REKONSTRUKSI PERTANGGUNGJAWABAN PIDANA ATAS KEJAHATAN *DEEPFAKE* BERBASIS KECERDASAN BUATAN GENERATIF DALAM HUKUM PIDANA INDONESIA.

Oleh: **Ruth Grace Meriana**
Universitas Negeri Jakarta, Jl. Rawamangun Muka, RT.11/RW.14, Rawamangu, Kec. Pulo Gadung, Kota Jakarta Timur
email: *gracegrabiell77@gmail.com

Informasi Artikel *Received*:	08-08-2026	*Accepted*:	25-10-2026

ABSTRAK

Perkembangan kecerdasan buatan generatif telah melahirkan teknologi *deepfake* yang mampu memanipulasi wajah, suara, dan ekspresi manusia secara sangat meyakinkan. Teknologi tersebut dapat dimanfaatkan untuk pencemaran nama baik, penipuan, kekerasan berbasis gender siber, serta disinformasi politik. Persoalan hukum muncul karena pertanggungjawaban pidana Indonesia pada dasarnya berorientasi pada kesalahan subjek hukum manusia, sedangkan pembuatan dan penyebaran *deepfake* dapat melibatkan pengguna, pengembang, dan penyedia platform melalui sistem yang memiliki tingkat otonomi tertentu. Penelitian ini bertujuan menganalisis kedudukan hukum kejahatan *deepfake* berbasis kecerdasan buatan generatif dalam hukum pidana Indonesia serta merumuskan model pertanggungjawaban pidana yang dapat membagi beban pertanggungjawaban berdasarkan peran setiap aktor. Penelitian menggunakan metode yuridis normatif dengan pendekatan perundang-undangan, konseptual, dan perbandingan hukum terhadap regulasi kecerdasan buatan di Uni Eropa dan Tiongkok. Hasil penelitian menunjukkan bahwa ketentuan pidana yang ada belum sepenuhnya memberikan mekanisme atribusi yang memadai terhadap hubungan antara tindakan pengguna, desain sistem, penyediaan layanan, dan luaran sintetis. Oleh karena itu, diperlukan rekonstruksi pertanggungjawaban pidana berbasis peran yang mencakup pertanggungjawaban berdasarkan desain, penyediaan, dan penggunaan sistem, disertai kewajiban pelabelan konten serta penandaan digital untuk mendukung pembuktian forensik elektronik. Model tersebut diarahkan untuk memperkuat kepastian hukum, perlindungan korban, dan akuntabilitas tanpa menghambat pengembangan teknologi.

Kata kunci: *deepfake*; kecerdasan buatan generatif; pertanggungjawaban pidana; kesalahan; regulasi kecerdasan buatan

ABSTRACT

The development of generative artificial intelligence has produced deepfake technology capable of manipulating human faces, voices, and expressions with a high degree of realism. The technology may be used for defamation, fraud, cyber gender-based violence, and political disinformation. A legal problem arises because Indonesian criminal liability is primarily oriented toward the fault of human legal subjects, while the creation and dissemination of deepfakes may involve users, developers, and platform providers through systems with a certain degree of autonomy. This study aims to analyze the legal position of generative artificial intelligence-based deepfake crimes under Indonesian criminal law and to formulate a criminal liability model that allocates responsibility according to the role of each actor. This study employs normative legal research using statutory, conceptual, and comparative approaches to artificial intelligence regulation in the European Union and China. The findings indicate that existing criminal law provisions do not yet provide a fully adequate mechanism for attributing liability across user conduct, system design, service provision, and synthetic outputs. Therefore, a role-based reconstruction of criminal liability is required, covering liability by design, liability by deployment, and liability by use, accompanied by mandatory content labeling and digital watermarking to support electronic forensic evidence. This model is intended to strengthen legal certainty, victim protection, and accountability without obstructing technological development.

Keywords: *deepfake; generative artificial intelligence; criminal liability; fault; artificial intelligence regulatio*

PENDAHULUAN

Rekonstruksi pertanggungjawaban pidana atas kejahatan *deepfake* berbasis kecerdasan buatan generatif merupakan kebutuhan penting dalam perkembangan hukum pidana di era transformasi digital. Perkembangan pembelajaran mendalam (deep learning) dan jaringan permusuhan generatif (*Generative Adversarial Networks*/GANs) memungkinkan terciptanya media audiovisual sintetis yang semakin sulit dibedakan dari media autentik. Kondisi tersebut memperluas bentuk kejahatan siber, antara lain pencemaran nama baik, penipuan identitas, kekerasan berbasis gender siber, serta disinformasi yang dapat memengaruhi kepercayaan publik (Chesney & Citron, 2019; Cole, 2022). Oleh karena itu, hukum pidana nasional perlu menata kembali kualifikasi perbuatan, hubungan kausal, dan subjek yang dapat dimintai pertanggungjawaban agar mampu merespons karakter kejahatan yang melibatkan teknologi generatif.

Sejumlah penelitian terdahulu telah memetakan implikasi hukum dari fenomena *deepfake* dan kejahatan berbasis kecerdasan buatan. Suhariyanto (2021) serta Rahardjo dan Santoso (2023) menunjukkan bahwa hukum pidana Indonesia masih bertumpu pada pertanggungjawaban pidana individual yang menghubungkan kesalahan dengan perbuatan pelaku. Cole (2022) serta Franks dan Waldman (2020) menyoroti kerugian

terhadap reputasi, privasi, dan kepentingan korban akibat penggunaan media sintetis. Kajian tersebut penting, tetapi sebagian besar masih menempatkan pengguna akhir sebagai pusat analisis. Persoalan menjadi lebih kompleks ketika suatu keluaran melanggar hukum terjadi melalui sistem generatif yang dirancang, dikembangkan, disediakan, dan digunakan oleh aktor yang berbeda.

Literatur lain menyoroti hambatan teknis dan yuridis dalam menuntut aktor di luar pengguna langsung. Karnow (2022) dan Harris (2023) menjelaskan bahwa penggunaan model generatif menimbulkan kesenjangan pertanggungjawaban karena pembuktian kesengajaan atau kelalaian tidak selalu dapat diarahkan secara langsung kepada sistem yang menghasilkan keluaran. Niko dan Wuryanto (2023) menunjukkan bahwa ketentuan mengenai manipulasi informasi elektronik menghadapi kesulitan ketika luaran berasal dari proses komputasional yang tidak sepenuhnya dapat diprediksi. Kusuma dan Adiputri (2023) serta Middleton (2022) juga memperlihatkan pentingnya menilai peran penyedia platform dan korporasi. Namun, belum terdapat perumusan yang memadai mengenai pembagian pertanggungjawaban berdasarkan tingkat kontribusi dan kendali setiap aktor.

Berdasarkan kesenjangan tersebut, penelitian ini menempatkan persoalan pertanggungjawaban sebagai masalah atribusi yang melibatkan sedikitnya tiga lapis peran, yaitu pihak yang merancang sistem, pihak yang menyediakan atau menerapkan sistem kepada publik, dan pihak yang menggunakan sistem untuk menghasilkan atau menyebarkan konten. Artikel ini bertujuan merekonstruksi konsep pertanggungjawaban pidana melalui pendekatan berbasis peran yang membedakan pertanggungjawaban berdasarkan desain, penyediaan, dan penggunaan. Rekonstruksi tersebut juga dikaitkan dengan kewajiban pelabelan konten dan penandaan digital sebagai mekanisme pendukung pembuktian, sehingga penegakan hukum tidak hanya berorientasi pada pemidanaan setelah kerugian terjadi, tetapi juga pada pencegahan dan keterlacakan proses produksi konten sintetis.

METODE

Penelitian ini merupakan penelitian hukum yuridis normatif yang menggunakan pendekatan perundang-undangan, pendekatan konseptual, dan pendekatan perbandingan hukum. Pendekatan perundang-undangan digunakan untuk menelaah ketentuan KUHP dan UU ITE yang berkaitan dengan kesalahan, pertanggungjawaban korporasi, pencemaran nama baik, serta manipulasi informasi atau dokumen elektronik. Pendekatan konseptual digunakan untuk menganalisis doktrin pertanggungjawaban pidana, hubungan antara *actus reus* dan *mens rea*, serta kemungkinan penerapan pertanggungjawaban berbasis peran dalam ekosistem kecerdasan buatan. Pendekatan perbandingan hukum digunakan dengan membandingkan prinsip pengaturan kecerdasan buatan di Uni Eropa melalui *EU Artificial Intelligence Act* dan di Tiongkok melalui pengaturan mengenai sintesis mendalam.

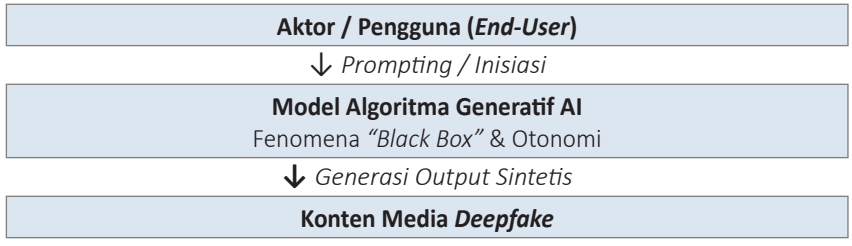
Bahan hukum yang digunakan terdiri atas bahan hukum primer dan sekunder. Bahan hukum primer meliputi peraturan perundang-undangan Indonesia serta instrumen regulasi kecerdasan buatan yang menjadi objek perbandingan. Bahan

hukum sekunder meliputi jurnal ilmiah, buku, dan dokumen kebijakan yang membahas *deepfake*, kecerdasan buatan, pertanggungjawaban pidana, serta pembuktian elektronik. Pengumpulan bahan hukum dilakukan melalui studi kepustakaan. Bahan hukum kemudian dianalisis secara kualitatif dengan teknik deskriptif-preskriptif, yaitu terlebih dahulu menguraikan keadaan hukum positif, mengidentifikasi celah normatif, kemudian merumuskan preskripsi hukum yang dapat digunakan untuk memperbaiki konstruksi pertanggungjawaban pidana. Analisis dilakukan secara sistematis, teleologis, dan komparatif agar rekonstruksi yang dihasilkan tetap berorientasi pada tujuan perlindungan hukum, kepastian hukum, dan keadilan. kebijakan terkait. Teknik analisis data mengaplikasikan metode analisis kualitatif deskriptif-preskriptif dengan menafsirkan bahan hukum secara sistematis, teleologis, dan komparatif guna menghasilkan preskripsi rekonstruksi pertanggungjawaban pidana yang adil dan berkeadilan.

