JURNAL KEPEMIMPINAN & PENGURUSAN SEKOLAH
Homepage : https://ejurnal.stkip-pessel.ac.id/index.php/kp
Email : jkps.stkippessel@gmail.com

p-ISSN : 2502-6445 ; e-ISSN : 2502-6437
Vol. 10, No. 4, December 2025
Page 2590-2599

Jurnal Kepemimpinan & Pengurusan Sekolah

PERTANGGUNGJAWABAN PIDANA ATAS PENYALAHGUNAAN *ARTIFICIAL INTELLIGENCE* (AI) DALAM PRODUKSI KONTEN HOAKS DAN DEEPFAKE DI MEDIA SOSIAL

Naila Cantika[1], Feliciya Archie Hidayat[2], Patricia Sherly Indriana [3], Frengki Riski Sirait [4], Yennie Agustin Mahroennisa Rasyid[5]
[1,2,3,4,5]Universitas Lampung, Indonesia

Email: nailacantika092@gmail.com

DOI: https://doi.org/10.34125/jkps.v10i4.1435

Sections Info
Article history:
Article history:
Submitted: 23 October 2025
Final Revised: 25 November 2025
Accepted: 28 November 2025
Published: 21 December 2025
Keywords:
Criminal liability
Artificial Intelligence
Hoax content
Social media

ABSTRAK

The advancement of artificial intelligence (AI) technology offers significant opportunities in digital content production, but its improper use for hoaxes and deepfakes on social media platforms leads to adverse effects, such as damaging reputations, spreading false news, and jeopardizing information security. This paper examines criminal liability for AI misuse by considering legal aspects, difficulties in verification, and the effectiveness of regulations in Indonesia, while emphasizing the need for adaptive rules to prevent manipulated content.The main discussion includes the concept of AI as a system capable of independent learning, hoaxes as inaccurate data disseminated, and deepfakes as techniques for altering media to resemble the original. The deepfake case involving President Prabowo illustrates the real dangers. The classification of criminal acts encompasses practices of fraud and defamation. The application of relevant articles in the Indonesian Criminal Code is evaluated, although challenges in proof include technical and legal obstacles. Criminal accountability involves penalties for individuals and corporations. Gaps in legislation urge regulatory changes to align with AI advancements.

ABSTRAK

Perkembangan teknologi kecerdasan buatan (AI) memberikan kesempatan besar dalam pembuatan konten digital, tetapi penggunaannya yang tidak tepat untuk hoaks dan deepfake di platform media sosial menghasilkan efek buruk, seperti merusak reputasi, menyebarkan berita palsu, dan membahayakan keamanan informasi. Tulisan ini meneliti tanggung jawab hukum pidana terhadap penyalahgunaan AI dengan mempertimbangkan dimensi hukum, kesulitan dalam verifikasi, serta kinerja peraturan di Indonesia, sambil menyoroti perlunya aturan yang adaptif untuk menghindari konten yang dimanipulasi.Pembahasan utama meliputi pengertian AI sebagai sistem yang mampu belajar secara mandiri, hoaks sebagai data tidak benar yang disebarluaskan, dan deepfake sebagai teknik pengubahan media yang menyerupai asli. Kasus deepfake yang menimpa Presiden Prabowo menggambarkan bahaya sebenarnya. Klasifikasi perbuatan pidana mencakup praktik penipuan dan fitnah. Penggunaan pasal-pasal terkait dalam KUHP dievaluasi, walaupun hambatan dalam pembuktian mencakup tantangan teknis dan hukum. Tanggung jawab pidana melibatkan hukuman untuk orang perseorangan maupun badan usaha. Celah dalam perundang-undangan mendorong perubahan regulasi guna menyesuaikan dengan kemajuan AI.

Kata kunci: *Pertanggungjawaban pidana, Artificial Intelligence (AI), Konten hoaks, Media sosial*

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah memberikan pengaruh besar pada bagaimana masyarakat mengakses, menghasilkan, dan menyebarkan informasi. Salah satu inovasi yang berkontribusi pada perubahan ini adalah Kecerdasan Buatan (AI), yakni teknologi yang memungkinkan sistem komputer meniru kemampuan berpikir manusia melalui proses pembelajaran, analisis, dan pengambilan keputusan secara otomatis (Pyndho Cevin Taraya & Aji Wibawa 2022). Walaupun memberikan manfaat besar dalam berbagai bidang, penggunaan AI juga membuka peluang terjadinya penyalahgunaan, terutama di ruang digital yang sangat dinamis. *Artificial Intelligence* (AI) sendiri memanfaatkan adanya metode pembelajaran *deep learning* untuk mengubah dan memanipulasi tampilan wajah dan audio seseorang pada konten digital yang di sebarluaskan pada media sosial. (Kasita, 2022).

