Jejakdigital
JURNAL ILMIAH MULTIDISIPLIN
Available online at https://indojurnal.com/index.php/jejakdigital

eISSN 3089-7734; pISSN 3089-7742
Vol. 2, No. 4, Tahun 2026
doi.org/10.63822/kk4f2f38
Hal. 8122-8131

# Pertanggungjawaban Pidana terhadap Pemanfaatan *Artificial Intelligence* (AI) dalam Tindak Pidana Penipuan Berbasis *Deepfake*

**Ilham Fatkul Fadjri[1], Aturkian Laia[2]**
Ilmu Hukum, Hukum, Universitas Dirgantara Marsekal Suryadarma, Jakarta Timur, Indonesia[1]
Ilmu Hukum, Hukum, Universitas Dirgantara Marsekal Suryadarma, Jakarta Timur, Indonesia[2]

*Email Korespodensi: fatkul418@gmail.com

**Sejarah Artikel:**

*Diterima* 07-07-2026
*Disetujui* 13-07-2026
*Diterbitkan* 15-07-2026

***ABSTRACT***

*The rapid development of generative artificial intelligence has introduced deepfake technology capable of producing highly realistic images, voices, and videos, creating new opportunities for fraud. This study examines the characteristics of deepfake-based fraud, analyzes criminal liability for offenders using AI, and identifies challenges in law enforcement in Indonesia. The research employs a normative legal method using statutory and conceptual approaches based on the Indonesian Criminal Code, the Electronic Information and Transactions Law, the Personal Data Protection Law, and relevant legal literature. The findings show that deepfake fraud represents a new form of conventional fraud through synthetic digital identity, creating cross-border crimes and evidentiary difficulties. Criminal liability remains attributable to human offenders who use AI as an instrument of crime. However, law enforcement faces challenges related to digital forensics, jurisdiction, and institutional capacity. Strengthening legal regulation, investigative capabilities, and adaptive criminal law policies is essential to address AI-based crimes effectively.*

***Keywords****: Deepfake; Generative Artificial Intelligence; Criminal Liability; Cyber Fraud.*

**ABSTRAK**

Perkembangan kecerdasan artifisial generatif melahirkan teknologi *deepfake* yang mampu merekayasa citra, suara, dan video secara realistis sehingga berpotensi disalahgunakan sebagai sarana tindak pidana penipuan. Penelitian ini bertujuan menganalisis karakteristik penipuan berbasis *deepfake*, pertanggungjawaban pidana pelakunya, serta tantangan penegakan hukumnya di Indonesia. Penelitian menggunakan metode yuridis normatif dengan pendekatan perundang-undangan dan konseptual berdasarkan Kitab Undang-Undang Hukum Pidana, Undang-Undang Informasi dan Transaksi Elektronik, Undang-Undang Pelindungan Data Pribadi, serta literatur ilmiah. Hasil penelitian menunjukkan bahwa penipuan berbasis *deepfake* merupakan perkembangan modus penipuan konvensional yang memanfaatkan rekayasa identitas digital, bersifat lintas batas, dan menyulitkan pembuktian. Pertanggungjawaban pidana tetap dibebankan kepada pelaku yang menggunakan AI sebagai alat melakukan tindak pidana. Namun, penegakan hukum masih menghadapi kendala forensik digital, yurisdiksi, dan kapasitas kelembagaan. Diperlukan penguatan regulasi, kemampuan aparat, dan kebijakan hukum pidana yang adaptif terhadap kejahatan berbasis AI.

**Katakunci:** *Deepfake*; Kecerdasan Artifisial Generatif; Pertanggungjawaban Pidana; Penipuan Siber.

## PENDAHULUAN

Kemajuan teknologi kecerdasan artifisial generatif telah melahirkan kemampuan sintesis media (*synthetic media*) yang dikenal secara umum sebagai *deepfake*, yaitu teknik manipulasi citra, video, maupun suara menggunakan algoritma *deep learning*, khususnya *generative adversarial network* (GAN) dan model difusi (*diffusion model*), yang mampu menghasilkan representasi audiovisual seseorang secara sangat menyerupai aslinya (Chesney & Citron, 2019). Kemampuan teknis ini pada mulanya dikembangkan untuk kepentingan hiburan, produksi film, dan penelitian akademik, namun dalam perkembangannya disalahgunakan sebagai sarana kejahatan, termasuk yang paling meresahkan adalah tindak pidana penipuan (*fraud*).