HASIL DAN PEMBAHASAN

Konstruksi hukum pidana Indonesia secara fundamental dibangun di atas dogma *geen straf zonder schuld* (tiada pidana tanpa kesalahan) yang mewajibkan penyatuan imperatif antara *actus reus* (elemen perbuatan wujud) dan *mens rea* (sikap batin ketercelaan) pada diri subjek hukum individual (*natuurlijke persoon*) (Rahardjo & Santoso, 2023). Bab II KUHP Baru (UU No. 1/2023) mempertahankan akar doktrinal ini dengan menempatkan kesengajaan (*dolus*) dan kesengajaan atau kelalaian (*culpa*) sebagai kualifikasi utama untuk menagih pertanggungjawaban pidana (Hasibuan & Yusuf, 2024). Kendati demikian, munculnya teknologi kecerdasan buatan generatif (*generative artificial intelligence*) khususnya yang memanfaatkan arsitektur *Generative Adversarial Networks* (GANs) dan *Transformer models* menimbulkan distorsi paradigmatik pada pemenuhan unsur-unsur pidana tradisional tersebut (Scharre, 2023).

Alur interaksi teknis dan implikasi yuridis antara aktor, model algoritma, dan luaran media sintesis yang membentuk fenomena kejahatan siber baru ini diilustrasikan pada Gambar 1 di bawah ini:



Gambar 1. Alur Interaksi Generasi Konten Sintetis *Deepfake* dan Implikasi Otonomi Algoritma

Gambar 1 menjelaskan bagaimana perantara teknis tidak lagi berkedudukan sebagai alat pasif biasa, melainkan sebagai sistem otonom yang memproduksi sintesis audiovisual (*deepfake*) dengan tingkat presisi tinggi sehingga mengaburkan batas antara realitas dan rekayasa digital (Chesney & Citron, 2019). Dalam kerangka UU ITE (UU No. 1/2024), formulasi tindak pidana seperti Pasal 27A jo. Pasal 45 ayat (4) mengenai pencemaran nama baik, serta Pasal 35 jo. Pasal 51 ayat (1) mengenai

manipulasi dokumen elektronik, mengasumsikan adanya kausalitas langsung antara perbuatan subjek hukum dengan akibat yang ditimbulkan (Niko & Wuryanto, 2023).

Problem hukum substantif muncul ketika perantara perbuatan (perantara) merupakan sistem otonom berbasis *deep learning* yang beroperasi melalui proses *black box phenomenon* (Hildebrandt, 2021). Fenomena *black box* ini menunjuk pada ketidakmampuan manusia termasuk perancang algoritma (*developer*) maupun pengguna (*end-user*) untuk memprediksi secara presisi bagaimana variabel masukan (*prompt*) ditransformasikan oleh jaringan saraf tiruan (*neural networks*) menjadi luaran sintesis yang terdistorsi atau melanggar hukum (Richards & King, 2023).

Keterbatasan regulasi positif dan doktrin pidana Indonesia saat ini dalam menjangkau ekosistem AI generatif dipetakan secara rinci pada Tabel 1 di bawah ini:

Tabel 1. Keterbatasan Doktrin dan Regulasi Hukum Positif Indonesia terhadap AI Generatif

Regulasi / Doktrin	Ketentuan Terkait	Keterbatasan dalam Menjangkau AI Generatif
Sub Kepala Tabel: Regulasi Nasional	Sub Kepala Tabel: Normatif	Sub Kepala Tabel: Celah Hukum (*Legal Gap*)
KUHP Baru (UU No. 1/2023)	Pasal 36–37 (Bentuk Kesengajaan & Kelalaian)	Mempersyaratkan kesadaran batin manusia secara spesifik; gagal mengatribusikan niat jahat (*mens rea*) pada luaran algoritma otonom yang dihasilkan melalui *deep learning* (Rahardjo & Santoso, 2023).
KUHP Baru (UU No. 1/2023)	Pasal 45–50 (Pertanggungjawaban Korporasi)	Terbatas pada perbuatan pengurus/korporasi untuk keuntungan korporasi; tidak menjangkau kegagalan teknis instruksional sistem AI otonom secara spesifik (Middleton, 2022).
UU ITE (UU No. 1/2024)	Pasal 27A & Pasal 35 (Pencemaran & Manipulasi Dokumen)	Membutuhkan perbuatan langsung (*direct act*); kesulitan membuktikan kausalitas ketika luaran dihasilkan oleh *generative model* yang terintervensi fenomena *black box* (Kusuma & Adiputri, 2023).
Sub Kepala Tabel: Doktrin Hukum	Sub Kepala Tabel: Ajaran Penyertaan	Sub Kepala Tabel: Hambatan Subjektif
Doktrin Daderschap	Konsep Penyertaan (*Deelneming*)	Kategori *doenplegen* (menyuruh melakukan) menuntut perantara berupa manusia yang tidak dapat dipidana (*wilsovereenstemming*); sistem AI bukan subjek hukum (*non-person*) yang memiliki kehendak (*wil*) (Karnow, 2022).

Catatan: Penilaian keterbatasan didasarkan pada analisis doktrinal terhadap normatifikasi KUHP Baru dan UU ITE.

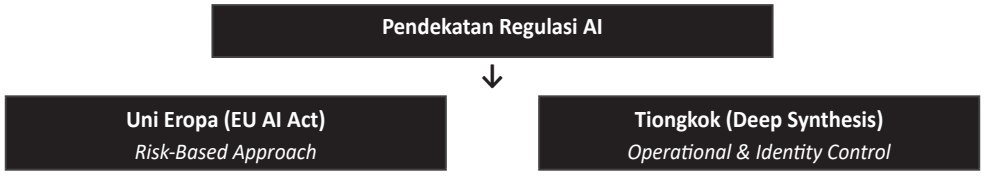
Berdasarkan pemetaan pada Tabel 1, terlihat jelas bahwa kondisi hukum saat ini menimbulkan *accountability gap* (kekosongan pertanggungjawaban) (Harris, 2023). Pengguna akhir (*end-user*) sering kali berargumen bahwa mereka hanya memasukkan *prompt* umum dan tidak menghendaki atau memprediksi luaran spesifik yang melanggar hukum (Scharre, 2023). Di sisi lain, pengembang (*developer*) berlindung di balik doktrin netralitas teknologi (*technology-neutrality doctrine*) dengan menyatakan bahwa model AI dipublikasikan sebagai alat serbaguna (*general-purpose AI*) (Koops,

2020). Ketimpangan ini membuktikan bahwa doktrin penyertaan (*deelneming*) seperti *doenplegen* (menyuruh melakukan) tidak dapat diterapkan secara analogis, sebab sistem AI bukanlah subjek hukum (*non-person*) yang dapat diposisikan sebagai alat yang disuruh (Karnow, 2022).

Selain masalah atribusi *mens rea*, kejahatan *deepfake* berdampak signifikan pada kekerasan berbasis gender siber (*cyber gender-based violence*) dalam bentuk *non-consensual deepfake pornography* (NCIFP) (Kirchengast, 2023). Manipulasi citra wajah atau tubuh korban yang digabungkan dengan konten eksploitasi seksual tidak dapat dijangkau secara optimal oleh ketentuan pornografi konvensional, sebab objek fisik asal tidak pernah secara nyata berada dalam adegan tersebut (Citron, 2019). Kerugian korban bersifat irreversibel dan merusak privasi seksual (*sexual privacy*), namun penjatuhan sanksi pidana terhambat oleh keraguan pembuktian mengenai siapa pelaku utama yang menginisiasi pembuatan dan penyebaran media sintesis tersebut (Franks & Waldman, 2020).

Perbandingan Hukum: Regulasi AI di Uni Eropa dan Tiongkok

Guna mengatasi keterbatasan doktrin pidana konvensional, penelusuran terhadap komparasi hukum internasional memberikan preseden tata kelola regulasi yang substansial (Sutherland, 2023). Pendekatan regulasi AI di tingkat global secara garis besar terbagi ke dalam dua paradigma utama sebagaimana ditunjukkan pada Gambar 2:



Gambar 2. Variasi Paradigma Pendekatan Regulasi AI Global

Sebagaimana diilustrasikan pada Gambar 2, perbedaan filosofis antara Uni Eropa dan Tiongkok menghasilkan instrumen mitigasi risiko yang dapat disintesiskan ke dalam sistem hukum Indonesia:

1. **Uni Eropa (*EU Artificial Intelligence Act*)**

   Uni Eropa mengadopsi pendekatan berbasis risiko (*risk-based approach*) melalui *Regulation (EU) 2024/1689* atau *EU Artificial Intelligence Act* (European Parliament, 2024; Kaminski, 2024). Regulasi ini mengklasifikasikan sistem AI ke dalam empat tingkatan risiko: *unacceptable risk* (dilarang), *high-risk*, *specific transparency risk*, dan *minimal risk* (Kaminski, 2024).

   - **Kewajiban Transparansi (*Transparency Obligations*):** Untuk AI generatif yang menghasilkan konten sintesis (*deepfake*), Uni Eropa menekankan kewajiban transparansi pada Pasal 50 *EU AI Act* (European Parliament, 2024).
   - **Standar Penandaan Teknis (*Machine-Readable Format*):** Penyedia sistem (*providers*) wajib memastikan bahwa konten buatan AI ditandai dalam format terstruktur yang dapat dideteksi secara teknis (*machine-readable format*), sehingga luaran dapat dikenali secara otomatis oleh sistem siber (Edwards, 2022).