Salah satu bentuk penyalahgunaan AI yang semakin menimbulkan kekhawatiran adalah pembuatan konten *hoax* dan *deepfake*. *Deepfake* adalah teknologi yang memanfaatkan kecerdasan buatan untuk memodifikasi wajah, suara, atau gerakan seseorang sehingga tampak sangat meyakinkan meskipun tidak pernah terjadi dalam kenyataan(Syafiq et al., 2025). Ketika disebarkan melalui media sosial, konten *deepfake* dapat menimbulkan kebingungan publik, merusak reputasi, hingga menimbulkan kerugian materiil maupun immateriil bagi pihak yang terdampak (Sari & Harwika 2022).

Fenomena tersebut semakin nyata terlihat dari kasus viral pada tahun 2025, yakni beredarnya video *deepfake* yang menampilkan wajah dan suara Presiden Prabowo Subianto serta Menteri Keuangan Sri Mulyani seolah-olah sedang meminta sejumlah uang kepada masyarakat. Video tersebut dibuat menggunakan teknologi AI secara manipulatif dan kemudian disebarkan untuk tujuan penipuan. Akibatnya, beberapa korban yang percaya terhadap konten tersebut mengalami kerugian finansial karena mentransfer uang sesuai instruksi dalam video. Kasus ini menunjukkan bahwa teknologi AI tidak hanya disalahgunakan untuk menghasilkan konten palsu, tetapi juga dapat dipakai sebagai alat melakukan tindak pidana yang memiliki dampak luas di ruang digital.

Dalam perspektif hukum pidana, penyalahgunaan AI menimbulkan tantangan serius. Konten yang dihasilkan AI sulit dibedakan dari konten asli sehingga menyulitkan proses identifikasi pelaku, penentuan bentuk kesalahan, serta pembuktian dalam proses peradilan. Selain itu, kerangka regulasi yang ada seperti KUHP dan UU ITE belum mengatur secara khusus tindak pidana yang melibatkan AI sebagai alat atau instrumen kejahatan, sehingga menimbulkan ruang abu-abu dalam penegakan hukum (Muslim Nugraha dkk 2025). Akibatnya, upaya penindakan terhadap kasus hoaks dan *deepfake* sering kali tidak sebanding dengan kompleksitas kejahatan yang dilakukan.

Melihat meningkatnya risiko penyalahgunaan teknologi AI dan dampak nyata yang ditimbulkannya, diperlukan kajian mendalam mengenai pertanggungjawaban pidana terhadap penyalahgunaan AI dalam pembuatan konten hoaks dan *deepfake* di media sosial. Kajian ini penting untuk mendorong pembentukan regulasi yang lebih adaptif, memberikan kepastian hukum, serta melindungi masyarakat dari ancaman kejahatan berbasis teknologi. Sejumlah penelitian merekomendasikan terhadap pentingnya pembentukan aturan khusus yang lebih ketat dan tegas dalam mengendalikan penggunaan teknologi. Adapun langkah-langkah yang bisa ditempuh meliputi peningkatan kesadaran akan kemampuan literasi digital di kalangan masyarakat, penguatan peraturan terhadap keamanan data pribadi, dan pengembangan sistem deteksi deepfake yang memiliki tingkat akurasi yang lebih mutakhir (Amelia et al., 2024).

METODE PENELITIAN

Penelitian ini menggunakan metode yuridis normatif yang menitikberatkan pada kajian terhadap peraturan perundang-undangan, doktrin, serta literatur yang berkaitan dengan pertanggungjawaban pidana atas penyalahgunaan *Artificial Intelligence* (AI) dalam pembuatan konten hoaks dan *deepfake*. Pendekatan ini digunakan karena permasalahan yang diteliti lebih banyak berkaitan dengan analisis norma hukum positif dan interpretasinya. Penelitian ini menggabungkan pendekatan perundang-undangan, untuk menelaah ketentuan dalam KUHP, UU ITE, dan regulasi terkait lainnya; pendekatan konseptual, guna memahami konsep *deepfake*, hoaks, unsur kesalahan, dan prinsip pertanggungjawaban; serta pendekatan kasus, dengan menjadikan contoh viral video *deepfake* Presiden Prabowo dan Menteri Sri Mulyani sebagai ilustrasi penerapan norma hukum pada situasi faktual.

Bahan hukum yang digunakan mencakup bahan hukum primer, seperti peraturan perundang-undangan dan putusan pengadilan yang terkait, serta bahan hukum sekunder, termasuk buku, artikel jurnal, dan temuan dari penelitian sebelumnya. Keseluruhan bahan hukum tersebut dikumpulkan melalui kajian pustaka guna mendukung analisis yang dilakukan. Bahan hukum yang telah terkumpul selanjutnya dianalisis dengan pendekatan deskriptif kualitatif, melalui penafsiran terhadap aturan hukum yang berlaku serta kaitannya dengan fenomena penyalahgunaan kecerdasan buatan. Pendekatan penalaran yang diterapkan adalah deduktif, yakni menyimpulkan dari ketentuan hukum umum menuju aplikasinya pada situasi spesifik.

