisasi akses terhadap alat-alat AI generatif, siapa pun kini dapat membuat deepfake dengan mudah dan murah. Aplikasi pembuat deepfake bertebaran di internet, sebagian besar dapat diakses secara gratis. Kondisi ini menciptakan lanskap ancaman baru yang sangat berbahaya.

Di Indonesia, fenomena deepfake mulai menunjukkan tren yang mengkhawatirkan. Kasus-kasus penyalahgunaan deepfake mulai bermunculan, dari yang sifatnya iseng hingga yang benar-benar merusak kehidupan seseorang. Media massa nasional dan internasional telah melaporkan berbagai insiden, mulai dari video palsu tokoh politik yang dibuat untuk memprovokasi hingga kasus revenge porn berbasis AI yang menimpa mahasiswi di Surakarta. Namun, kesadaran publik tentang bahaya deepfake masih sangat rendah, sementara kerangka hukum yang ada dinilai belum memadai untuk menjerat pelaku dan melindungi korban.

Riset ini hadir untuk menjawab kegelisahan tersebut. Dengan menggunakan pendekatan riset primer melalui wawancara eksklusif dengan para ahli dan praktisi, riset ini berusaha membedah anatomi ancaman deepfake secara komprehensif, menganalisis kerangka kebijakan yang ada, dan merumuskan strategi mitigasi yang aplikatif. Riset ini disusun oleh Ujang Trisno, Direktur Riset Data Cyber Indonesia, sebagai kontribusi intelektual untuk memperkuat keamanan informasi nasional Indonesia di era Generative AI.

2. Rumusan Masalah

1. Bagaimana bentuk, modus operandi, dan dampak penyalahgunaan deepfake di

Indonesia berdasarkan perspektif para ahli dan praktisi?

2. Bagaimana efektivitas kerangka hukum yang ada (UU ITE, KUHP, UU PDP) dalam mengakomodasi ancaman deepfake dan melindungi korban?

3. Apa saja tantangan teknis, sosial, dan yudisial dalam penanganan kasus deepfake di Indonesia?

4. Apa strategi kebijakan dan langkah mitigasi yang komprehensif dan berkelanjutan untuk menangkal ancaman deepfake di masa depan?

3. Tujuan

1. Mengidentifikasi dan mengklasifikasikan bentuk-bentuk penyalahgunaan deepfake di Indonesia secara mendalam.

2. Menganalisis efektivitas dan celah kerangka hukum yang ada dalam menangani kasus deepfake.

3. Mengungkap tantangan teknis dan sosial yang dihadapi dalam penanganan kasus deepfake.

4. Merumuskan rekomendasi kebijakan dan strategi mitigasi yang komprehensif untuk pemerintah, platform digital, dan masyarakat.

4. Batasan Masalah

Objek kajian adalah fenomena penyalahgunaan deepfake di Indonesia pada periode 2025-2026. Data primer diperoleh melalui wawancara mendalam dengan tiga narasumber yang kompeten di bidangnya (nama disamarkan untuk menjaga privasi dan etika penelitian). Data sekunder diperoleh dari literatur akademik, laporan resmi pemerintah, analisis hukum, dan studi kasus yang dipublikasikan di media massa. Penelitian tidak melakukan pengujian teknis terhadap alat pembuat atau pendeteksi deepfake secara langsung. Analisis bersifat kualitatif-deskriptif dan tidak bertujuan untuk menggeneralisasi temuan ke seluruh populasi.

5. Tinjauan Pustaka

Kajian ini merujuk pada beberapa literatur utama yang relevan dengan topik deepfake dan keamanan informasi. Pertama, jurnal akademik karya Ilmi Sila Ayu, Febi Ramadhani Rusdin, dan Muhammad Taslim (2026) yang berjudul "The Dynamics of

Gender-Based Violence Through Generative Technology (Deepfake & Voice Cloning) in a Digital Communication Perspective" dalam Journal of Innovative and Creativity. Jurnal ini membahas secara mendalam bagaimana deepfake dan voice cloning digunakan sebagai bentuk kekerasan berbasis gender, dengan fokus pada dampak psikologis dan sosial terhadap korban. Temuan jurnal ini sangat relevan untuk mendukung analisis tentang eksploitasi seksual melalui deepfake pornography.