- **Sanksi dan Akuntabilitas Pidana/Perdata:** Kegagalan memenuhi standar mitigasi risiko ini berakibat pada sanksi administratif finansial yang sangat berat (hingga puluhan juta Euro) serta membuka ruang bagi tanggung jawab perdata dan pidana berbasis kelalaian (*negligence*) bagi pengembang yang mengabaikan *duty of care* (Tzanou, 2023).

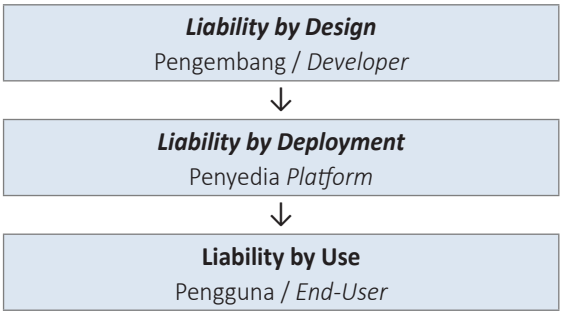
2. Tiongkok (*Deep Synthesis Provisions & Generative AI Measures*)

Berbeda dengan Uni Eropa yang berfokus pada hak asasi manusia dan kategorisasi risiko, Tiongkok menerapkan kontrol operasional dan identitas yang ketat melalui *Provisions on the Administration of Deep Synthesis of Internet Information Services* (Zhao, 2023).

- **Registrasi Identitas Asli (*Real-Name Registration*):** Setiap pengguna layanan sintesis mendalam wajib memverifikasi identitas resmi berbasis dokumen kewarganegaraan sebelum menggunakan fitur generatif guna menghilangkan anonimitas siber (Zhao, 2023).
- **Pendaftaran Algoritma (*Algorithm Registration*):** Penyedia layanan wajib mendaftarkan algoritma dan spesifikasi model mereka pada badan siber nasional (*Cyberspace Administration of China*) (Zhao, 2023).
- **Mandat *Watermarking* Digital:** Terdapat kewajiban memasang label terstruktur yang terlihat (*explicit mark*) dan label digital yang tertanam dalam metadata (*implicit watermark*) yang tidak dapat dihapus (*indelible*), sehingga rekam jejak pembuatan konten dapat dilacak secara instan (Zhao, 2023).

Rekonstruksi Pertanggungjawaban Pidana Berlapis Berbasis Peran *(Role-Based Liability)*

Berdasarkan kelemahan regulasi domestik dan diferensiasi pendekatan internasional, dirumuskan modifikasi teori pertanggungjawaban baru berupa pertanggungjawaban pidana fungsional yang membagi kriteria ketercelaan ke dalam tiga strata peran (*role-based liability*) (Floridi, 2023). Struktur pertanggungjawaban ini ditunjukkan pada Gambar 3:



Gambar 3. Model dan Hierarki Model Pertanggungjawaban Pidana Berlapis (*Role-Based Liability*)

Gambar 3 menguraikan pembagian beban kelalaian dan kesengajaan secara proporsional sesuai dengan tingkatan kontribusi aktor dalam ekosistem AI generatif:

1. ***Liability by Design* (Pengembang / *Developer*)**

   Pertanggungjawaban pengembang tidak didasarkan pada *mens rea* atas kejahatan spesifik yang dilakukan pengguna, melainkan pada *constructive intent* atau *recklessness* (kebangatan/kecerobohan ekstrem) dalam fase perancangan sistem (Wintgens & Van den Brande, 2023). Pengembang dapat dikenakan pertanggungjawaban pidana fungsional apabila (Wintgens & Van den Brande, 2023):

   a. Secara sengaja mematikan atau tidak memasang batas pengaman dasar (*safety guardrails*/*red-teaming*) yang berfungsi mencegah eksekusi *prompt* ilegal.
   b. Mengembangkan model algoritma yang secara spesifik dirancang untuk tujuan ilegal (seperti pembuatan konten eksploitasi seksual tanpa konsensual atau pemalsuan identitas finansial).
   c. Mengabaikan standar keamanan teknis yang patut (*duty of care*) dalam mencegah penyalahgunaan algoritma secara luas.

2. ***Liability by Deployment* (Penyedia *Platform* / Intermediari)**

   Penyedia platform yang mendistribusikan API (*Application Programming Interface*) atau antarmuka aplikasi kepada publik memikul kewajiban pengawasan korporasi (*corporate duty to monitor*) (Simanjuntak & Farida, 2023). Dalam kerangka pertanggungjawaban korporasi (Pasal 45–50 KUHP Baru), platform dapat dijatuhi pidana korporasi melalui doktrin *vicarious liability* atau *strict liability* jika (Simanjuntak & Farida, 2023):

   a. Pembiaran terhadap pola pemanfaatan kejahatan yang terdeteksi (*failure to act upon notice*).
   b. Ketiadaan mekanisme pemfilteran konten (*content moderation and abuse-filtering*) yang memadai untuk menyaring rekayasa media ilegal.
   c. Ketidaktersediaan saluran pelaporan dan penindakan cepat (*notice-and-takedown mechanism*) yang memfasilitasi penghentian penyebaran konten manipulatif.

3. ***Liability by Use* (Pengguna Akhir / *End-User*)**

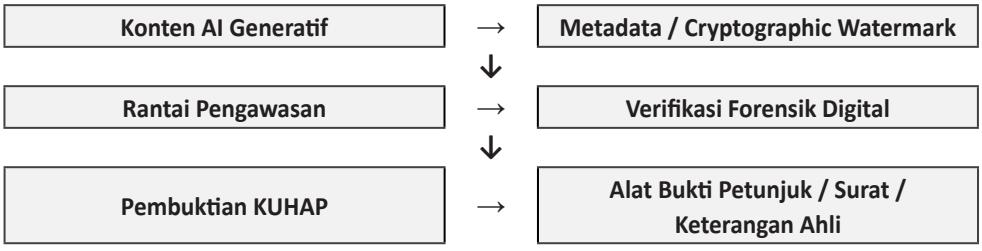
   Pengguna akhir memikul pertanggungjawaban pidana utama (*direct liability*) sebagai pelaku materiil (*pleger*) (Cole, 2022). *Mens rea* pengguna dinilai dari manipulasi masukan data (*prompt engineering*) dan niat jahat untuk merealisasikan akibat yang dilarang (Harris, 2023). Jika pengguna secara sadar mengarahkan AI untuk mereplikasi identitas seseorang tanpa izin guna menipu atau mencemarkan nama baik, perbuatan tersebut memenuhi unsur kesengajaan penuh (*dolus directus*) (Niko & Wuryanto, 2023).

Kewajiban Pelabelan Konten AI, *Digital Watermarking,* dan Integrasi Forensik Elektronik

Rekonstruksi pertanggungjawaban pidana tidak akan berjalan efektif tanpa didukung oleh pembaruan hukum acara pidana siber, khususnya terkait pembuktian (Pratama, 2023). Fenomena manipulasi media sintesis berpotensi mengikis kepercayaan terhadap bukti digital secara umum (*liar's dividend*), di mana pelaku kejahatan nyata

dapat berargumen bahwa bukti rekaman asli mereka adalah hasil rekayasa *deepfake* (Chesney & Citron, 2019).

Proses integrasi pembuktian teknis dan forensik digital dalam KUHAP digambarkan pada alur Gambar 4 berikut:



Gambar 4. Alur Integrasi Bukti Digital *Watermarking* ke dalam Pembuktian Pidana KUHAP

1. Inkorporasi *Digital Watermarking* sebagai Alat Bukti Sah

Pencantuman *watermark* kriptografis berbasis steganografi atau konsorsium C2PA (*Coalition for Content Provenance and Authenticity*) ke dalam setiap luaran media sintesis harus ditetapkan sebagai standar teknis wajib (Radley & Nissenbaum, 2023). Metadata yang tertanam memuat riwayat pembuatan (*content provenance*), identitas algoritma, dan stempel waktu (*timestamp*) (Lessig, 2024). Dalam KUHAP dan UU ITE, penjelajahan metadata ini berkualifikasi sebagai alat bukti surat, informasi elektronik, dan/atau alat bukti petunjuk yang sah di pengadilan (Pratama, 2023).

2. Penguatan *Cyber Chain of Custody*

Dengan adanya kewajiban pelabelan, penyidik forensik digital dapat melakukan ekstraksi sidik jari digital (*algorithmic fingerprinting*) untuk melacak rantai penguasaan barang bukti elektronik (*chain of custody*) secara presisi mulai dari tahap formulasi *prompt* oleh *end-user* hingga eksekusi oleh platform (Franke, 2022). Hal ini mengeliminasi hambatan *black box* dan memberikan kepastian hukum dalam pembuktian persidangan, sekaligus mendukung ketercapaian *Sustainable Development Goals* (SDGs) Tujuan Ke-16 terkait penguatan institusi hukum yang tangguh.

Dengan demikian, rekonstruksi pertanggungjawaban pidana *deepfake* seharusnya diarahkan pada pembentukan rantai akuntabilitas yang dapat ditelusuri. Setiap tahap, mulai dari desain sistem, penyediaan akses, penggunaan, pembuatan konten, hingga penyebaran, harus memiliki indikator kewajiban dan kesalahan yang jelas. Rantai tersebut akan memudahkan penegak hukum menentukan aktor yang harus dimintai pertanggungjawaban sekaligus mencegah pembebanan tanggung jawab kepada pihak yang tidak memiliki kendali yang memadai. Pendekatan ini melengkapi doktrin penyertaan konvensional dengan kerangka yang lebih sesuai terhadap ekosistem teknologi yang melibatkan manusia, korporasi, dan sistem komputasional.

Model ini juga memperlihatkan bahwa asas netralitas teknologi tidak dapat dipahami sebagai alasan untuk mengabaikan kewajiban kehati-hatian. Teknologi yang dapat digunakan untuk tujuan yang sah dan tidak sah tetap dapat menimbulkan

kewajiban tertentu ketika risiko penyalahgunaannya dapat diperkirakan dan aktor terkait memiliki kemampuan untuk mengendalikannya. Namun, kewajiban tersebut harus dibatasi secara jelas agar tidak menimbulkan beban yang tidak proporsional terhadap pengembang atau penyedia layanan. Dengan mempertemukan prinsip kehati-hatian, proporsionalitas, dan kesalahan, hukum pidana dapat menjaga keseimbangan antara perlindungan masyarakat dan keberlanjutan inovasi kecerdasan buatan (Koops, 2020; Sutherland, 2023).