HASIL DAN PEMBAHASAN

1. Konsep AI, Hoaks, dan *DeepFake*

a. *Artificial Intelligence* (AI)

Menurut pendapat Holmes, kecerdasan buatan, yang biasa dikenal sebagai *Artificial Intelligence* (AI), di dalam konteks pendidikan mengacu pada sistem yang secara khusus diciptakan untuk mendukung dan mempermudah proses belajar mengajar. Dampak kecerdasan buatan atau *Artificial Intelligence* (AI) terhadap perkembangan pendidikan sangat signifikan. Kecerdasan buatan yang memainkan peran penting ini berpotensi untuk merevolusi fungsi kecerdasan manusia (Pakpahan & Roida. 2021). Pergeseran ini menunjukkan peralihan fungsi kecerdasan manusia dari tugas-tugas tertentu menjadi diambil alih oleh kecerdasan buatan atau *Artificial Intelligence* (AI). Dengan kemajuan teknologi yang tak terelakkan, manusia tampaknya bersaing dengan inovasi yang mereka hasilkan sendiri. Seperti yang telah dijelaskan sebelumnya, perkembangan teknologi bertujuan untuk mempermudah kehidupan manusia, bukan untuk menggantikan posisi mereka yang utama. Dalam bidang pendidikan, pengaplikasian AI sangat tergantung pada kemampuan pengajar untuk mengadopsi teknologi ini untuk para siswa. Begitu pula, sebaiknya mahasiswa memanfaatkan AI sebagai alat untuk memaksimalkan potensi diri mereka, sehingga kualitas dan produktivitas mereka dapat berkembang sesuai dengan kemajuan zaman. Dengan AI, para mahasiswa dapat meningkatkan kemampuan mereka sejalan dengan tujuan pendidikan yang tercantum dalam Sistem Pendidikan Nasional. Teknologi AI menawarkan aksesibilitas terhadap berbagai materi ajar yang bersifat personal dan teradaptasi, sehingga mahasiswa bisa mengasah keterampilan dan pengetahuan mereka dengan cara yang lebih efisien. Selain itu, integrasi AI dalam proses pembelajaran dapat memberikan umpan balik yang cepat, membantu mahasiswa dalam memahami materi, serta merancang pengalaman belajar yang lebih responsif. Melalui pendekatan ini, implementasi AI dalam pendidikan berperan krusial dalam meningkatkan mutu pendidikan sesuai dengan standar nasional yang telah ditetapkan..

b. Hoaks

Pada abad ke-20, penyebaran berita palsu atau *hoax* mulai meluas, meskipun kata "*hoax*" telah ada sejak tahun 1808. Istilah ini berasal dari "*hocus*," yang berarti menipu, dan kata tersebut sering muncul dalam trik sulap atau pertunjukan ilusi (Sarjito & Aris. 2024). Saat ini, hoax sering dikaitkan dengan berita palsu atau informasi yang asal-usulnya tidak jelas, membuat pembaca merasa kebingungan dan cepat terprovokasi. Saat ini, memperoleh berita atau informasi melalui internet menjadi sangat mudah. Dalam lingkungan digital di Indonesia, keadaan ini menyebabkan terjadinya krisis etika, di mana isu *hoax* bisa dengan cepat berkembang dan menyebar melalui internet, menjadikan masyarakat mudah terpengaruh. *Hoax* biasanya digunakan untuk mengendalikan opini publik mengenai isu SARA dan politik, yang pada gilirannya menciptakan perpecahan sosial serta menghambat kemajuan bangsa(Rustanta et al., 2025). Penyebaran berita *hoax* di Indonesia tampaknya tidak mengenal batas, dengan sejumlah opini yang mendorong kebencian dan informasi yang salah terus-menerus dihasilkan oleh pihak yang tidak bertanggung jawab. Diskusi tentang efek berita *hoax* menunjukkan bahwa jenis informasi ini sangat merugikan bagi pengguna internet, karena tidak hanya menyebar informasi yang salah tetapi juga mengandung ujaran kebencian yang dapat memicu kekacauan dan merusak solidaritas serta persatuan bangsa Indonesia. Sebagai pengguna media sosial, kita harus lebih waspada terhadap *hoax*. Masyarakat bisa memfilter berita di internet dengan melakukan verifikasi, seperti menelusuri asal-usul berita tersebut. Oleh karena itu, penting bagi masyarakat Indonesia untuk mendapatkan pendidikan media sosial yang dapat meningkatkan kesadaran menjadi pengguna yang lebih bertanggung jawab dan dapat menyaring konten di platform tersebut. Dengan demikian, masyarakat dapat mengakses informasi yang valid dan tidak mudah terpengaruh oleh berita *hoax*.