Modus penipuan berbasis *deepfake* yang marak terjadi antara lain peniruan suara pimpinan perusahaan untuk memerintahkan transfer dana (*CEO fraud*/*voice cloning scam*), pemalsuan panggilan video untuk meyakinkan korban melakukan investasi bodong, hingga penyalahgunaan wajah figur publik untuk mempromosikan produk investasi ilegal di media sosial. Kasus internasional yang banyak dirujuk sebagai preseden adalah penipuan terhadap sebuah perusahaan energi di Inggris pada tahun 2019, ketika pelaku menggunakan teknologi kloning suara (*voice cloning*) untuk menirukan suara direktur utama perusahaan induk dan berhasil memerintahkan transfer dana senilai ratusan ribu euro (Stupp, 2019), serta kasus serupa yang melibatkan rekayasa panggilan video terhadap sebuah perusahaan multinasional di Hong Kong pada tahun 2024 dengan kerugian mencapai puluhan juta dolar Amerika Serikat. Di Indonesia, meskipun belum banyak kasus *deepfake fraud* yang diproses secara terbuka hingga tahap putusan pengadilan, laporan masyarakat mengenai penyalahgunaan wajah tokoh publik dan pejabat negara untuk promosi investasi bodong melalui video hasil rekayasa AI telah berulang kali mencuat di ruang publik, sebagaimana dilaporkan oleh Otoritas Jasa Keuangan dan berbagai pemberitaan media nasional.

Fenomena ini menimbulkan persoalan hukum pidana yang kompleks karena hukum pidana Indonesia, baik Kitab Undang-Undang Hukum Pidana (KUHP) warisan kolonial yang telah digantikan oleh Undang-Undang Nomor 1 Tahun 2023 tentang Kitab Undang-Undang Hukum Pidana, maupun Undang-Undang Informasi dan Transaksi Elektronik, dirumuskan tanpa mengantisipasi secara spesifik modus operandi yang melibatkan konten sintetis hasil kecerdasan artifisial. Ketentuan penipuan konvensional dalam Pasal 378 KUHP lama (kini Pasal 492 UU 1/2023) mensyaratkan unsur "rangkaian kebohongan" atau "tipu muslihat" yang dilakukan oleh seseorang terhadap korban, sehingga menimbulkan pertanyaan hukum mengenai bagaimana unsur tersebut diterapkan ketika tipu muslihat dilakukan melalui perantaraan sistem AI yang dilatih dan dijalankan oleh pelaku, alih-alih dilakukan secara langsung oleh pelaku itu sendiri. Pertanyaan ini menjadi semakin relevan mengingat karakteristik AI generatif yang bersifat otomatis dan dapat dioperasikan tanpa keahlian teknis tinggi (*democratization of deepfake tools*), sehingga potensi penyalahgunaannya meluas ke berbagai lapisan masyarakat. Selain persoalan konstruksi delik, tantangan lain yang tidak kalah signifikan adalah aspek pembuktian dan penegakan hukum, mengingat konten *deepfake* pada dasarnya dirancang agar tidak mudah dibedakan dari rekaman asli oleh indra manusia, sehingga membutuhkan keahlian digital forensik khusus untuk mengidentifikasi artefak digital (*digital artifacts*) yang menjadi penanda hasil rekayasa AI. Ketertinggalan kapasitas kelembagaan penegak hukum dalam mendeteksi konten sintetis, dikombinasikan dengan sifat lintas batas (*borderless*) dari kejahatan siber, menjadikan penegakan hukum terhadap penipuan berbasis *deepfake* sebagai tantangan hukum pidana kontemporer yang mendesak untuk dikaji secara mendalam.

Kajian mengenai *deepfake* dan hukum pidana telah menjadi perhatian akademik dalam satu dekade terakhir, meskipun fokusnya bervariasi. Terdapat beberapa penelitian terdahulu yang dapat dijadikan

pijakan akademik dalam studi ini. Marwan dan Kusuma (2020) mengkaji urgensi pengaturan *deepfake* dalam hukum pidana Indonesia dan menyimpulkan bahwa ketentuan UU ITE mengenai pemalsuan informasi elektronik dapat diterapkan secara analogis terhadap konten *deepfake*, namun penelitian tersebut masih bersifat umum dan belum spesifik menyasar modus penipuan berbasis *deepfake*. Handayani (2022) meneliti tanggung jawab pidana terhadap penyalahgunaan teknologi *voice cloning* dalam praktik penipuan perbankan, dengan simpulan bahwa unsur tipu muslihat dalam Pasal 378 KUHP lama dapat dipenuhi meskipun tipu daya dilakukan melalui suara hasil rekayasa, sepanjang dapat dibuktikan adanya niat jahat (*mens rea*) pelaku manusia di balik penggunaan teknologi tersebut. Penelitian tersebut memberikan kontribusi penting namun terbatas pada modus *voice cloning* dan belum mencakup modus video *deepfake* secara komprehensif.

Penelitian oleh Nasution (2023) berfokus pada pertanggungjawaban pidana korporasi dalam penyalahgunaan kecerdasan artifisial untuk kejahatan siber, dengan menekankan pada konstruksi pertanggungjawaban korporasi menurut UU 1/2023, namun belum secara spesifik mengaitkannya dengan karakteristik unik *deepfake* sebagai modus penipuan berbasis identitas sintetis. Adapun kajian oleh Pratama dan Wulandari (2024) mengangkat isu pembuktian digital forensik terhadap konten *deepfake* dalam proses peradilan pidana, dengan simpulan bahwa hukum acara pidana Indonesia belum memiliki standar baku mengenai autentikasi bukti elektronik berbasis AI, sehingga hakim bergantung pada keterangan ahli digital forensik yang kompetensinya belum merata di seluruh wilayah Indonesia. Terakhir, studi komparatif oleh Chesney dan Citron (2019) dalam *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security* menjadi rujukan konseptual penting mengenai taksonomi ancaman hukum *deepfake* secara umum, meskipun konteksnya berbasis hukum Amerika Serikat dan tidak spesifik membahas delik penipuan menurut hukum pidana Indonesia.