Dari perspektif pembaruan hukum, rekonstruksi yang ditawarkan tidak harus diwujudkan melalui satu aturan pidana baru yang berdiri sendiri. Pembaruan dapat dilakukan melalui harmonisasi ketentuan mengenai tindak pidana yang sudah ada dengan aturan khusus mengenai tata kelola AI generatif. Ketentuan tersebut setidaknya perlu memperjelas definisi konten sintetis yang relevan secara hukum, kewajiban aktor berdasarkan tingkat kendali, standar pelabelan, mekanisme preservasi bukti digital, serta kondisi yang dapat menjadi dasar pertanggungjawaban korporasi. Dengan cara demikian, hukum pidana tidak hanya bereaksi terhadap akibat, tetapi juga membangun mekanisme pencegahan dan pembuktian yang sesuai dengan karakter teknologi generatif.

Perlindungan korban harus menjadi bagian integral dari rekonstruksi. Kejahatan *deepfake* tidak hanya menghasilkan kerugian ekonomi, tetapi juga dapat menyebabkan kerusakan reputasi, hilangnya privasi, tekanan sosial, dan kekerasan berbasis gender. Dalam kasus konten seksual sintetis, korban bahkan dapat kehilangan kendali atas representasi dirinya karena wajah atau tubuhnya ditempatkan dalam konteks yang tidak pernah terjadi. Oleh karena itu, sistem hukum perlu memberikan jalur pelaporan dan penghapusan konten yang cepat, sekaligus menjamin bahwa proses penghapusan tidak menghilangkan bukti yang diperlukan untuk penyidikan. Keseimbangan antara pemulihan korban dan preservasi barang bukti menjadi bagian penting dari penegakan hukum terhadap *deepfake* (Citron, 2019; Kirchengast, 2023).

Kewajiban pelabelan juga memiliki fungsi preventif. Ketika masyarakat dapat membedakan konten autentik dan konten sintetis, risiko kesalahan persepsi dan penyebaran informasi palsu dapat dikurangi. Pendekatan Uni Eropa menunjukkan pentingnya kewajiban transparansi terhadap konten yang dihasilkan atau dimanipulasi AI, sedangkan Tiongkok menekankan identifikasi pengguna, pendaftaran algoritma, dan penandaan digital. Kedua pendekatan tersebut tidak perlu diadopsi secara identik oleh Indonesia. Prinsip yang relevan dapat disesuaikan dengan kebutuhan nasional, terutama untuk membangun keterlacakan tanpa mengabaikan perlindungan data pribadi, kebebasan berekspresi, dan inovasi teknologi (European Parliament, 2024; Zhao, 2023). Implikasi Rekonstruksi terhadap Penegakan Hukum dan Perlindungan Korban

Rekonstruksi pertanggungjawaban pidana perlu disertai mekanisme pembuktian yang dapat menghubungkan suatu konten *deepfake* dengan proses pembuatannya. Hal ini penting karena pelaku dapat menghapus akun, mengubah metadata, atau memindahkan konten melalui berbagai platform. Pelabelan konten yang dihasilkan AI dan penandaan digital dapat membantu penyidik mengidentifikasi asal atau riwayat

suatu konten, meskipun mekanisme tersebut tidak dapat diposisikan sebagai bukti tunggal mengenai kesalahan pelaku. Metadata, riwayat pembuatan, waktu pembuatan, identitas layanan, dan rekam jejak distribusi harus ditempatkan sebagai bagian dari rangkaian alat bukti yang diverifikasi melalui pemeriksaan forensik (Franke, 2022; Pratama, 2023).

Model tiga lapis tersebut perlu dilengkapi dengan prinsip proporsionalitas. Satu peristiwa *deepfake* dapat melibatkan beberapa aktor dengan tingkat pengetahuan, kendali, dan kontribusi yang berbeda. Karena itu, penegak hukum perlu memisahkan perbuatan pembuatan sistem, penyediaan akses, penggunaan sistem, dan penyebaran konten. Pemisahan tersebut mencegah dua risiko sekaligus, yaitu kekosongan pertanggungjawaban ketika setiap aktor saling melempar tanggung jawab dan perluasan pertanggungjawaban yang berlebihan terhadap pihak yang hanya menyediakan teknologi secara netral. Prinsip proporsionalitas juga menuntut adanya pembuktian individual terhadap unsur kesalahan, sehingga model berbasis peran tidak berubah menjadi pertanggungjawaban otomatis berdasarkan status sebagai pengembang atau penyedia platform.

Sementara itu, pengguna akhir merupakan aktor yang paling dekat dengan perbuatan materiil ketika ia secara sadar memberikan masukan, mengarahkan sistem, menghasilkan konten, dan menyebarkannya untuk mencapai akibat yang dilarang. Penilaian kesalahan terhadap pengguna dapat mempertimbangkan isi perintah yang diberikan kepada sistem, tujuan penggunaan, pengetahuan mengenai identitas atau kondisi korban, serta tindakan setelah konten dihasilkan. Apabila pengguna secara sadar membuat konten palsu untuk menipu, mencemarkan nama baik, atau mengeksploitasi seseorang, keberadaan sistem AI sebagai sarana tidak menghilangkan pertanggungjawaban pidananya. Dalam konteks ini, teknologi diperlakukan sebagai instrumen yang digunakan oleh pelaku, sedangkan unsur kesalahan tetap dinilai pada diri pengguna (Cole, 2022; Harris, 2023).

Pada tingkat penyedia platform, fokus atribusi bergeser pada pengendalian operasional. Penyedia platform memiliki posisi yang berbeda dari pengembang karena berhubungan langsung dengan distribusi layanan, pengelolaan akses, moderasi konten, dan mekanisme pelaporan. Oleh sebab itu, kewajiban hukum dapat diarahkan pada pencegahan dan penghentian penyalahgunaan setelah terdapat pemberitahuan yang memadai. Kegagalan bertindak tidak dapat secara otomatis dipersamakan dengan keterlibatan langsung dalam kejahatan, tetapi dapat menjadi dasar pertanggungjawaban apabila penyedia memiliki kapasitas nyata untuk mencegah penyebaran, mengetahui adanya pola penyalahgunaan, dan tetap membiarkannya. Kerangka demikian memungkinkan penerapan pertanggungjawaban korporasi secara proporsional tanpa mengabaikan batas antara perbuatan pidana dan kegagalan administratif (Middleton, 2022; Simanjuntak & Farida, 2023).

Pada tingkat pengembang, ukuran pertanggungjawaban terutama berkaitan dengan keputusan desain dan pengendalian risiko yang secara wajar dapat diperkirakan. Pengembang tidak semestinya dipidana hanya karena suatu model serbaguna kemudian

disalahgunakan oleh pihak lain. Pertanggungjawaban dapat dibebankan apabila terdapat pengetahuan atau keadaan yang semestinya diketahui mengenai risiko penyalahgunaan yang serius, tetapi pengembang secara sadar mengabaikan pengamanan yang layak. Misalnya, pengembang dengan sengaja menghilangkan mekanisme pencegahan yang secara teknis tersedia, merancang model untuk menghasilkan konten terlarang, atau mengabaikan risiko yang telah diketahui secara luas. Dengan kriteria tersebut, pertanggungjawaban tidak berubah menjadi pertanggungjawaban mutlak, melainkan tetap berlandaskan tingkat kesalahan dan kewajiban kehati-hatian yang relevan (Wintgens & Van den Brande, 2023; Scharre, 2023).

Kriteria Atribusi Pertanggungjawaban Pidana dalam Ekosistem *Deepfake* Rekonstruksi pertanggungjawaban pidana tidak berarti menghapus asas kesalahan atau menjadikan setiap pihak yang berada dalam rantai teknologi secara otomatis bertanggung jawab. Prinsip tersebut justru harus dipertahankan dengan menyesuaikan cara menilai kesalahan terhadap posisi, kemampuan mengendalikan risiko, dan kontribusi faktual masing-masing aktor. Dengan demikian, pertanggungjawaban berbasis peran harus dipahami sebagai mekanisme atribusi yang tetap mensyaratkan adanya hubungan antara tindakan atau pembiaran aktor dengan risiko serta akibat yang dilarang. Pendekatan ini sejalan dengan gagasan akuntabilitas pada sistem otonom yang menuntut kejelasan mengenai siapa yang memiliki kemampuan untuk mencegah atau mengurangi risiko pada setiap tahap siklus hidup teknologi (Hildebrandt, 2021; Richards & King, 2023).

Dari sisi kelembagaan, penegakan hukum terhadap *deepfake* juga membutuhkan kapasitas penyidik, penuntut umum, dan hakim dalam memahami bukti digital serta karakteristik sistem AI. Pembuktian tidak cukup hanya menunjukkan bahwa suatu gambar, suara, atau video merupakan hasil manipulasi. Penegak hukum harus dapat menjelaskan hubungan antara konten, akun atau perangkat yang digunakan, proses pembuatannya, dan tindakan aktor yang diduga bertanggung jawab. Karena itu, standar pemeriksaan forensik, dokumentasi rantai penguasaan barang bukti, dan kompetensi ahli menjadi unsur penting dalam memastikan bahwa rekonstruksi pertanggungjawaban dapat diterapkan secara konsisten. Tanpa penguatan kapasitas tersebut, perubahan norma berisiko tidak menghasilkan peningkatan efektivitas penegakan hukum sebagaimana diharapkan (Pratama, 2023; Franke, 2022).