c. *Deepfake*

Deepfake adalah teknologi yang memanfaatkan kecerdasan buatan untuk membuat atau mengubah gambar, suara, atau video agar terlihat sangat asli, padahal sebenarnya tidak benar. Nama "deepfake" berasal dari kombinasi kata "*deep learning*" dan "*fake*. " Teknologi ini bekerja dengan metode pembelajaran mesin yang canggih, terutama melalui deep learning serta jaringan saraf buatan. Ada beberapa teknik utama dalam penggunaan teknologi ini, yaitu: 1. Jaringan Adversarial Generatif (GAN) adalah metode yang paling sering digunakan untuk menghasilkan deepfake. GAN terdiri dari dua jaringan saraf yang berinteraksi satu sama lain: - Generator: menciptakan gambar atau video palsu - Discriminator: mencoba menentukan apakah konten tersebut asli atau tidak. Kedua jaringan ini bersaing. *Generator* terus memperbaiki kemampuannya untuk membuat konten yang terlihat lebih nyata, sedangkan discriminator berusaha semakin cakap dalam membedakan antara yang asli dan yang palsu. Proses ini dilakukan berulang kali sampai generator dapat menghasilkan konten yang sangat sulit dibedakan dari yang asli. 2. *Autoencoders* adalah metode yang digunakan untuk pertukaran wajah. Sistem ini mempelajari bagaimana mengkodekan fitur wajah seseorang dan kemudian menggabungkannya dengan fitur wajah orang lain, sehingga menciptakan video dengan wajah yang tertukar. 3. Untuk menghasilkan deepfake yang berkualitas tinggi, sistem memerlukan pelatihan yang meliputi: - ratusan atau ribuan gambar atau video dari subjek yang ingin dipalsukan - waktu pemrosesan yang lama (bisa memakan waktu berjam-jam hingga berhari-hari) - komputer dengan kapasitas pemrosesan yang tinggi (biasanya menggunakan GPU) Salah satu kasus yang berkaitan dengan deepfake melibatkan Presiden Prabowo. Dittipidsiber Bareskrim Polri telah menangkap seorang pria bernama AMA berusia 29 tahun terkait dugaan pembuatan dan penyebaran video *deepfake* yang melibatkan nama Presiden Prabowo Subianto, Wakil Presiden Gibran Rakabuming Raka, dan Menteri Keuangan

Sri Mulyani. AMA menggunakan kecerdasan buatan untuk membuat video palsu dan menyebarkannya melalui berbagai platform media sosial. Video tersebut berisi informasi yang mengklaim memberikan dukungan pemerintah kepada masyarakat dan mencantumkan nomor WhatsApp untuk calon korban yang ingin menghubungi pelaku. Setelah menghubungi, korban diarahkan untuk mendaftar sebagai penerima bantuan. Namun, mereka diminta mengirimkan sejumlah uang untuk alasan "biaya administrasi. " Meskipun dijanjikan pencairan bantuan, uang itu tidak pernah diterima karena program bantuan yang dijanjikan tidak ada. AMA mengaku telah melakukan penipuan ini sejak tahun 2020 hingga 16 Januari 2025. Selama penyelidikan, ditemukan setidaknya 11 korban yang telah mentransfer uang kepada pelaku dengan jumlah antara Rp 250. 000 hingga Rp 1. 000. 000 per orang. Penyidik juga menemukan jaringan lain, termasuk seorang individu berinisial FA yang masuk dalam daftar pencarian orang (DPO). Secara hukum, AMA dikenakan Pasal 51 ayat (1) jo. Pasal 35 Undang-Undang Nomor 1 Tahun 2024 (perubahan kedua UU ITE) serta Pasal 378 Kitab Undang-Undang Hukum Pidana mengenai penipuan. Selain itu, dalam perkembangan kasus ini, aparat juga menangkap pelaku lain bernama JS berusia 25 tahun di Kabupaten Pringsewu, Lampung, pada 4 Februari 2025. JS mengelola akun Instagram @indoberbagi2025 yang dipakai untuk menyebarkan video deepfake yang sama dan menjaring korban. Analisis forensik digital terhadap video yang digunakan oleh JS menunjukkan bahwa kontennya sepenuhnya merupakan hasil manipulasi dengan teknologi *deepfake*. Meskipun teknologi ini memiliki potensi tujuan penggunaan yang bersifat positif dan negatif, pada kenyataannya *deepfake* kerap kali disalahgunakan untuk aktivitas perbuatan melawan hukum seperti : penipuan identitas, perusakan reputasi seseorang, pencemaran nama baik, dan penyebarluasan berita palsu atau hoaks (Afnan et al., 2022).

2. Penerapan Pasal-Pasal Yang Relevan Terhadap Penyalahgunaan *Artificial Intelligence* (AI) dalam Pembuatan Konten Hoaks dan *Deepfake* di Media Sosial

Pada hakikatnya terdapat beberapa aturan yaitu pasal-pasal berdasarkan hukum positif yang diterapkann di negara indonesia berkaitan dengan perbuatan tindak pidana terkhususnya tindak pidana *deepfake* yang relevan dan dapat diterapkan, diantaranya sebagai berikut:

1) UUD NRI Tahun 1945 pasal 28G ayat (1).
2) UU RI Nomor 12 Tahun 2005 mengatur tentang ratifikasi *International Covenant On Civil And Political Rights* Pasal 17.
3) UU RI Nomor 1 Tahun 2024 tentang perubahan kedua atas UU Nomor 11 Tahun 2008 mengatur tentang ITE Pasal 27A Pasal 28 (1).
4) UU RI Nomor 27 Tahun 2022 tentang perlindungan data pribadi diatur dalam Pasal 65 (3) Pasal 66, Pasal 67 (3), dan Pasal 68.