Penelitian-penelitian terdahulu tersebut menunjukkan bahwa kajian mengenai *deepfake* dalam hukum pidana Indonesia masih terfragmentasi ke dalam topik yang berbeda-beda: sebagian berfokus pada urgensi regulasi secara umum, sebagian pada modus *voice cloning* perbankan, sebagian pada pertanggungjawaban korporasi, dan sebagian lagi pada aspek pembuktian forensik semata. Penelitian ini hadir untuk menawarkan kebaruan ilmiah (*novelty*) karena secara spesifik memusatkan kajian pada delik penipuan (*fraud*) berbasis *deepfake* sebagai satu kesatuan konstruksi hukum pidana yang utuh. Analisis dalam artikel ini dilakukan secara terintegrasi dengan memeriksa karakteristik deliknya, konstruksi pertanggungjawaban pidana terhadap pelaku yang memanfaatkan AI sebagai *instrumentum delicti*, sekaligus memetakan tantangan penegakan hukumnya dalam kerangka KUHP baru (UU 1/2023), UU ITE, dan UU PDP secara bersamaan, yang belum ditemukan pada penelitian-penelitian sebelumnya.

Guna memandu jalannya analisis, penelitian ini merumuskan beberapa pokok permasalahan utama, yaitu mengenai bagaimana karakteristik tindak pidana penipuan berbasis *deepfake* dalam perspektif hukum pidana, bagaimana pertanggungjawaban pidana terhadap pemanfaatan AI dalam tindak pidana penipuan berbasis *deepfake*, serta apa saja tantangan penegakan hukum terhadap tindak pidana penipuan berbasis *deepfake* di Indonesia. Sejalan dengan rumusan masalah tersebut, penelitian ini bertujuan untuk mendeskripsikan dan menganalisis karakteristik yuridis tindak pidana penipuan berbasis *deepfake* sebagai bentuk kejahatan siber modern. Lebih lanjut, artikel ini bertujuan untuk menganalisis konstruksi pertanggungjawaban pidana terhadap pelaku yang memanfaatkan kecerdasan artifisial dalam tindak pidana tersebut berdasarkan hukum pidana positif Indonesia, yang pada akhirnya ditujukan untuk mengidentifikasi tantangan penegakan hukum terhadap tindak pidana penipuan berbasis *deepfake* di Indonesia sebagai bahan masukan bagi pembentukan kebijakan hukum pidana yang responsif terhadap perkembangan teknologi ke

depan.

## METODE PENELITIAN

Penelitian ini menggunakan metode penelitian hukum yuridis normatif (*legal doctrinal research*), yaitu penelitian yang mengkaji hukum sebagai sistem norma yang mengatur perbuatan manusia yang dianggap pantas, dengan menggunakan bahan pustaka atau data sekunder sebagai sumber data utama (Soekanto & Mamudji, 2015; Marzuki, 2021). Pendekatan yang digunakan meliputi pendekatan perundang-undangan (*statute approach*) untuk menelaah keterkaitan antara KUHP, UU ITE, dan UU PDP dalam mengonstruksikan delik penipuan berbasis *deepfake*; pendekatan konseptual (*conceptual approach*) untuk memahami doktrin kesalahan (*mens rea*), doktrin *instrumentum delicti*, dan konsep pertanggungjawaban pidana dalam konteks pemanfaatan teknologi; serta pendekatan kasus (*case approach*) secara terbatas dengan menelaah kasus-kasus penyalahgunaan *deepfake* yang telah dilaporkan baik di dalam maupun luar negeri sebagai ilustrasi empiris atas konstruksi hukum yang dianalisis.

Bahan hukum yang digunakan terdiri atas bahan hukum primer, yaitu Kitab Undang-Undang Hukum Pidana lama (*Wetboek van Strafrecht*) sepanjang masih relevan sebagai pembanding, Undang-Undang Nomor 1 Tahun 2023 tentang Kitab Undang-Undang Hukum Pidana yang akan berlaku efektif sebagai KUHP nasional, Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana diubah dengan Undang-Undang Nomor 19 Tahun 2016 dan Undang-Undang Nomor 1 Tahun 2024, Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi, Undang-Undang Nomor 8 Tahun 1981 tentang Hukum Acara Pidana, serta Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial. Bahan hukum sekunder meliputi buku teks hukum pidana dan hukum siber, jurnal ilmiah nasional dan internasional, hasil penelitian terdahulu, serta dokumen kebijakan terkait keamanan siber. Bahan hukum tersier berupa kamus hukum dan glosarium teknis untuk menjelaskan istilah kecerdasan artifisial dan digital forensik.