Selain itu, rekonstruksi harus memperhatikan pembagian fungsi antara hukum pidana dan mekanisme nonpidana. Tidak setiap pelanggaran kewajiban pelabelan atau moderasi harus berujung pada pidana. Sanksi administratif, kewajiban perbaikan sistem, penghapusan konten, dan mekanisme pemulihan korban dapat ditempatkan sebagai instrumen pelengkap, sedangkan hukum pidana digunakan terhadap perbuatan yang memenuhi tingkat ketercelaan dan unsur tindak pidana. Pembedaan tersebut penting untuk mencegah kriminalisasi berlebihan sekaligus memberikan insentif kepada pengembang dan penyedia platform untuk menerapkan pengamanan sejak tahap desain. Pengalaman Uni Eropa yang menekankan pendekatan berbasis risiko dan pengalaman Tiongkok yang menekankan identifikasi serta pengendalian

operasional menunjukkan bahwa tata kelola AI yang efektif membutuhkan kombinasi instrumen hukum, bukan hanya ketentuan pidana (European Parliament, 2024; Zhao, 2023; Kaminski, 2024).

Batas Pertanggungjawaban dan Asas Kesalahan Penerapan model pertanggung-jawaban berbasis peran harus tetap dibatasi oleh asas kesalahan. Hal ini penting karena hubungan seseorang dengan teknologi tidak dengan sendirinya membuktikan adanya niat atau kelalaian yang dapat dipidana. Dalam perkara *deepfake*, pembuktian harus membedakan antara penggunaan teknologi untuk tujuan yang sah, penyalahgunaan yang tidak disengaja, dan penggunaan yang sejak awal diarahkan untuk menghasilkan akibat yang dilarang. Pembatasan tersebut dapat dilakukan dengan menilai pengetahuan pelaku mengenai sifat konten, tujuan pembuatan, tingkat kendali terhadap proses, tindakan setelah mengetahui adanya pelanggaran, serta kemampuan yang tersedia untuk mencegah atau menghentikan akibat. Pendekatan demikian mempertahankan prinsip bahwa pidana harus didasarkan pada kesalahan sekaligus menyesuaikan metode atribusi dengan perubahan struktur perbuatan yang dimediasi teknologi (Rahardjo & Santoso, 2023; Scharre, 2023).

Rekonstruksi tersebut pada akhirnya harus menghasilkan norma yang mudah diterapkan. Norma tidak cukup menyebut bahwa pengembang, platform, dan pengguna dapat bertanggung jawab, tetapi harus menjelaskan indikator kesalahan, bentuk kewajiban, serta keadaan yang menghubungkan pelanggaran kewajiban dengan akibat yang dilarang. Kejelasan indikator tersebut akan mengurangi perbedaan penafsiran dan memberikan pedoman bagi aparat penegak hukum dalam menentukan kapan suatu peristiwa *deepfake* masuk ke ranah pidana. Dengan demikian, pembaruan hukum tidak hanya bersifat reaktif, tetapi membentuk sistem akuntabilitas yang dapat bekerja sejak tahap perancangan teknologi sampai tahap penggunaan dan penyebaran konten.

Dengan kerangka tersebut, setiap aktor memperoleh batas kewajiban yang lebih terukur dan korban memperoleh dasar perlindungan yang lebih jelas. Kerangka ini sekaligus menjaga agar inovasi teknologi tidak dipandang sebagai sumber kesalahan semata, melainkan sebagai ruang yang memerlukan tata kelola dan akuntabilitas.

Kedudukan *Deepfake* sebagai Perbuatan yang Relevan dalam Hukum Pidana Indonesia

Deepfake pada dasarnya tidak merupakan satu jenis tindak pidana yang berdiri sendiri, melainkan teknologi atau metode yang dapat digunakan untuk merealisasikan berbagai bentuk perbuatan yang telah dikenal dalam hukum pidana. Karakter tersebut menimbulkan konsekuensi penting bagi kebijakan kriminal. Hukum tidak seharusnya semata-mata mempidanakan keberadaan teknologi atau penggunaan kecerdasan buatan, tetapi harus mengidentifikasi perbuatan konkret, kepentingan hukum yang dilanggar, akibat yang ditimbulkan, serta hubungan antara aktor dan akibat tersebut. Dengan demikian, pendekatan terhadap *deepfake* harus berangkat dari prinsip bahwa teknologi merupakan sarana, sedangkan pertanggungjawaban pidana tetap diarahkan kepada subjek hukum yang melakukan, memerintahkan, memfasilitasi, atau secara relevan membiarkan terjadinya perbuatan yang memenuhi unsur tindak pidana.

Karakter multifungsi *deepfake* menyebabkan satu konten dapat sekaligus berkaitan dengan lebih dari satu kepentingan hukum. Video palsu yang menggunakan wajah seseorang, misalnya, dapat dibuat untuk merusak reputasi, mengambil keuntungan ekonomi, melakukan penipuan, mengintimidasi korban, atau menyebarkan informasi politik yang menyesatkan. Dalam kasus tertentu, satu rangkaian tindakan bahkan dapat mencakup proses memperoleh data korban, membuat materi sintetis, mengunggahnya ke platform, kemudian menyebarkannya melalui beberapa akun. Karena itu, konstruksi hukum harus membedakan antara tahap penciptaan, penguasaan, dan distribusi konten. Pembedaan ini penting agar penyidik dan hakim tidak berhenti pada pertanyaan mengenai siapa yang pertama kali membuat konten, tetapi juga dapat menilai siapa yang memiliki kontribusi dan kendali pada setiap tahap.

Dalam konteks UU ITE, persoalan utama bukan hanya apakah *deepfake* dapat dikategorikan sebagai informasi elektronik yang dimanipulasi, melainkan bagaimana membuktikan bahwa manipulasi tersebut merupakan perbuatan yang dilakukan dengan kesalahan oleh terdakwa. Ketika seseorang menggunakan perangkat lunak untuk mengubah suara atau wajah, terdapat tindakan manusia yang dapat diamati, seperti memilih bahan, menentukan sasaran, memasukkan *prompt*, melakukan penyuntingan, menyimpan hasil, dan menyebarluaskan konten. Rangkaian tindakan tersebut dapat menjadi dasar untuk membangun *actus reus*. Persoalan *mens rea* muncul ketika pelaku menyatakan bahwa hasil akhir sepenuhnya dihasilkan sistem dan tidak dapat diprediksi. Dalih tersebut tidak dapat secara otomatis menghapus pertanggungjawaban, terutama apabila bukti menunjukkan bahwa pelaku memberikan instruksi yang secara jelas diarahkan untuk menghasilkan konten tertentu dan kemudian melakukan tindakan lanjutan untuk mencapai akibat yang dilarang.

Sebaliknya, tidak tepat pula apabila setiap pengguna langsung dianggap bertanggung jawab atas seluruh keluaran sistem. Sistem generatif dapat menghasilkan variasi yang tidak seluruhnya dapat diprediksi dari suatu *input*. Dalam penggunaan yang sah, seseorang dapat membuat simulasi, karya seni, pendidikan, penelitian, atau hiburan tanpa memiliki tujuan untuk merugikan orang lain. Oleh karena itu, pertanggungjawaban harus mempertimbangkan tujuan penggunaan, tingkat pengetahuan, bentuk instruksi, tindakan setelah mengetahui hasil, serta hubungan kausal antara perilaku pengguna dan akibat yang dilarang. Prinsip ini menjaga agar pembaruan hukum tidak berubah menjadi *strict liability* yang terlalu luas dan bertentangan dengan asas kesalahan.

Rekonstruksi *Actus Reus* dan *Mens Rea* dalam Kejahatan *Deepfake*

Doktrin *actus reus* dan *mens rea* tetap relevan dalam menghadapi teknologi generatif, tetapi cara mengidentifikasinya perlu disesuaikan dengan struktur perbuatan digital. Pada tindak pidana konvensional, hubungan antara pelaku dan alat yang digunakan biasanya lebih mudah ditelusuri. Dalam kejahatan *deepfake*, terdapat lapisan tambahan berupa model algoritma, server, antarmuka, dataset, sistem penyaringan, dan *platform* distribusi. Lapisan tersebut tidak berarti bahwa seluruh teknologi harus diperlakukan sebagai pelaku. Sebaliknya, teknologi perlu diposisikan sebagai bagian dari rantai kausal yang membantu menjelaskan bagaimana tindakan manusia menghasilkan akibat tertentu.

Pada pengguna akhir, *actus reus* dapat ditemukan dalam tindakan yang secara sadar menginisiasi pembuatan konten, memilih identitas korban, mengunggah bahan sumber, memasukkan prompt, melakukan penyuntingan, atau mendistribusikan hasil. Semakin spesifik dan terarah tindakan tersebut, semakin kuat hubungan faktual antara perilaku pengguna dan hasil. Mens rea kemudian dapat dinilai dari konteks tindakan. *Prompt* yang secara eksplisit meminta sistem untuk menempatkan wajah korban pada konten tertentu, misalnya, memiliki nilai pembuktian yang berbeda dengan *prompt* umum yang tidak mengarah kepada individu tertentu. Demikian pula, tindakan menghapus label AI, mengganti nama file, menggunakan akun anonim, atau menyebarkan konten berulang kali setelah menerima keberatan dari korban dapat menjadi keadaan yang relevan dalam menilai pengetahuan dan tujuan pelaku.

Pada pengembang, *actus reus* tidak dapat dirumuskan hanya berdasarkan fakta bahwa model yang dibuat kemudian disalahgunakan. Pengembang pada dasarnya menyediakan teknologi yang dapat memiliki penggunaan sah. Oleh sebab itu, dasar atribusi harus lebih spesifik, yaitu tindakan atau pembiaran dalam desain, pengujian, distribusi, dan pengendalian risiko. Pertanggungjawaban dapat menjadi relevan ketika pengembang mengetahui risiko serius yang dapat diperkirakan, mempunyai kemampuan teknis untuk mengurangi risiko, tetapi dengan sengaja menghilangkan pengamanan yang wajar atau bahkan merancang sistem untuk memfasilitasi tujuan ilegal. Pendekatan ini sejalan dengan prinsip proporsionalitas karena menghubungkan tanggung jawab dengan tingkat kontrol yang benar-benar dimiliki pengembang.