Meninjau berbagai instrumen hukum positif di tanah air, terlihat jelas bahwa regulasi yang tersedia belum menyentuh aspek tindak pidana *deepfake* secara menyeluruh. Padahal, jika bersandar pada konstitusi tertinggi yakni UUD NRI 1945, negara memiliki kewajiban utama untuk memproteksi seluruh rakyatnya tanpa terkecuali. Mandat tersebut tertuang lebih spesifik dalam Pasal 28G ayat (1), yang menegaskan bahwa privasi merupakan bagian tak terpisahkan dari hak asasi manusia, termasuk perlindungan data pribadi sebagai hak fundamental warga.

Sebagai bentuk implementasi Pasal 28G tersebut, hadirlah Undang-Undang Nomor 1 Tahun 2024 yang merombak aturan mengenai Informasi dan Transaksi Elektronik. Namun secara empiris, UU ITE tersebut masih didominasi oleh pengaturan terkait aktivitas niaga

digital. Hal ini menjadi celah krusial, mengingat fenomena *deepfake* adalah bentuk manipulasi konten yang mengandung informasi palsu. Kejahatan ini mengeksploitasi identitas individu secara melawan hukum demi menghancurkan martabat seseorang. Maka, penguatan payung hukum sangat mendesak demi menjamin keamanan digital masyarakat.

Apabila menilik Pasal 27A, terdapat aturan mengenai tindakan yang merugikan serta merusak martabat seseorang melalui pencemaran reputasi. Selain itu, Pasal 28 ayat (1) secara spesifik menjangkau penyebaran berita bohong atau fenomena hoaks di ruang digital. Selain instrumen tersebut, aturan pelaksana dari Pasal 28G ayat (1) UUD NRI 1945 yang sangat relevan dengan isu *deepfake* adalah Undang-Undang Nomor 27 Tahun 2022 mengenai Perlindungan Data Pribadi (UU PDP). Regulasi yang diresmikan sejak Oktober 2022 ini memuat poin-poin krusial, terutama pada Pasal 65 sampai Pasal 68.

Relevansi ini muncul karena aktivitas *deepfake* menjadikan data biometrik korban sebagai instrumen utama dalam melancarkan kejahatan. Merujuk Pasal 4 huruf b UU PDP, data biometrik mencakup identitas personal yang unik seperti rekaman suara, pemindaian wajah, hingga sidik jari. Realitas tersebut menunjukkan bahwa penyalahgunaan *deepfake* bukan sekadar penipuan elektronik atau penghinaan biasa, melainkan bentuk eksploitasi data pribadi secara ilegal demi tujuan kriminal. Pasal 65 ayat (3) secara tegas melarang penggunaan data milik orang lain, yang ancaman pidananya tertuang dalam Pasal 67 ayat (3). Sementara itu, Pasal 66 melarang fabrikasi data pribadi palsu yang berisiko mencelakai pihak lain, dengan sanksi hukuman yang diatur pada Pasal 68 (Noerman & Ibrahim 2024).

3. Tantangan Pembuktian Terhadap Penyalahgunaan *Artificial Intelligence* (AI) dalam Pembuatan Konten Hoaks dan *Deepfake* di Media Sosial

Dalam sistem hukum pidana Indonesia, setiap individu atau pribadi yang secara sengaja menyebarkan konten yang isinya memuat unsur penemaran nama baik seseorang dapat dimintakan sanksi pidana berdasarkan pada ketentuan dalam Pasal 310, 311 KUHP jo. Pasal 27 (3) Undang-Undang ITE. Dalam konteks ini, pengguna media *deepfake* yang terbukti melakukan perbuatan memanipulasi gambar atau suara milik seseorang untuk mencoreng nama baik serta martabat korban juga dapat dikenakan pasal tersebut. Meskipun demikian, tidak adanya pengaturan spesifik terkait dengan deepfake yang menyulitkan aparat penegak hukum dalam menyesuaikan unsur-unsur tindak pidana dengan realitas teknologinya medianya (Putri et al., 2024).

Perkembangan teknologi media *deepfake* menjadi bagian dari kecerdasan buatan yang menuntut adanya reformulasi aturan dalam hukum pidana di negara Indonesia. Kejahatan berbasis kecerdasan buatan termasuk pencemaran nama baik melalui *deepfake*, belum diatur sebagai tindak pidana khusus dalam hukum positif di negara indonesia. Reformulasi adalah *criminalization*, yaitu proses penetapan perbuatan sebagai tindak pidana. Dalam konteks *deepfake* aspek pembuktian dan yurisdiksi menjadi tantangan penting dalam reformulasi hukum pidana. Dalam kejahatan digital seperti *deepfake*, pembuktian tidak dapat hanya mengandalkan dari alat bukti konvensional, tetapi juga harus melibatkan keahlian forensik digital, pelacakan metadata, dan keterlibatan dari saksi ahli. Menurut Bambang Poernomo, pembuktian dalam hukum pidana harus menyesuaikan dengan sifat kejahatan yang sedang dihadapi agar tercapainya asas kepastian dan keadilan hukum. Selain itu, yurisdiksi dalam kejahatan siber diperluas karena pelaku bisa saja berada di luar wilayah hukum Indonesia, sehingga dibutuhkan adanya pembaharuan instrumen perjanjian internasional dan kerjasama lintas negara. Kendala utama dalam menghadapi media *deepfake* terletak pada kecanggihan manipulasi visualnya yang membuat pembedaan antara fakta asli dan konten fabrikasi