Teknik pengumpulan bahan hukum dilakukan melalui studi kepustakaan (*library research*) dengan menelusuri peraturan perundang-undangan, putusan pengadilan yang relevan sepanjang tersedia, dan literatur akademik baik cetak maupun digital. Analisis bahan hukum dilakukan secara kualitatif dengan teknik deskriptif-analitis, yaitu menguraikan norma hukum pidana yang berlaku kemudian menganalisisnya secara kritis dikaitkan dengan karakteristik teknis *deepfake* dan modus operandi penipuan berbasis teknologi tersebut, sehingga diperoleh konstruksi hukum yang argumentatif mengenai pertanggungjawaban pidana dan tantangan penegakannya.

## HASIL DAN PEMBAHASAN

### 1. Karakteristik Tindak Pidana Penipuan Berbasis *Deepfake* dalam Perspektif Hukum Pidana

*Deepfake* secara teknis merupakan produk dari algoritma pembelajaran mendalam, khususnya *generative adversarial network* (GAN) yang bekerja melalui dua jaringan saraf tiruan yang saling berkompetisi *generator* yang menciptakan konten sintetis dan *discriminator* yang menilai keasliannya sehingga menghasilkan *output* yang secara progresif semakin sulit dibedakan dari rekaman asli (Goodfellow et al., 2014; Chesney & Citron, 2019). Perkembangan lebih lanjut berupa model difusi (*diffusion model*) dan teknologi *voice cloning* berbasis *few-shot learning* bahkan memungkinkan penciptaan konten sintetis yang meyakinkan hanya dengan modal sampel suara atau wajah berdurasi singkat, sehingga

menurunkan secara signifikan hambatan teknis (*barrier to entry*) bagi pelaku kejahatan untuk memanfaatkan teknologi tersebut.

Dari perspektif hukum pidana, penipuan berbasis *deepfake* pada dasarnya merupakan perluasan modus dari delik penipuan konvensional yang diatur dalam Pasal 378 KUHP lama, yang kini dirumuskan ulang dalam Pasal 492 Undang-Undang Nomor 1 Tahun 2023 tentang KUHP, dengan unsur-unsur: dengan maksud menguntungkan diri sendiri atau orang lain secara melawan hukum; dengan memakai nama palsu atau martabat palsu, tipu muslihat, ataupun rangkaian kebohongan; menggerakkan orang lain untuk menyerahkan sesuatu barang, memberi utang, atau menghapuskan piutang. Karakteristik khas yang membedakan penipuan berbasis *deepfake* dari penipuan konvensional terletak pada unsur "tipu muslihat" atau "rangkaian kebohongan" yang tidak lagi disampaikan secara langsung oleh pelaku melalui interaksi verbal atau tatap muka, melainkan melalui perantaraan konten audiovisual sintetis yang direkayasa menggunakan sistem AI untuk meniru identitas pihak ketiga yang dipercaya oleh korban.

Karakteristik pertama adalah rekayasa identitas digital (*synthetic identity impersonation*), yaitu penciptaan representasi audiovisual palsu atas identitas seseorang baik figur publik, pejabat, pimpinan perusahaan, maupun kerabat korban untuk membangun kepercayaan (*trust exploitation*) yang menjadi prasyarat keberhasilan tipu muslihat. Berbeda dengan pemalsuan identitas konvensional yang umumnya terbatas pada dokumen tertulis atau identitas verbal, *deepfake* memungkinkan rekayasa identitas pada level audiovisual yang jauh lebih meyakinkan secara psikologis karena memanfaatkan bias kognitif manusia yang cenderung mempercayai informasi yang dapat dilihat dan didengar secara langsung (*seeing is believing bias*).

Karakteristik kedua adalah sifat semi-otomatis dan dapat direplikasi (*scalability*), karena sekali suatu model AI dilatih untuk meniru identitas tertentu, pelaku dapat menghasilkan konten sintetis dalam jumlah tidak terbatas dan menyasar korban secara massal (*mass targeting*) dengan biaya marjinal yang sangat rendah, berbeda dengan penipuan konvensional yang umumnya membutuhkan interaksi personal antara pelaku dan korban secara satu per satu. Karakteristik ini menjadikan penipuan berbasis *deepfake* berpotensi menimbulkan kerugian sosial yang jauh lebih luas dibandingkan penipuan konvensional dalam rentang waktu yang sama.