Penyedia *platform* memiliki posisi antara pengembang dan pengguna. *Platform* mungkin tidak menciptakan model dasar, tetapi dapat mengendalikan akses, menetapkan ketentuan penggunaan, menyediakan sistem pelaporan, melakukan moderasi, menyimpan *log*, dan mengambil tindakan terhadap konten. Karena itu, ukuran pertanggungjawaban *platform* harus difokuskan pada kewajiban yang berada dalam lingkup kontrolnya. Kegagalan melakukan tindakan setelah menerima pemberitahuan yang jelas mengenai konten ilegal dapat memiliki bobot yang berbeda dengan kegagalan yang terjadi ketika *platform* sama sekali tidak mengetahui adanya penyalahgunaan. Pembedaan ini penting agar kewajiban korporasi tidak berubah menjadi kewajiban absolut untuk mengetahui dan mencegah seluruh konten yang dibuat oleh pengguna.

Masalah Kausalitas dan *Black Box*

Hubungan kausal menjadi salah satu persoalan paling penting dalam pertanggungjawaban pidana atas *deepfake*. Pada satu sisi, keluaran sistem merupakan hasil dari proses komputasional yang dipengaruhi data pelatihan, arsitektur model, parameter, *prompt*, dan proses pascaproduksi. Pada sisi lain, hukum pidana membutuhkan hubungan yang cukup jelas antara perbuatan terdakwa dan akibat yang dilarang. Fenomena *black box* tidak berarti hubungan kausal sama sekali tidak dapat ditemukan. Yang berubah adalah objek pembuktian. Penegak hukum perlu menelusuri keseluruhan rangkaian tindakan dan tidak hanya berusaha memahami setiap proses internal jaringan saraf.

Dalam praktik pembuktian, pertanyaan yang lebih relevan adalah siapa yang menginisiasi proses, perangkat apa yang digunakan, akun apa yang digunakan untuk mengakses layanan, bahan sumber apa yang dimasukkan, prompt apa yang diberikan, kapan konten dihasilkan, apakah terdapat proses penyuntingan, dan siapa yang pertama kali mendistribusikannya. *Log* akses, metadata, riwayat akun, perangkat, waktu pembuatan, serta bukti komunikasi dapat membantu membangun rangkaian tersebut. Dengan demikian, *black box* lebih tepat dipahami sebagai tantangan teknis terhadap pembuktian, bukan alasan normatif untuk menghapus pertanggungjawaban.

Kausalitas juga harus dibedakan dari prediktabilitas. Seorang pengguna mungkin tidak dapat memprediksi setiap detail visual yang dihasilkan model, tetapi dapat memprediksi tujuan dan karakter umum dari output ketika ia memberikan instruksi yang sangat spesifik. Dalam keadaan demikian, ketidakpastian teknis tidak otomatis memutus hubungan antara tindakan dan akibat. Sebaliknya, apabila output yang melanggar hukum muncul secara tidak terduga dari penggunaan yang wajar, kemudian pengguna tidak mengetahui dan tidak menyebarkannya, dasar untuk pertanggungjawaban pidana akan jauh lebih lemah. Penilaian tersebut menunjukkan pentingnya pendekatan kasus per kasus.

Pada pengembang dan penyedia *platform*, kausalitas harus dinilai melalui konsep kontribusi dan pengendalian risiko. Tidak setiap kondisi yang memungkinkan terjadinya kejahatan dapat dianggap sebagai penyebab hukum yang cukup. Mesin pencari, penyedia komputasi awan, penyedia perangkat keras, dan layanan internet secara umum dapat menjadi bagian dari lingkungan teknologi tanpa otomatis bertanggung jawab atas tindak pidana pengguna. Oleh karena itu, rekonstruksi yang diusulkan harus memiliki ambang yang jelas: adanya pengetahuan atau keadaan yang semestinya diketahui, adanya kemampuan untuk melakukan pencegahan atau mitigasi, adanya pelanggaran terhadap kewajiban yang relevan, serta adanya hubungan yang cukup antara pelanggaran tersebut dan risiko atau akibat yang dilarang.

Pertanggungjawaban Korporasi dan Pembagian Peran

Perkembangan AI generatif memperkuat pentingnya pertanggungjawaban korporasi karena sebagian besar model dan *platform* dikembangkan serta dioperasikan oleh badan usaha. Namun, pengenaan pidana terhadap korporasi harus tetap memperhatikan prinsip individualisasi kesalahan. Status sebagai badan usaha yang memperoleh keuntungan dari teknologi tidak cukup untuk membuktikan keterlibatan pidana. Harus terdapat dasar normatif yang menghubungkan perbuatan orang dalam struktur korporasi, kebijakan perusahaan, kegagalan pengawasan, atau keputusan organisasi dengan tindak pidana yang terjadi.

Dalam model *liability by design*, penilaian terhadap korporasi dapat diarahkan pada proses manajemen risiko. Apabila perusahaan telah mengetahui bahwa sistem tertentu memiliki risiko tinggi untuk digunakan membuat konten ilegal, tetapi secara sengaja mengabaikan pengujian keamanan, menonaktifkan mekanisme penyaringan, atau memasarkan fitur tersebut secara khusus untuk tujuan yang melanggar hukum, maka tingkat ketercelaannya berbeda dari perusahaan yang telah menerapkan pengamanan

memadai tetapi tetap mengalami penyalahgunaan oleh pengguna. Ukuran tersebut membantu membedakan negligence biasa, pelanggaran administratif, dan kesalahan yang layak masuk ke ranah pidana.

Liability by deployment memiliki dimensi berbeda. Penyedia *platform* dapat memiliki informasi yang tidak tersedia bagi pengembang, misalnya pola penggunaan, laporan pengguna, frekuensi pelanggaran, atau jaringan akun yang menyebarkan konten. Oleh karena itu, kewajiban platform dapat berkembang seiring meningkatnya pengetahuan dan kemampuan deteksi. Ketika suatu platform menerima pemberitahuan yang kredibel mengenai *deepfake* yang merugikan korban dan mempunyai sarana teknis untuk membatasi penyebarannya, kegagalan tanpa alasan yang dapat dibenarkan dapat menjadi faktor penting dalam menilai pertanggungjawaban. Namun, kewajiban tersebut harus dirumuskan secara proporsional dan disertai prosedur agar tidak menimbulkan penghapusan konten yang sah secara sewenang-wenang.

Pemisahan tanggung jawab juga diperlukan ketika satu *deepfake* dibuat oleh satu orang tetapi disebarkan secara masif oleh pihak lain. Pengguna pertama dapat menjadi pelaku utama jika ia memiliki niat jahat dalam pembuatan. Orang yang kemudian mengetahui sifat ilegal konten dan secara aktif menyebarkannya dapat memiliki bentuk pertanggungjawaban yang berbeda. Platform yang tidak terlibat dalam pembuatan tidak boleh diperlakukan sama dengan aktor yang sengaja mengembangkan fitur untuk tujuan ilegal. Dengan demikian, role-based liability tidak dimaksudkan untuk memperbanyak pelaku, tetapi untuk menghindari dua ekstrem, yaitu impunitas akibat rantai teknologi yang kompleks dan *overcriminalization* terhadap pihak yang kontribusinya netral.

Perlindungan Korban dan Pemulihan Kerugian

Persoalan *deepfake* tidak dapat diselesaikan hanya dengan menentukan siapa yang dipidana. Korban sering kali membutuhkan penghentian penyebaran, pemulihan reputasi, penghapusan konten, perlindungan terhadap identitas, dan pemulihan kerugian. Dalam kasus *deepfake* seksual, kerugian bahkan dapat berlangsung lama karena salinan konten dapat dibuat kembali dan disebarkan melalui layanan yang berbeda. Oleh sebab itu, mekanisme perlindungan korban harus dirancang bersamaan dengan mekanisme pertanggungjawaban.

Prosedur *notice-and-takedown* menjadi penting, tetapi harus dirancang agar penghapusan konten tidak menghilangkan bukti. Ketika korban melaporkan konten, platform idealnya dapat melakukan dua tindakan secara paralel: membatasi akses publik terhadap konten dan mempertahankan data yang relevan bagi proses hukum sesuai ketentuan yang berlaku. Pendekatan ini memungkinkan kepentingan korban dan kepentingan penyidikan berjalan bersamaan. Jika konten langsung dihapus tanpa preservasi, bukti dapat hilang. Sebaliknya, jika konten dibiarkan tersebar selama proses verifikasi yang terlalu panjang, kerugian korban dapat semakin besar.

Perlindungan korban juga perlu memperhatikan risiko *secondary victimization*. Korban *deepfake* seksual, misalnya, tidak seharusnya dipaksa berulang kali menjelaskan atau memperlihatkan materi yang bersifat sensitif hanya untuk memperoleh

perlindungan. Sistem peradilan perlu mengembangkan prosedur pemeriksaan yang menjaga privasi dan membatasi penyebaran ulang barang bukti. *Digital forensic examination* harus dilakukan secara terukur sehingga kebutuhan pembuktian tidak menjadi alasan untuk memperluas distribusi materi yang merugikan korban.

Selain itu, pemulihan reputasi harus dipahami sebagai bagian dari kebijakan hukum. *Deepfake* dapat menyebar lebih cepat daripada kemampuan korban melakukan klarifikasi. Sekalipun konten kemudian dinyatakan palsu, sebagian penerima mungkin tetap mempercayainya. Oleh karena itu, pelabelan, klarifikasi, penghapusan, dan mekanisme koreksi perlu berjalan secara terintegrasi. Pendekatan ini memperlihatkan bahwa kebijakan kriminal terhadap *deepfake* harus memiliki orientasi restoratif dan preventif, bukan semata-mata represif.

Formulasi Model Rekonstruksi dalam Hukum Nasional

Berdasarkan keseluruhan analisis, rekonstruksi pertanggungjawaban pidana dapat dituangkan dalam beberapa prinsip normatif. Pertama, hukum perlu memberikan definisi kerja mengenai konten sintetis yang relevan secara hukum. Definisi tersebut tidak boleh terlalu luas sehingga mencakup seluruh karya hasil AI, tetapi harus menitikberatkan pada manipulasi atau sintesis identitas, representasi, atau informasi yang digunakan untuk menimbulkan akibat yang dilarang. Definisi yang jelas diperlukan agar masyarakat dapat membedakan penggunaan teknologi yang sah dari penggunaan yang berpotensi pidana.