menjadi sangat sulit. Teknologi ini berisiko mencederai orisinalitas sebuah karya, sebab hasil rekayasanya mampu menampilkan realitas semu yang terlihat sangat meyakinkan bagi khalayak luas. Situasi tersebut kian mengkhawatirkan akibat distribusi disinformasi yang bergerak masif di jejaring media sosial, sehingga masyarakat awam kehilangan kemampuan untuk memverifikasi kebenaran informasi tersebut. Dampaknya, tingkat kepercayaan publik terhadap ekosistem informasi digital mengalami degradasi, yang pada akhirnya memicu normalisasi terhadap penyebaran berita palsu di ruang publik (Putri et al., 2024). Jika ditinjau dari perspektif hukum pidana nasional, aktivitas penyebaran *deepfake* dapat dikategorikan sebagai pelanggaran ITE. Dasar hukumnya merujuk pada Undang-Undang Nomor 19 Tahun 2016 yang merupakan revisi atas UU Nomor 11 Tahun 2008. Secara spesifik, Pasal 27 ayat (3) UU ITE menegaskan larangan bagi setiap individu untuk mendistribusikan, mentransmisikan, atau membuat dokumen elektronik yang bermuatan penghinaan serta pencemaran nama baik dapat diakses publik.

Ketentuan ini menjadi instrumen hukum untuk menjerat pelaku yang memanfaatkan *deepfake* demi merusak martabat seseorang, seperti melalui pembuatan video fitnah. Pelanggar pasal ini terancam hukuman penjara maksimal empat tahun atau denda mencapai tujuh ratus lima puluh juta rupiah. Meski demikian, penegakan hukum menghadapi batu sandungan besar dalam membuktikan unsur penghinaan tersebut. Hal ini dikarenakan *deepfake* sering kali dibalut dengan kemasan satire, humor, atau sindiran halus yang kerap dijadikan celah oleh pelaku untuk mengelak dari jeratan hukum. Di sisi lain, apabila konten hasil rekayasa digital tersebut mengandung unsur asusila, maka pelaku dapat dijatuhi sanksi berat berdasarkan Pasal 29 Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi (Prayoga & Tuasikal 2025). Sinkronisasi berbagai regulasi ini sangat krusial guna membendung dampak destruktif dari penyalahgunaan teknologi kecerdasan buatan di Indonesia.

4. Kategori dalam Penggunaan *Artificial Intelligence* pada *Deepfake* sebagai Salah Satu Tindak Pidana dalam *Cybercrime*

Dalam hal ini terdapat 2 (dua) serangan-serangan yang biasanya digunakan oleh para pelaku dalam melancarkan aksinya, diantaranya :

1) Serangan melalui *Generative Adversarial Networks* (GAN's)

Ian J. Goodfelow bersama dengan ketujuh rekannya pertama kali menemukan suatu penemuan tentang *Generative Adversarial Networks* (GAN's) pada tahun 2014 yang silam. Ian J. Goodfelow dan rekannya menuangkan penemuan mereka ke dalam sebuah jurnal yang bertajuk *Generative Adversarial Nets*, yang dalam penggunaannya menggunakan sebuah Algoritma *Deep Learning* konvensional lainnya, dari adanya perbedaan tersebut memanfaatkan secara optimal penggunaan data set terlatih untuk mengoptimalkan hasil dari akumulasi pengujian yang dilakukan, dengan semakin bertambah banyaknya data set maka semakin baik pula hasil akhirnya. Pada dasarnya keunikan dari GAN's *(Generative Adversarial Networks)* terletak pada pendekatannya yang berbeda dari konsep *Deep Learning* pada umumnya dikarenakan algoritma yang digunakan mencakup dua jaringan *neural artificial* yang bekerja secara bersamaan untuk menyelesaikan suatu permasalahan. Kedua jaringan tersebut terdiri dari : *Generator* yang bertugas untuk menghasilkan sampel data, serta Diskriminator yang berperan dalam mengidentifikasi dan memvalidasi keaslian dari sampel tersebut.

2) Serangan melalui Botnet Berbasis *Artificial Intelligence* (AI)

Botnet merupakan kumpulan perangkat komputer yang telah terindikasi mengandung virus berbahaya *(malware)* yang biasanya dioperasikan dan dikendalikan oleh seorang

botmaster. Beberapa jenis perangkat yang kemungkinan bisa terinfeksi oleh virus berbahaya *(malware)*, antara lain : kamera web, komputer, cctv, serta perangkat seluler yang sekiranya bisa dipantau, dioperasikan dan dikendalikan oleh seorang botmaster. Botnet sendiri memang sudah mempunyai target operasi pada sistem *windows* dengan menyebarkan *spam* yang berisi jaringan virus berbahaya pada *e-mail* dan *file* unduhan secara terus-menerus. Tentunya dalam hal ini mempunyai tujuan untuk menyerang target atau korbannya yang berupa pemerasan dengan dalih penyebaran konten hoaks melalui peretasan.