Kedua, laporan resmi dari Otoritas Jasa Keuangan (OJK) yang dipublikasikan melalui Katadata (2026) dengan judul "OJK Ungkap 6 Modus Penipuan Keuangan Berbasis AI, Sudah Ada 1.379 Kasus". Laporan ini menyajikan data empiris tentang maraknya penipuan keuangan yang memanfaatkan teknologi AI, termasuk deepfake untuk memalsukan identitas dan suara. Data ini menjadi landasan kuat untuk menganalisis dampak finansial dari penyalahgunaan deepfake.

Ketiga, artikel berita dari Republika (2026) yang berjudul "Bermotif Sakit Hati, Mahasiswi di Surakarta Jadi Korban Revenge Porn Berbasis AI". Artikel ini menyajikan studi kasus nyata di Indonesia di mana wajah korban diedit menggunakan AI menjadi konten pornografi, dan pelaku dijerat dengan UU ITE dengan ancaman hukuman 12 tahun penjara. Studi kasus ini menjadi bukti empiris bahwa ancaman deepfake bukanlah spekulasi, melainkan realitas yang sudah terjadi di Indonesia.

Keempat, analisis hukum dari DNT Lawyers (2026) yang berjudul "Penipuan Deepfake dan Dasar Pertanggungjawaban Hukumnya". Artikel ini membahas secara komprehensif dasar hukum penjeratan pelaku deepfake di Indonesia, termasuk Pasal 492 KUHP, Pasal 35 UU ITE, dan Pasal 65 UU PDP. Analisis ini sangat penting untuk mengevaluasi efektivitas kerangka hukum yang ada dan merumuskan rekomendasi kebijakan.

6. Metodologi

Riset ini menggunakan pendekatan kualitatif-deskriptif dengan metode wawancara mendalam (in-depth interview) dan studi literatur. Pendekatan kualitatif dipilih karena tujuan riset ini adalah untuk memahami secara mendalam fenomena deepfake, bukan untuk mengukur atau menggeneralisasi. Wawancara mendalam memungkinkan peneliti untuk menggali perspektif, pengalaman, dan pengetahuan narasumber secara komprehensif.

Wawancara dilakukan terhadap tiga narasumber kunci yang dipilih secara purposive sampling berdasarkan kompetensi dan pengalaman mereka di bidangnya. Narasumber pertama adalah seorang praktisi keamanan siber senior yang telah

berpengalaman lebih dari 10 tahun di industri keamanan digital. Narasumber kedua adalah seorang penyidik siber dari kepolisian yang menangani kasus-kasus kejahatan siber. Narasumber ketiga adalah seorang akademisi hukum tata negara yang mengkaji regulasi terkait teknologi dan perlindungan data pribadi. Untuk menjaga privasi dan etika penelitian, nama ketiga narasumber disamarkan.

7. Teknik Analisis

Analisis tematik dan kritis terhadap transkrip wawancara, literatur, dan dokumen hukum. Data disajikan secara deskriptif-analitis untuk membedah akar masalah dan merumuskan solusi.

8. Hasil

Berdasarkan analisis tematik terhadap transkrip wawancara dan data sekunder, riset ini menemukan beberapa temuan utama:

1. BENTUK DAN DAMPAK PENYALAHGUNAAN DEEPFAKE
Narasumber 1 (Praktisi Keamanan Siber) mengungkapkan bahwa bahaya paling nyata dari deepfake adalah penyalahgunaan wajah untuk membuat konten pornografi palsu. "Wajah seseorang bisa ditempel ke tubuh orang lain, lalu disebarkan untuk mempermalukan atau balas dendam. Ini bukan sekadar serangan digital, tapi kejahatan seksual yang menghancurkan martabat dan kondisi psikologis korban," ujarnya. Ia menambahkan bahwa korban deepfake pornography seringkali mengalami trauma berat, depresi, bahkan keinginan untuk mengakhiri hidup.