Kedua, hukum perlu menggunakan prinsip kendali dan kontribusi sebagai dasar pembagian tanggung jawab. Semakin besar kemampuan seorang aktor untuk mengendalikan risiko dan semakin spesifik kontribusinya terhadap perbuatan, semakin kuat dasar untuk menilai pertanggungjawaban. Pengguna yang secara sadar membuat dan menyebarkan konten palsu untuk menipu berada pada posisi berbeda dari pengembang model serbaguna. Pengembang yang secara sadar menyediakan sistem untuk tujuan ilegal juga berbeda dari pengembang yang telah menerapkan pengamanan layak tetapi modelnya disalahgunakan secara tidak terduga.

Ketiga, kewajiban transparansi perlu ditempatkan sebagai instrumen pendukung. Pelabelan dan *watermarking* tidak dengan sendirinya membuktikan bahwa seseorang telah melakukan tindak pidana. Namun, keduanya dapat membantu membedakan konten sintetis dari konten autentik, memudahkan penelusuran asal konten, dan meningkatkan kualitas pembuktian. Karena itu, kewajiban teknis harus disertai standar keamanan, interoperabilitas, dan prosedur preservasi agar dapat digunakan secara efektif dalam penegakan hukum.

Keempat, pengaturan pidana perlu diintegrasikan dengan sanksi administratif dan mekanisme perdata. Pelanggaran kewajiban pelabelan yang bersifat teknis tidak selalu membutuhkan pidana. Sanksi administratif dapat digunakan untuk pelanggaran yang tidak mencapai tingkat ketercelaan pidana, sedangkan hukum pidana diarahkan pada tindakan yang memiliki kesalahan dan akibat serius. Pembagian ini penting agar hukum tidak menimbulkan efek jera yang justru menghambat pengembangan teknologi dan penelitian.

Kelima, perlu ada pedoman pembuktian yang mengatur bagaimana data AI diperlakukan dalam proses penyidikan dan persidangan. Bukti tidak seharusnya hanya berupa hasil akhir video atau gambar. Bukti yang lebih lengkap mencakup perangkat, akun, *log, prompt*, metadata, komunikasi, riwayat penyimpanan, dan informasi distribusi. Seluruh data tersebut harus diperoleh dan dipertahankan melalui prosedur yang menjamin integritasnya. Ahli digital forensik diperlukan untuk menjelaskan keterkaitan antara bukti teknis dan tindakan terdakwa.

Keenam, hukum perlu mempertahankan asas kesalahan sebagai batas utama. Rekonstruksi bukanlah alasan untuk menerapkan pertanggungjawaban otomatis. Dalam setiap perkara harus dibuktikan apakah terdakwa mengetahui, menghendaki, atau setidaknya lalai terhadap risiko yang relevan sesuai rumusan tindak pidana yang diterapkan. Untuk pengembang dan *platform*, kewajiban kehati-hatian harus dirumuskan secara konkret sehingga seseorang dapat mengetahui apa yang diwajibkan hukum. Hal ini sekaligus menjaga asas legalitas dan kepastian hukum.

Penerapan Model melalui Skenario Kasus

Untuk melihat penerapan model tersebut, dapat digunakan skenario hipotetis. Seorang pengguna memperoleh foto seseorang dari media sosial, memasukkannya ke sistem generatif, memberikan instruksi agar wajah korban ditempatkan pada video seksual, kemudian mengunggah hasilnya ke beberapa platform dengan keterangan palsu. Dalam skenario ini, pengguna merupakan aktor yang paling dekat dengan perbuatan. Pemilihan korban, penggunaan materi sumber, instruksi yang diberikan, serta tindakan penyebaran menunjukkan adanya rangkaian perilaku manusia yang dapat dinilai dari sisi *actus reus* dan *mens rea*. Kehadiran AI sebagai alat tidak menghilangkan tanggung jawab pengguna.

Skenario kedua berbeda. Pengembang menyediakan model serbaguna yang memiliki pengamanan terhadap permintaan tertentu. Seorang pengguna berhasil melakukan jailbreak dan menghasilkan konten ilegal. Apabila pengembang telah menerapkan pengujian dan pengamanan yang layak, tidak mengetahui penyalahgunaan tersebut, dan tidak memiliki alasan konkret untuk memperkirakannya dalam konteks tertentu, dasar pertanggungjawaban pidana pengembang menjadi lemah. Hal ini menunjukkan pentingnya prinsip proporsionalitas. Pengembang tidak boleh dipidana hanya karena teknologinya dapat disalahgunakan.

Skenario ketiga adalah ketika pengembang mengetahui bahwa model secara sistematis digunakan untuk membuat *deepfake* seksual terhadap korban tertentu, menerima laporan berulang, tetapi secara sengaja menghilangkan filter yang sebelumnya dapat membatasi penyalahgunaan tersebut karena pertimbangan keuntungan. Dalam keadaan seperti ini, posisi pengembang berbeda. Pengetahuan mengenai risiko, kemampuan teknis untuk melakukan mitigasi, keputusan untuk tidak bertindak, dan keberlanjutan penyalahgunaan dapat menjadi faktor yang relevan untuk menilai kesalahan.

Skenario keempat melibatkan *platform* distribusi. *Platform* menerima pemberitahuan yang memadai mengenai video *deepfake* ilegal, memiliki mekanisme untuk membatasi penyebaran, dan dapat mengidentifikasi akun yang bertanggung jawab. Jika *platform* secara sadar membiarkan penyebaran berlangsung demi mempertahankan keterlibatan pengguna, pertanyaan pertanggungjawaban dapat muncul. Namun, jika *platform* belum menerima pemberitahuan, sistem deteksinya tidak mampu mengidentifikasi konten secara wajar, dan tidak ada bukti bahwa *platform* mengetahui pola penyalahgunaan, dasar untuk pertanggungjawaban pidana harus diuji secara lebih hati-hati.

Skenario-skenario tersebut menunjukkan bahwa role-based liability bukan sistem yang menetapkan hukuman berdasarkan jabatan. Model tersebut justru berfungsi sebagai kerangka analisis. Penyidik perlu mengidentifikasi aktor, tindakan, pengetahuan, kemampuan pengendalian, kewajiban yang berlaku, serta akibat. Penuntut umum kemudian harus membuktikan unsur tindak pidana dan kesalahan secara individual. Hakim pada akhirnya menilai seluruh bukti dan memastikan bahwa pemidanaan tidak didasarkan pada asumsi bahwa teknologi selalu berbahaya.

Implikasi Kelembagaan bagi Aparat Penegak Hukum

Rekonstruksi norma tidak akan efektif tanpa peningkatan kapasitas institusi. Penyidik perlu memiliki kemampuan untuk mengamankan perangkat dan data digital, memahami log layanan AI, mengidentifikasi metadata, serta bekerja sama dengan ahli. Penuntut umum perlu mampu menghubungkan bukti teknis dengan unsur delik, sedangkan hakim membutuhkan pemahaman dasar mengenai cara kerja sistem generatif agar dapat menilai argumentasi para pihak secara kritis. Peningkatan kapasitas ini tidak berarti seluruh aparat harus menjadi ahli pemrograman, tetapi mereka harus memahami konsep teknis yang relevan dengan pembuktian.

Kerja sama lintas lembaga juga diperlukan karena *deepfake* dapat melibatkan penyedia layanan yang berada di yurisdiksi berbeda. Mekanisme pertukaran informasi, permintaan preservasi data, dan kerja sama internasional menjadi penting ketika bukti berada pada server atau *platform* asing. Dalam konteks ini, pendekatan hukum nasional perlu tetap memperhatikan aturan mengenai yurisdiksi, perlindungan data, dan hak asasi manusia. Penegakan hukum yang efektif harus berjalan bersama dengan perlindungan terhadap privasi dan kebebasan berekspresi.

Lebih lanjut, pemerintah perlu mendorong standardisasi teknis. Pelabelan AI akan kurang efektif apabila setiap penyedia menggunakan format yang tidak kompatibel. Penggunaan standar *provenance* dan *watermarking* yang dapat diverifikasi lintas platform dapat meningkatkan kemampuan pelacakan. Namun, standar tersebut perlu diuji terhadap risiko manipulasi, penghapusan metadata, dan perubahan format. Karena itu, watermarking seharusnya diposisikan sebagai salah satu lapisan pembuktian, bukan satu-satunya metode untuk menentukan keaslian konten.

Tantangan terhadap Kebebasan Berekspresi dan Netralitas Teknologi

Rekonstruksi hukum *deepfake* juga harus memperhatikan kebebasan berekspresi. Tidak semua konten sintetis merupakan konten ilegal. Satire, parodi, kritik politik, karya

seni, pendidikan, penelitian, dan hiburan dapat menggunakan teknologi manipulasi secara sah. Apabila hukum memberikan definisi terlalu luas dan menjadikan setiap *deepfake* sebagai tindak pidana, regulasi dapat menimbulkan *chilling effect*. Masyarakat dapat menghindari penggunaan teknologi yang sebenarnya sah karena khawatir terkena sanksi.

Oleh karena itu, unsur tujuan, konteks, kerugian, dan kesalahan perlu diperhatikan. Pelabelan dapat menjadi solusi yang lebih proporsional untuk konten yang tidak ilegal tetapi berpotensi menimbulkan kebingungan. Dengan label yang jelas, masyarakat memperoleh informasi bahwa suatu materi merupakan hasil sintetis tanpa harus melarang keberadaannya. Pendekatan tersebut lebih sesuai dengan prinsip bahwa hukum pidana merupakan ultimum remedium dan harus digunakan terhadap perbuatan yang benar-benar memenuhi tingkat ketercelaan yang layak dipidana.