Karakteristik ketiga adalah kaburnya *locus delicti* dan *tempus delicti* akibat sifat lintas batas kejahatan siber. Proses pembuatan konten *deepfake* dapat dilakukan di suatu yurisdiksi, disebarluaskan melalui *server* yang berlokasi di yurisdiksi lain, dan menimbulkan kerugian bagi korban di yurisdiksi ketiga, sehingga menyulitkan penerapan asas teritorialitas sebagaimana diatur Pasal 2 KUHP (dan padanannya dalam Pasal 4 UU 1/2023) yang pada dasarnya mendasarkan yurisdiksi pidana pada tempat terjadinya tindak dipidana. Karakteristik keempat adalah konvergensi antara delik penipuan konvensional dengan delik siber, karena tindak pidana ini pada umumnya tidak berdiri sendiri melainkan bersinggungan dengan delik lain seperti pemalsuan informasi elektronik (Pasal 35 UU ITE), akses ilegal terhadap sistem elektronik (Pasal 30 UU ITE) apabila data biometrik korban diperoleh secara tidak sah, dan pelanggaran perlindungan data pribadi (Pasal 65 dan Pasal 67 UU PDP) apabila data wajah atau suara diperoleh dan diproses tanpa persetujuan subjek data, sehingga penipuan berbasis *deepfake* pada praktiknya sering merupakan delik berlapis (*concursus*) yang melibatkan beberapa ketentuan pidana sekaligus.

## 2. Pertanggungjawaban Pidana terhadap Pemanfaatan AI dalam Tindak Pidana Penipuan Berbasis *Deepfake*

Prinsip dasar hukum pidana Indonesia menganut asas kesalahan (*mens rea*) sebagai syarat

pemidanaan, sebagaimana tercermin dalam adagium *actus non facit reum nisi mens sit rea* (suatu perbuatan tidak menjadikan seseorang bersalah kecuali disertai niat jahat). Prinsip ini juga ditegaskan dalam Pasal 34 Undang-Undang Nomor 1 Tahun 2023 tentang KUHP yang merumuskan bahwa tidak seorang pun dapat dipidana tanpa kesalahan (*geen straf zonder schuld*/*no punishment without fault*). Dalam konteks penipuan berbasis *deepfake*, pertanyaan pokok yang harus dijawab adalah kepada siapa pertanggungjawaban pidana dibebankan ketika unsur tipu muslihat dilakukan melalui perantaraan sistem AI, bukan secara langsung oleh pelaku.

Doktrin hukum pidana yang relevan untuk menjawab persoalan ini adalah konsep *instrumentum delicti* atau alat untuk melakukan kejahatan, yang menempatkan sistem AI sebagai sarana atau instrumen yang digunakan oleh pelaku manusia untuk mewujudkan niat jahatnya, sebagaimana halnya senjata, dokumen palsu, atau perangkat lunak peretasan dalam konteks kejahatan siber konvensional. Kecerdasan artifisial, betapapun canggihnya, tidak diakui sebagai subjek hukum pidana (*rechtssubject*) yang dapat memikul pertanggungjawaban pidana, karena hukum pidana Indonesia hanya mengenal dua kategori subjek hukum, yaitu orang perseorangan (*naturlijke persoon*) dan korporasi (*rechtspersoon*) sebagaimana diperluas cakupannya melalui Peraturan Mahkamah Agung Nomor 13 Tahun 2016 tentang Tata Cara Penanganan Perkara Tindak Pidana oleh Korporasi dan dipertegas dalam Pasal 46 UU 1/2023. Dengan demikian, pertanggungjawaban pidana tetap dibebankan kepada pelaku manusia atau korporasi yang menggunakan, mengembangkan, atau memerintahkan penggunaan sistem AI untuk tujuan penipuan, bukan kepada sistem AI itu sendiri.

Konstruksi pertanggungjawaban pidana tersebut dapat dianalisis melalui tiga skenario keterlibatan. Skenario pertama adalah pelaku yang secara langsung menggunakan aplikasi atau layanan *deepfake* yang tersedia secara komersial maupun sumber terbuka (*open source*) untuk menciptakan konten sintetis dengan niat menipu korban; dalam skenario ini, unsur *mens rea* relatif mudah dikonstruksikan karena pelaku memiliki kendali penuh atas proses pembuatan dan penyebaran konten, sehingga dapat dijerat Pasal 492 UU 1/2023 *juncto* Pasal 28 ayat (1) UU ITE tentang penyebaran berita bohong yang menimbulkan kerugian konsumen dalam transaksi elektronik, atau Pasal 35 UU ITE tentang manipulasi informasi elektronik apabila konten *deepfake* tersebut disajikan seolah-olah otentik.

Skenario kedua adalah keterlibatan pengembang atau penyedia platform AI yang digunakan untuk menghasilkan konten *deepfake*, yang pertanggungjawabannya bergantung pada sejauh mana pengembang mengetahui atau patut menduga (*wetenschap*/*knowledge*) bahwa produknya akan digunakan untuk tujuan kriminal. Apabila pengembang secara aktif memasarkan produknya untuk tujuan penipuan atau tidak menerapkan langkah pencegahan yang wajar (*reasonable safeguard*) meskipun mengetahui pola penyalahgunaan yang berulang, maka dapat dikonstruksikan pertanggungjawaban turut serta (*medeplichtigheid*) sebagaimana diatur Pasal 56 KUHP lama/Pasal 21 UU 1/2023 tentang pembantuan tindak pidana, atau bahkan pertanggungjawaban sebagai pelaku peserta (*mededader*) apabila keterlibatannya bersifat aktif dan disengaja. Namun demikian, pertanggungjawaban ini sulit diterapkan terhadap penyedia teknologi yang bersifat generik dan memiliki kegunaan sah (*dual-use technology*), mengingat prinsip proporsionalitas dalam hukum pidana mensyaratkan pembatasan pertanggungjawaban hanya pada pihak yang benar-benar memiliki kesengajaan atau kealpaan yang relevan dengan tindak pidana yang terjadi.