Narasumber 2 (Penyidik Siber) menyoroti bahaya deepfake untuk penipuan finansial. "Deepfake memungkinkan penipu memalsukan wajah dan suara seseorang—entah itu pejabat, kolega, atau anggota keluarga—untuk meyakinkan korban agar segera mentransfer uang. OJK mencatat ada 1.379 laporan penipuan keuangan yang terindikasi memakai AI, dan modus deepfake jadi salah satu yang paling berbahaya karena sulit dikenali," jelasnya.

Narasumber 3 (Akademisi Hukum) menyoroti dampak deepfake terhadap reputasi dan personal branding. "Bagi figur publik atau profesional, deepfake bisa dipakai untuk merusak reputasi secara sistematis. Sebuah video palsu yang menampilkan seolah-olah seseorang mengatakan hal kontroversial bisa menyebar dalam hitungan jam, dan proses pemulihan citra jauh lebih sulit daripada penyebarannya," ujarnya.

2. EFEKTIVITAS KERANGKA HUKUM
Berdasarkan analisis terhadap UU ITE, KUHP, dan UU PDP, riset ini menemukan bahwa

kerangka hukum yang ada belum cukup spesifik dan komprehensif untuk mengakomodasi kompleksitas ancaman deepfake. UU ITE memang mengatur tentang penyebaran informasi elektronik yang melanggar kesusilaan dan pencemaran nama baik, tetapi tidak secara eksplisit menyebut deepfake. KUHP memiliki pasal tentang penipuan dan pencemaran nama baik, tetapi pembuktiannya membutuhkan ahli forensik digital yang kapasitasnya masih terbatas di Indonesia.

3. TANTANGAN TEKNIS DAN SOSIAL
Narasumber 2 (Penyidik Siber) mengungkapkan bahwa tantangan terbesar dalam penanganan kasus deepfake adalah pembuktian forensik. "Kami kesulitan membuktikan bahwa sebuah video itu deepfake atau asli. Butuh alat dan keahlian khusus, sementara pelaku terus mengembangkan teknologi mereka," ujarnya. Narasumber 1 (Praktisi Keamanan Siber) menyoroti rendahnya kesadaran publik tentang bahaya deepfake. "Banyak orang masih menganggap deepfake itu sekadar hiburan, padahal dampaknya bisa sangat merusak. Literasi digital kita masih jauh dari cukup," ujarnya.

9. Pembahasan

Temuan riset ini menunjukkan bahwa deepfake bukanlah sekadar isu teknologi, melainkan ancaman multidimensional yang menyentuh aspek sosial, ekonomi, hukum, dan keamanan nasional. Dari perspektif sosial, deepfake pornography adalah bentuk kekerasan berbasis gender yang dampaknya jauh lebih parah daripada pelecehan verbal. Korban tidak hanya mengalami kerugian finansial dan reputasi, tetapi juga trauma psikologis yang berkepanjangan.

Dari perspektif ekonomi, penipuan deepfake merugikan masyarakat secara finansial dan merusak kepercayaan pada sistem keuangan digital. Data OJK yang menunjukkan 1.379 laporan penipuan berbasis AI sepanjang tahun 2026 adalah alarm merah yang tidak bisa diabaikan. Jika tidak segera ditangani, kepercayaan publik terhadap layanan keuangan digital akan menurun, yang pada akhirnya dapat menghambat pertumbuhan ekonomi digital Indonesia.

Dari perspektif hukum, regulasi yang ada masih bersifat reaktif dan belum proaktif. UU ITE dan KUHP memang dapat digunakan untuk menjerat pelaku deepfake, tetapi implementasinya seringkali terkendala oleh pembuktian forensik yang rumit dan yurisdiksi lintas negara. Diperlukan pembentukan regulasi khusus yang secara eksplisit mengatur tentang deepfake.