Prinsip netralitas teknologi juga harus dipertahankan dengan pengertian yang tepat. Teknologi tidak boleh dianggap bersalah hanya karena memiliki potensi penyalahgunaan. Namun, netralitas tidak berarti pengembang dan penyedia bebas dari seluruh kewajiban. Ketika suatu risiko dapat diperkirakan, dapat dikurangi dengan langkah yang wajar, dan berada dalam kendali aktor tertentu, kewajiban kehati-hatian dapat dibebankan. Dengan demikian, keseimbangan antara inovasi dan perlindungan hukum tidak dicapai dengan membebaskan teknologi dari tanggung jawab, tetapi dengan mengaitkan tanggung jawab pada perilaku manusia dan organisasi yang mengendalikan teknologi.

Agenda Pembaruan Hukum ke Depan

Dalam jangka panjang, Indonesia memerlukan kerangka tata kelola AI yang dapat menghubungkan regulasi teknologi dengan hukum pidana. Kerangka tersebut setidaknya mencakup klasifikasi risiko, kewajiban transparansi, standar pengujian keamanan, mekanisme pelaporan, kewajiban preservasi bukti, dan pengaturan tanggung jawab berdasarkan peran. Hukum pidana kemudian berfungsi sebagai bagian dari sistem tersebut, terutama terhadap penggunaan teknologi yang dilakukan dengan kesengajaan atau kelalaian yang memenuhi unsur tindak pidana.

Pembaharuan juga perlu dilakukan secara adaptif. Teknologi generatif berubah cepat, sehingga norma yang terlalu bergantung pada nama aplikasi atau jenis algoritma tertentu akan mudah tertinggal. Norma sebaiknya berfokus pada fungsi, risiko, tingkat kendali, dan akibat. Dengan pendekatan berbasis fungsi, perubahan dari satu model AI ke model lain tidak selalu memerlukan perubahan norma dasar. Regulasi dapat tetap relevan meskipun teknologi berkembang.

Akhirnya, keberhasilan rekonstruksi dapat diukur bukan hanya dari banyaknya pelaku yang dipidana, tetapi dari kemampuan sistem hukum mencegah penyalahgunaan, memberikan perlindungan cepat kepada korban, menghasilkan pembuktian yang dapat dipercaya, dan menjaga ruang penggunaan teknologi yang sah. Tujuan tersebut menuntut integrasi antara hukum pidana materiil, hukum acara, tata kelola teknologi, kebijakan platform, serta pendidikan literasi digital masyarakat. Dalam perspektif

tersebut, *deepfake* tidak semata-mata merupakan persoalan teknologi, melainkan persoalan tata kelola risiko dan pembagian tanggung jawab dalam masyarakat digital.

KESIMPULAN

Penelitian ini menyimpulkan bahwa konstruksi hukum pidana Indonesia belum sepenuhnya mampu menjangkau kompleksitas kejahatan *deepfake* berbasis kecerdasan buatan generatif, terutama ketika proses pembuatan dan penyebaran konten melibatkan beberapa aktor dengan tingkat kendali yang berbeda. Kesenjangan tersebut terlihat pada kesulitan menghubungkan kesalahan dengan luaran yang dihasilkan melalui sistem generatif serta pada keterbatasan doktrin pertanggungjawaban yang masih berorientasi pada hubungan langsung antara pelaku, perbuatan, dan akibat. Sebagai rekonstruksi, penelitian ini menawarkan model pertanggungjawaban pidana berbasis peran yang membedakan tanggung jawab berdasarkan desain sistem, penyediaan atau penerapan sistem, dan penggunaan sistem. Pengembang dapat dimintai pertanggungjawaban apabila terdapat kesalahan dalam keputusan desain dan pengendalian risiko; penyedia platform dapat dimintai pertanggungjawaban sesuai tingkat kendali dan kegagalannya memenuhi kewajiban pengawasan; sedangkan pengguna akhir tetap menjadi pelaku utama apabila dengan sengaja menggunakan sistem untuk menghasilkan atau menyebarkan konten yang dilarang. Model tersebut perlu didukung pelabelan konten dan penandaan digital untuk memperkuat keterlacakan serta pembuktian forensik elektronik. Pemerintah dan pembentuk undang-undang perlu mengharmonisasikan pengaturan pidana, tata kelola AI, dan pembuktian elektronik dengan tetap mempertahankan asas kesalahan, proporsionalitas, kepastian hukum, perlindungan korban, dan ruang bagi inovasi teknologi. undang untuk mengadopsi formulasi akuntabilitas tripartit ini ke dalam amandemen regulasi siber nasional, yang didukung oleh kewajiban teknis *digital watermarking* dan integrasi forensik metadata dalam pembuktian KUHAP guna memperkuat ketahanan institusi hukum nasional.

REFERENSI

Chesney, R., & Citron, D. K. (2019). Deepfakes and the new disinformation threat. *California Law Review*, 107(6), 1753–1822.

Citron, D. K. (2019). Sexual privacy. *Yale Law Journal*, 128(7), 1870–1960.

Cole, S. (2022). Deepfakes and the legal landscape: Synthetic media and the limits of current law. *Journal of Free Speech Law*, 2(1), 115–150.

Edwards, L. (2022). Regulating AI in Europe: From the AI Act to Product Liability Directive. *Computer Law & Security Review*,47, Article 105724. https://doi.org/10.1016/j.clsr.2022.105724

European Parliament. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L 168, 1–110.

Floridi, L. (2023). AI ethics and governance: An architectural approach. *Philosophy & Technology*, 36(2), Article 201. https://doi.org/10.1007/s13347-023-00620-3

Franke, U. (2022). Digital forensics for synthetic media: Technical and legal dimensions of deepfake detection. *Forensic Science International: Digital Investigation*, 40, Article 301318.

Franks, M. A., & Waldman, A. E. (2020). Sex, lies, and videotape: Deepfakes and the urgent need for accountability. *Harvard Journal of Law & Technology*, 33(2), 511–562.

Harris, D. (2023). Deepfake technology and criminal responsibility: Applying mens rea to generative AI systems. *Artificial Intelligence and Law*, 31(4), 689–712.

Hasibuan, A. Z., & Yusuf, M. (2024). Perkembangan hukum pidana siber di Indonesia pasca pelaksanaan Undang-Undang ITE baru. *Jurnal Pembangunan Hukum Indonesia*, 6(2), 180–198.

Hildebrandt, M. (2021). Law for algorithmic decision-making: Embodiment, intelligence, and liability. *Law, Innovation and Technology*, 13(1), 1–25.

Kaminski, M. E. (2024). The EU Artificial Intelligence Act: A risk-based approach to regulatory compliance. *European Law Journal*, 29(3), 310–335.

Karnow, C. E. A. (2022). The application of criminal liability to autonomous artificial intelligence. *Santa Clara High Technology Law Journal*, 38(2), 145–182.

Kirchengast, T. (2023). Deepfakes, image-based sexual abuse, and criminal justice reform. *Criminology & Criminal Justice*, 23(5), 712–730.

Koops, B.-J. (2020). Should ICT regulation be technology-neutral? The case of deepfake synthetic media. *Duke Law & Technology Review*, 18(1), 101–135.

Kusuma, H., & Adiputri, R. (2023). Tanggung jawab hukum perantara platform digital atas konten deepfake berdasarkan UU ITE. Jurnal Hukum & Pembangunan, 53(4), 890–908. Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news: Addressing misinformation and deepfakes. *Science*, 359(6380), 1094–1096. https://doi.org/10.1126/science.aoo2998

Lessig, L. (2024). Code and other laws of cyberspace in the era of generative AI. *Harvard Law Review*, 137(5), 1401–1435.

Middleton, B. (2022). Corporate criminal liability for AI-generated crimes: Re-evaluating vicarious liability. *Criminal Law Review*, 2022(8), 620–641.

Niko, A., & Wuryanto, T. (2023). Penegakan hukum pidana terhadap tindak pidana manipulasi informasi melalui deepfake technology. *Jurnal Pembaharuan Hukum*, 10(2), 215–230.

Pratama, B. D. (2023). Pelabelan konten dan watermarking digital sebagai alat bukti forensik elektronik dalam tindak pidana siber. *Jurnal Penegakan Hukum dan Keadilan*, 4(2), 112–131.

Radley, T., & Nissenbaum, H. (2023). Contextual integrity and synthetic media: Regulating AI deepfakes.*Information, Communication & Society*, 26(4), 845–862.

Rahardjo, S., & Santoso, B. (2023). Rekonstruksi asas tiada pidana tanpa kesalahan dalam kejahatan berbasis teknologi otonom. *Jurnal Masalah-Masalah Hukum*, 52(3), 301–315.

Richards, N. M., & King, J. H. (2023). Big data ethics and generative AI accountability. *Wake Forest Law Review*, 58(2), 411–440.

Scharre, P. (2023). Autonomous systems and criminal intent: Rethinking the doctrine of mens rea. *Harvard National Security Journal*, 14(1), 89–128.

Simanjuntak, A., & Farida, M. (2023). Analisis yuridis doktrin vicarious liability pada penyedia jasa platform digital atas penyebaran konten manipulatif. *Jurnal Hukum Ius Quia Iustum*, 30(2), 334–355.

Suhariyanto, B. (2021). Tantangan penegakan hukum pidana terhadap kejahatan siber berbasis kecerdasan buatan. *Jurnal Hukum Respublica*, 21(1), 45–60.

Sutherland, E. (2023). Governance of generative artificial intelligence and deepfake technologies. *Telecommunications Policy*, 47(8), Article 102590. https://doi.org/10.1016/j.telpol.2023.10259 0

Tzanou, M. (2023). European data protection and generative AI: Deepfakes, fundamental rights, and regulatory responses. *Computer Law & Security Review*, 49, Article 105810.

Wintgens, L. J., & Van den Brande, S. (2023). Attributing criminal intent to developers of generative algorithms: A European perspective. *European Journal of Crime, Criminal Law and Criminal Justice*, 31(3), 245–270.

Zhao, J. (2023). China's regulatory approach to deep synthesis: Identity authentication and algorithmic governance. *Asian Journal of Law and Society*, 10(2), 380–402.