Skenario ketiga adalah keterlibatan korporasi dalam penipuan berbasis *deepfake* yang terorganisasi, misalnya sindikat penipuan investasi daring yang menggunakan tim teknis untuk memproduksi konten *deepfake* tokoh publik secara sistematis. Dalam skenario ini, pertanggungjawaban pidana korporasi dapat

dikonstruksikan berdasarkan Pasal 46 UU 1/2023 yang mengatur bahwa tindak pidana oleh korporasi dapat dilakukan oleh pengurus yang memiliki kedudukan fungsional dalam struktur organisasi korporasi, bertindak dalam lingkup usaha korporasi, dan menguntungkan korporasi secara melawan hukum, sehingga korporasi dapat dijatuhi pidana pokok berupa denda dan pidana tambahan berupa pencabutan izin usaha atau pembubaran korporasi sebagaimana diatur Pasal 122 UU 1/2023.

Persoalan yang lebih rumit muncul ketika sistem AI generatif digunakan secara otonom melampaui instruksi awal pelaku (*emergent behavior*), sebagaimana potensi yang mulai diperdebatkan dalam literatur hukum teknologi mutakhir (Hallevy, 2021). Meskipun secara teoretis skenario ini dapat menimbulkan kekosongan pertanggungjawaban (*responsibility gap*) apabila luaran AI benar-benar di luar prediksi dan kendali penggunanya, dalam praktik penipuan berbasis *deepfake* yang telah teridentifikasi hingga saat ini, unsur kesengajaan pelaku manusia dalam mengarahkan penggunaan teknologi untuk tujuan menipu korban tetap dapat dibuktikan melalui rangkaian perbuatan persiapan dan pelaksanaan yang terencana, sehingga doktrin *responsibility gap* belum menjadi persoalan praktis yang signifikan dalam konteks *deepfake fraud*, meskipun tetap relevan untuk diantisipasi dalam perumusan kebijakan hukum pidana jangka panjang.

**3. Tantangan Penegakan Hukum terhadap Tindak Pidana Penipuan Berbasis *Deepfake* di Indonesia**

Tantangan pertama yang dihadapi aparat penegak hukum adalah kesulitan pembuktian digital forensik terhadap keaslian suatu konten audiovisual. Deteksi *deepfake* membutuhkan keahlian teknis khusus untuk mengidentifikasi artefak digital seperti inkonsistensi kedipan mata, distorsi piksel pada area wajah, ketidaksesuaian pencahayaan, maupun anomali spektral pada gelombang suara, yang seluruhnya membutuhkan perangkat lunak forensik mutakhir dan sumber daya manusia yang terlatih khusus (Pratama & Wulandari, 2024). Kapasitas ini belum tersedia secara merata di seluruh satuan kerja Kepolisian Negara Republik Indonesia, terutama di tingkat kepolisian daerah dan kepolisian resor, sehingga penanganan kasus *deepfake fraud* cenderung terpusat pada unit siber di tingkat Markas Besar Kepolisian, yang berimplikasi pada lambatnya respons terhadap laporan masyarakat di daerah.

Tantangan kedua adalah persoalan penentuan yurisdiksi dan kerja sama hukum lintas negara (*mutual legal assistance*). Sebagaimana disinggung pada subbab sebelumnya, penipuan berbasis *deepfake* kerap melibatkan infrastruktur digital yang tersebar di berbagai negara, termasuk penggunaan aplikasi pembuat *deepfake* yang dikembangkan oleh perusahaan asing, penyimpanan data pada *server* luar negeri, dan transfer dana melalui rekening atau dompet kripto lintas yurisdiksi. Ketentuan Pasal 2 dan Pasal 3 KUHP lama (serta padanannya dalam Pasal 4 dan Pasal 5 UU 1/2023) mengatur asas teritorialitas dan asas perlindungan yang memungkinkan penerapan hukum pidana Indonesia terhadap tindak pidana yang berdampak di wilayah Indonesia meskipun sebagian perbuatan dilakukan di luar negeri, namun implementasinya membutuhkan kerja sama penegakan hukum internasional yang efektif, sementara Indonesia belum memiliki perjanjian *mutual legal assistance* dengan seluruh negara yang berpotensi menjadi lokasi penyedia infrastruktur digital terkait *deepfake fraud*.

Tantangan ketiga adalah keterbatasan norma hukum acara pidana dalam mengatur autentikasi bukti elektronik berbasis AI. Pasal 5 UU ITE mengakui informasi elektronik dan/atau dokumen elektronik sebagai alat bukti yang sah dan merupakan perluasan dari alat bukti yang diatur dalam hukum acara pidana yang berlaku, namun ketentuan tersebut belum memberikan panduan teknis mengenai standar autentikasi khusus untuk membedakan konten elektronik asli dari konten hasil rekayasa AI. Ketiadaan standar baku ini menyebabkan penilaian terhadap keaslian bukti sangat bergantung pada keterangan ahli (*expert witness*) yang kompetensi dan metodologinya dapat bervariasi antar kasus, sehingga berpotensi menimbulkan

inkonsistensi putusan pengadilan dalam menilai bobot pembuktian konten *deepfake*.