Kekuatan botnet bersumber dari kapasitas server yang mumpuni guna menghindar dari adanya deteksi sistem pengamanan serta melancarkan serangan masif dengan menggunakan teknik penyamaran paket data dan enkripsi. Botnet sendiri dapat melindungi diri dari adanya sistem perlindungan data dengan menunggu proses yang tidak sepenuhnya merata. Tentunya hal ini menjadi salah satu sarana bagi para pelaku dalam melancarkan aksi tindak pidana lewat peretasan akun (Patricia et al.,2025).

5. Pertanggungjawaban Pidana atas Kasus Penggunaan *Deepfake* terhadap Pelaku dalam Perspektif Hukum Pidana Indonesia

Berdasarkan sistem hukum pidana Indonesia, bagi setiap individu yang secara sadar dan dengan sengaja mendistribusikan konten yang bermuatan pencemaran nama baik dapat dikenakan sanksi yang berupa tanggung jawab pidana berdasarkan atas regulasi hukum yang tercantum dalam Pasal 310 dan 311 KUHP, serta Pasal 27 ayat (3) UU ITE. Dalam hal ini dijelaskan, bahwa pengguna teknologi yang menggunakan *deepfake* yang secara nyata dan terbukti memanipulasi gambar maupun suara seseorang untuk merusak citra kehormatan dan martabat korban dapat dikenakan pasal-pasal tersebut. Meskipun dalam implementasinya tidak secara eksplisit mengatur tentang konten *deepfake*, karena dalam hal ini menyulitkan pihak aparat penegak hukum dalam menyesuaikan apa yang menjadi unsur-unsur tindak pidana merujuk pada realitas teknologinya.

Dalam hukum pidana sendiri, konsep pertanggungjawaban pidana yang mensyaratkan adanya kedua unsur kesalahan, yaitu (*dolus* atau *culpa*) serta hubungan kausal antara tindakan pelaku serta apa akibat yang ditimbulkan. Pada contoh kasus konten *deepfake*, pelaku yang dengan sengaja membuat dan menyebarluaskan konten palsu dengan mencemarkan nama baik dapat diklasifikasikan sebagai pelaku tindak pidana dengan alasan kesengajaan *(dolus)*. Namun, dalam pembuktian niat jahat dan identitas pelaku memerlukan adanya dukungan forensik digital dan adanya regulasi yang mampu mengatur aspek teknis konten yang berbasis *Artificial Intelligence* atau AI.

Salah satu akibat dari penyebaran konten *deepfake* ini menyebabkan sulitnya aparat penegak hukum dalam mengidentifikasi identitas pelaku yang sesungguhnya, mengingat bahwa dalam peluncuran konten *deepfake* yang dapat berupa video atau foto yang dihasilkan sangat persis menyerupai realitanya, dalam hal ini pula semakin diperparah dengan adanya perputaran penyebaran konten digital yang dengan cepat tersebar luas pada media sosial. Sehingga menurut Ardiyani yang berpendapat bahwa belum adanya peraturan khusus yang mengatur dengan jelas dan spesifik terkait adanya *deepfake* yang menjadi tuntutan sebagai hambatan utama dalam dasar hukum yang berkekuatan hukum tetap, kuat, dan mengikat untuk memproses adanya pertanggungjawaban pelaku (Putu et al., 2024).

KESIMPULAN

Perkembangan teknologi Kecerdasan Buatan (AI) memberikan keuntungan yang signifikan, namun juga diperparah dengan adanya risiko penyalahgunaan yang serius,

terutama dalam menciptakan konten palsu dan *deepfake* di *platform* sosial media. Teknologi *deepfake*, yang mengandalkan data biometrik dan kemampuan mengedit digital, dapat menipu publik, merusak reputasi individu, serta menjadi alat untuk melakukan kejahatan seperti penipuan. Indonesia secara normatif telah memiliki berbagai peraturan hukum seperti KUHP, UU ITE, dan UU Perlindungan Data Pribadi. Akan tetapi, regulasi tersebut belum menangani secara khusus kejahatan yang menggunakan AI sebagai alat utama, yang menyebabkan kesulitan dalam pembuktian, pengenalan pelaku, dan penegakan hukum yang efektif. Kasus *deepfake* yang menimpa Presiden Prabowo Subianto serta pejabat negara lainnya menunjukkan bahwa penyalahgunaan AI telah berkembang ke tingkat yang dapat mengancam keamanan digital dan kepercayaan publik. Oleh sebab itu, dibutuhkan pembaruan regulasi yang lebih sensitif terhadap inovasi teknologi, pengembangan keterampilan forensik digital bagi aparat penegak hukum, serta peningkatan literasi digital masyarakat agar mampu mengenali dan memverifikasi informasi yang beredar. Langkah ini sangat penting untuk memastikan adanya perlindungan hukum yang efektif, mencegah kejahatan berbasis teknologi, dan melindungi integritas ruang digital di Indonesia.