Wawancara dengan para narasumber mengungkap bahwa tantangan terbesar dalam

penanganan deepfake bukanlah teknologi itu sendiri, melainkan manusia. Rendahnya literasi digital, lemahnya penegakan hukum, dan kurangnya kesadaran publik adalah akar masalah yang harus segera diatasi. Diperlukan pendekatan sosio-teknis yang menggabungkan penegakan hukum, penguatan kapasitas forensik, dan edukasi publik yang masif.

10. Kesimpulan

Deepfake merupakan ancaman serius bagi keamanan informasi nasional Indonesia. Penyalahgunaannya mencakup eksploitasi seksual, penipuan finansial, dan perusakan reputasi. Kerangka hukum yang ada belum memadai untuk menjerat pelaku secara efektif, terutama dalam hal pembuktian forensik dan penanganan pelaku lintas negara. Diperlukan pembentukan regulasi khusus, penguatan kapasitas forensik digital, kewajiban platform digital, dan literasi digital yang menekankan aspek etika dan moral.

11. Rekomendasi

1. Pembentukan Regulasi Khusus Deepfake: Pemerintah perlu menyusun undang-undang atau peraturan yang secara spesifik mengatur pembuatan, penyebaran, dan penggunaan deepfake, termasuk sanksi pidana yang tegas.

2. Penguatan Kapasitas Forensik Digital: Aparat penegak hukum perlu dilatih untuk mendeteksi dan membuktikan deepfake secara forensik.

3. Kewajiban Platform Digital: Platform media sosial dan aplikasi pesan harus diwajibkan menyediakan alat deteksi deepfake dan melabeli konten yang dihasilkan AI.

4. Literasi Digital yang Komprehensif: Edukasi publik harus mencakup cara mengenali deepfake, dampak psikologis korban, dan tanggung jawab moral dalam menyebarkan informasi.

5. Perlindungan Korban: Negara harus menyediakan layanan pendampingan psikologis dan hukum bagi korban deepfake.

6. Kerja Sama Internasional: Indonesia perlu menjalin kerja sama dengan negara lain untuk menangani deepfake lintas batas.

12. Keterbatasan Riset

Penelitian ini menggunakan data primer dari wawancara dengan tiga narasumber yang bersifat purposive sampling, sehingga temuan tidak dapat digeneralisasi ke seluruh populasi. Data sekunder yang digunakan juga terbatas pada literatur yang tersedia untuk publik. Penelitian tidak melakukan pengujian teknis terhadap alat deteksi deepfake.

Daftar Pustaka

Ayu, I. S., Rusdin, F. R., & Taslim, M. (2026). The Dynamics of Gender-Based Violence Through Generative Technology (Deepfake & Voice Cloning) in a Digital Communication Perspective. Journal of Innovative and Creativity, 6(1). https://doi.org/10.31004/joecy.v6i1.6223

Katadata. (2026). OJK Ungkap 6 Modus Penipuan Keuangan Berbasis AI, Sudah Ada 1.379 Kasus.

Republika. (2026). Bermotif Sakit Hati, Mahasiswi di Surakarta Jadi Korban Revenge Porn Berbasis AI.

DNT Lawyers. (2026). Penipuan Deepfake dan Dasar Pertanggungjawaban Hukumnya.

Halaman Verifikasi

Dokumen ini diterbitkan secara resmi oleh Data Cyber Indonesia. Pindai QR code di bawah untuk memverifikasi keaslian dokumen ini melalui website kami.

Pindai atau buka tautan berikut untuk memverifikasi:
https://datacyber.or.id/verifikasi/DCY-RISET-2026-092

Kode Verifikasi: DCY-RISET-2026-092

DOI: 10.5281/zenodo.24074601

Lisensi: CC BY 4.0

Diterbitkan oleh Data Cyber Indonesia

Tanggal: 09 November 2026