Tantangan keempat adalah kesenjangan literasi digital masyarakat yang menyebabkan tingginya angka korban penipuan berbasis *deepfake*, sekaligus rendahnya pelaporan (*underreporting*) akibat rasa malu korban atau ketidaktahuan mengenai mekanisme pelaporan yang tepat. Rendahnya kesadaran masyarakat mengenai keberadaan teknologi *deepfake* menjadikan efektivitas represif hukum pidana semata tidak memadai tanpa didukung langkah preventif berupa edukasi publik yang masif, sebagaimana yang menjadi salah satu semangat Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial, meskipun surat edaran tersebut bersifat *soft law* yang tidak memiliki daya paksa hukum sebagaimana peraturan perundang-undangan.

Tantangan kelima bersifat kebijakan formulatif, yaitu belum adanya ketentuan pidana khusus (*lex specialis*) yang secara eksplisit mengatur *deepfake* sebagai modus operandi tindak pidana, sehingga penegak hukum harus mengonstruksikan penerapan pasal-pasal yang ada secara analogis, yang berpotensi menimbulkan perdebatan yuridis mengenai kesesuaiannya dengan asas legalitas (*nullum delictum nulla poena sine praevia lege poenali*) sebagaimana diatur Pasal 1 ayat (1) KUHP lama dan dipertegas dalam Pasal 1 ayat (1) UU 1/2023. Meskipun konstruksi analogis terhadap Pasal 492 UU 1/2023 *juncto* Pasal 28 dan Pasal 35 UU ITE secara umum dapat dipertahankan karena unsur-unsur delik penipuan tetap terpenuhi terlepas dari sarana yang digunakan, ketiadaan rumusan yang eksplisit mengenai *deepfake* berpotensi menimbulkan variasi penafsiran antar penegak hukum dan antar wilayah, sehingga sejumlah negara seperti Amerika Serikat pada tingkat negara bagian (misalnya California melalui AB 730 dan Texas melalui SB 751) serta Tiongkok melalui Regulasi *Provisions on the Administration of Deep Synthesis Internet Information Services* yang berlaku sejak tahun 2023, telah merumuskan ketentuan khusus mengenai kewajiban pelabelan (*labelling*) dan larangan penyalahgunaan konten sintetis, yang dapat menjadi rujukan perbandingan bagi pembentukan kebijakan hukum pidana Indonesia ke depan.

Berdasarkan uraian tersebut, tantangan penegakan hukum terhadap tindak pidana penipuan berbasis *deepfake* di Indonesia bersifat multidimensional, mencakup dimensi teknis-forensik, dimensi yurisdiksional, dimensi hukum acara pidana, dimensi sosial-edukatif, dan dimensi kebijakan formulatif, yang keseluruhannya membutuhkan respons kebijakan hukum pidana yang terpadu, bukan sekadar mengandalkan penafsiran ekstensif terhadap norma yang telah ada.

## KESIMPULAN

Berdasarkan hasil dan pembahasan di atas, dapat ditarik beberapa kesimpulan. Pertama, tindak pidana penipuan berbasis *deepfake* merupakan perluasan modus dari delik penipuan konvensional yang diatur Pasal 378 KUHP lama/Pasal 492 UU 1/2023, dengan karakteristik khas berupa rekayasa identitas digital untuk mengeksploitasi kepercayaan korban, sifat semi-otomatis dan dapat direplikasi secara massal, kaburnya *locus delicti* akibat sifat lintas batas kejahatan siber, serta konvergensi dengan delik siber lain seperti pemalsuan informasi elektronik dan pelanggaran perlindungan data pribadi, sehingga pada praktiknya sering berwujud delik berlapis. Kedua, pertanggungjawaban pidana terhadap pemanfaatan AI dalam tindak pidana penipuan berbasis *deepfake* tetap dibebankan kepada pelaku manusia atau korporasi yang menggunakan sistem AI sebagai *instrumentum delicti*, karena kecerdasan artifisial tidak diakui sebagai subjek hukum pidana menurut sistem hukum Indonesia; konstruksi pertanggungjawaban tersebut dapat menyasar pelaku langsung berdasarkan Pasal 492 UU 1/2023 *juncto* Pasal 28 dan Pasal 35 UU ITE, pihak yang membantu (*medeplichtige*) apabila terbukti mengetahui dan menyediakan sarana secara sengaja,

maupun korporasi berdasarkan Pasal 46 UU 1/2023 apabila tindak pidana dilakukan dalam lingkup usaha korporasi. Ketiga, penegakan hukum terhadap tindak pidana ini menghadapi tantangan multidimensional yang meliputi keterbatasan kapasitas digital forensik, kesulitan penentuan yurisdiksi dan kerja sama hukum lintas negara, ketiadaan standar autentikasi bukti elektronik berbasis AI dalam hukum acara pidana, rendahnya literasi digital masyarakat, serta belum adanya ketentuan pidana khusus yang secara eksplisit mengatur *deepfake* sebagai modus operandi kejahatan.