REFERENSI

Afnan, H. A., Wardah Yuspin, & M. Kn. (2022). Perlindungan hukum penyalahgunaan artificial intelligence deepfake pada layanan pinjaman online (Disertasi, Universitas Muhammadiyah Surakarta).

Amelia, Y. F., Kaimuddin, A., & Ashsyarofi, H. L. (2024). Pertanggungjawaban pidana pelaku terhadap korban penyalahgunaan artificial intelligence deepfake menurut hukum positif Indonesia. Dinamika, 30(1), 9675–9691.

Banfatin, P. M., Medan, K. K., & Fallo, D. F. (2025). Pengaturan Hukum Pidana di Indonesia terhadap Penyalahgunaan Teknologi Artificial Intelligence Deepfake dalam melakukan Tindak Pidana Cybercrime. Pemuliaan Keadilan. https://doi.org/10.62383/pk.v2i1.402

Chatterjee, S., & Mohanta, A. (2024). Navigating AI Liability in Criminal Law. 311-334. https://doi.org/10.4018/979-8-3693-7580-8.ch014

Eurike Hailtik, A. G., & Afifah, W. (2024). Criminal responsibility of artificial intelligence committing deepfake crimes in Indonesia. Asian Journal of Social and Humanities, 2(1), 776–783. https://doi.org/10.59888/ajosh.v2i4.222

Kasita, I. D. (2022). Deepfake pornografi: Tren kekerasan gender berbasis online (KGBO) di era pandemi COVID-19. Jurnal Wanita dan Keluarga, 3(1), 16–26. https://doi.org/10.22146/jwk.5202

Laporan kasus penipuan video deepfake Presiden Prabowo dan Sri Mulyani (2025). Pemberitaan nasional dan klarifikasi aparat kepolisian. https://tribratanews.maluku.polri.go.id/index.php/informasi/berita/baca/dittipidsiber-tangkap-pelaku-deepfake-presiden-prabowo-serta-pejabat-negara-lainnya diakses 2 Desember 2025.

Long, D., & Magerko, B. (2020). Apa itu Literasi AI? Kompetensi dan Pertimbangan Desain. Konferensi tentang Faktor Manusia dalam Sistem Komputasi - Prosiding. https://doi.org/10.1145/3313831.3376727

Muslim Nugraha, Amanda Sela Sadina, Viraliza Ramadonna, & Keyssyah Aulia Hidayat. (2025). Analisis Unsur Perbuatan Melanggar Hukum atas Penggunaan Artificial Intelligence dalam Kasus Konten Deepfake. Legal System Journal, 1-12. https://doi.org/10.70656/lsj.v2i1.392

Noerman, C. T., & Ibrahim, A. L. (2024). Kriminalisasi deepfake di Indonesia sebagai bentuk pelindungan negara. Jurnal USM Law Review. https://doi.org/10.26623/julr.v7i2.8995

Prayoga, H., & Tuasikal, H. (2025). Penyebaran Konten Deepfake Sebagai Tindak Pidana: Analisis Kritis Terhadap Penegakan Hukum dan Perlindungan Publik di Indonesia. Abdurrauf Law and Sharia. https://doi.org/10.70742/arlash.v2i1.194

Putri, N. P. M., Hartono, M. S., & Yudiawan, I. D. G. H. (2024). Analisis Reformulasi Pertanggungjawaban Pidana Pengguna Teknologi Deepfake Dalam Tindak Pidana Pencemaran Nama Baik Berbasis Artificial Intelligence. Jurnal Pacta Sunt Servanda. https://doi.org/10.23887/jpss.v5i2.5807

Putri, S. M. I., Salsabila, N., & Hosnah, A. U. (2024). Kriminalisasi Penggunaan Deepfake Dalam Tindak Pidana Penipuan Dan Pencemaran Nama Baik: Tantangan dan Solusi Hukum. Jurnal Hukum Legalita. https://doi.org/10.47637/legalita.v6i2.1453

Pyndho Cevin Taraya, & Aji Wibawa. (2022). Mewujudkan Society 5.0 Melalui Pemanfaatan Teknologi Kecerdasan Buatan. Jurnal Inovasi Teknologi dan Edukasi Teknik, 386–398. https://doi.org/10.17977/um068v2i82022p378-385

Rustanta, A., Dwi Putranto, S., & Huang, P. (2025). Menjaga Ruang Publik Digital: Tantangan Etika Komunikasi dan Regulasi di Era TikTok. Jurnal Komunikasi, 17(1), 63–83. https://doi.org/10.24912/jk.v17i1.32927

Sari, A., & Harwika, D. (2022). Legal Liability of Artificial Intelligence in Perspective of Civil Law in Indonesia. International Journal of Social Science Research and Review, 5(2), 57-60. https://doi.org/10.47814/ijssrr.v5i2.191

Undang-Undang Nomor 44 Tahun 2008 yang mengatur tentang Pornografi, Undang-Undang Nomor 11 Tahun 2008 mengenai Informasi dan Transaksi Elektronik, Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Perlindungan Data Pribadi