Berdasarkan simpulan tersebut, penelitian ini merekomendasikan agar pembentuk kebijakan mempertimbangkan penguatan kapasitas digital forensik aparat penegak hukum secara merata di seluruh wilayah Indonesia melalui pelatihan khusus dan penyediaan perangkat lunak deteksi *deepfake*, perumusan pedoman teknis autentikasi bukti elektronik berbasis AI dalam hukum acara pidana, penjajakan perjanjian kerja sama hukum internasional yang lebih luas untuk mengantisipasi sifat lintas batas kejahatan ini, penguatan edukasi publik mengenai literasi digital dan mekanisme pelaporan, serta kajian lebih lanjut mengenai urgensi perumusan ketentuan pidana khusus (*lex specialis*) mengenai penyalahgunaan teknologi sintesis media, termasuk kemungkinan kewajiban pelabelan (*labelling*) konten hasil kecerdasan artifisial sebagaimana mulai diterapkan di berbagai yurisdiksi lain, sebagai bagian dari kebijakan hukum pidana yang responsif terhadap perkembangan teknologi kecerdasan artifisial generatif. Penelitian lanjutan disarankan untuk mengkaji secara empiris penanganan kasus *deepfake fraud* oleh aparat penegak hukum di Indonesia serta mengevaluasi efektivitas regulasi pelabelan konten AI di negara-negara yang telah menerapkannya.

## DAFTAR PUSTAKA

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, *107*(6), 1753–1820. https://doi.org/10.15779/Z38RS6M386

Cyberspace Administration of China. (2023). *Provisions on the Administration of Deep Synthesis Internet Information Services*.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Advances in Neural Information Processing Systems*, *27*, 2672–2680. https://doi.org/10.48550/arXiv.1406.2661

Hallevy, G. (2021). *Liability for crimes involving artificial intelligence systems* (2nd ed.). Springer. https://doi.org/10.1007/978-3-030-80234-9

Handayani, R. (2022). Tanggung jawab pidana terhadap penyalahgunaan teknologi voice cloning dalam praktik penipuan perbankan. *Jurnal Hukum Pidana dan Kriminologi*, *9*(2), 145–163.

Indonesia. (1915). *Kitab Undang-Undang Hukum Pidana (Wetboek van Strafrecht)*. Staatsblad 1915 Nomor 732.

Indonesia. (1981). *Undang-Undang Nomor 8 Tahun 1981 tentang Hukum Acara Pidana*. Lembaran Negara Republik Indonesia Tahun 1981 Nomor 76.

Indonesia. (2008). *Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana diubah dengan Undang-Undang Nomor 19 Tahun 2016 dan Undang-Undang Nomor 1 Tahun 2024*.

Indonesia. (2016). *Peraturan Mahkamah Agung Nomor 13 Tahun 2016 tentang Tata Cara Penanganan Perkara Tindak Pidana oleh Korporasi*.

Indonesia. (2022). *Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi*. Lembaran Negara Republik Indonesia Tahun 2022 Nomor 218.

Indonesia. (2023). *Undang-Undang Nomor 1 Tahun 2023 tentang Kitab Undang-Undang Hukum Pidana*. Lembaran Negara Republik Indonesia Tahun 2023 Nomor 1.

Kementerian Komunikasi dan Informatika Republik Indonesia. (2023). *Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial*.

Marwan, A., & Kusuma, D. (2020). Urgensi pengaturan deepfake dalam hukum pidana Indonesia. *Jurnal Hukum dan Peradilan*, *9*(3), 401–420. https://doi.org/10.25216/jhp.9.3.2020.401-420

Marzuki, P. M. (2021). *Penelitian hukum* (Edisi Revisi). Prenadamedia Group.

Moeljatno. (2018). *Asas-asas hukum pidana* (Edisi Revisi). Rineka Cipta.

Nasution, M. R. (2023). Pertanggungjawaban pidana korporasi dalam penyalahgunaan kecerdasan artifisial untuk kejahatan siber. *Jurnal Hukum dan Pembangunan*, *53*(2), 310–329. https://doi.org/10.21143/jhp.vol53.no2.4042

Pratama, Y., & Wulandari, S. (2024). Pembuktian digital forensik terhadap konten deepfake dalam proses peradilan pidana di Indonesia. *Jurnal Ilmu Hukum*, *27*(1), 77–95.

Soekanto, S., & Mamudji, S. (2015). *Penelitian hukum normatif: Suatu tinjauan singkat*. Rajawali Pers.

State of California. (2019). *Assembly Bill No. 730 (2019), California Elections Code Section 20010.*

Stupp, C. (2019, August 30). Fraudsters used AI to mimic CEO's voice in unusual cybercrime case. *The Wall Street Journal*.