

Hukum Siber Di Era Digital: Rekonstruksi Penanggulangan Kejahatan Deepfake Dalam Perspektif Hukum Pidana Indonesia Dan Maqāṣid Al-Syarī'ah

Muhammad Hafiduddin¹, Afif Hasbullah², Ahmad Munir³, Ahmad Zeeya Ulya Eldavi⁴, Desi Lusiawati⁵

^{1,2,3,4,5} Program Doktorat Universitas Islam Darul Ulum (UNISDA), Lamongan
Sekolah Tinggi Ilmu Hukum IBLAM, Jakarta; Universitas Ngudi Waluyo, Semarang
Email: 2505301006@mhs.unisda.ac.id; afif@unisda.ac.id; ahmadmunir@unisda.ac.id;
zeeyaulyaeldavi@gmail.com; desizee26sh@gmail.com



e-ISSN: 2964-0962

SEIKAT: Jurnal Ilmu Sosial, Politik dan Hukum

<https://ejournal.45mataram.ac.id/index.php/seikat>

Vol. 5, No. 4, Agustus 2026

Page: 1325-1336

Available at:

<https://ejournal.45mataram.ac.id/index.php/seikat/article/view/3370>

DOI:

<https://doi.org/10.55681/seikat.v5i4.3370>

Article History:

Received: 13-07-2026

Revised: 04-08-2026

Accepted: 04-08-2026

Abstract : *The rapid development of generative artificial intelligence has enabled deepfake technology to create synthetic audio, images, and videos that closely resemble authentic content. Its malicious use threatens identity, privacy, reputation, property, public trust, and the integrity of digital evidence. This study examines the legal response to deepfake crime by integrating Indonesian criminal law with the ethical-normative framework of maqāṣid al-sharī'ah. It employs normative legal research using statutory, conceptual, and Islamic-law approaches. Primary legal materials include the Electronic Information and Transactions Law as last amended in 2024, the 2023 National Criminal Code as adjusted by Law Number 1 of 2026, the Personal Data Protection Law, and the Sexual Violence Crime Law. The study finds that Indonesian law can address particular consequences of deepfakes through provisions on electronic-data manipulation, fraud, defamation, pornography, electronic sexual violence, and unlawful use of personal data. Nevertheless, these provisions remain dispersed, consequence-oriented, and insufficiently responsive to synthetic identity manipulation, rapid content removal, victim restoration, evidentiary attribution, and platform accountability. Islamic law classifies harmful deepfake conduct as ja'irah ta'zir because it embodies falsehood, deception, fraud, defamation, and harm. The study proposes a complementary reconstruction model consisting of a precise legal definition, tiered criminalization based on intent and harm, rapid victim-centred remedies, digital-evidence standards, and preventive duties grounded in tabayyun and the protection of intellect, life, honour, lineage, and property.*

Keywords: *Deepfake; Cybercrime; Criminal Law; Islamic Law; Maqāṣid al-Sharī'ah.*

Abstrak : Perkembangan pesat kecerdasan buatan generatif memungkinkan teknologi deepfake menghasilkan audio, gambar, dan video sintetis yang menyerupai konten autentik. Penyalahgunaannya mengancam identitas, privasi, reputasi, harta, kepercayaan publik, dan integritas alat bukti digital. Penelitian ini menganalisis penanggulangan kejahatan deepfake melalui integrasi hukum pidana Indonesia dengan kerangka etis-normatif maqāṣid al-syarī'ah. Penelitian menggunakan metode hukum

normatif dengan pendekatan perundang-undangan, konseptual, dan hukum Islam. Bahan hukum primer meliputi UU Informasi dan Transaksi Elektronik sebagaimana terakhir diubah pada 2024, KUHP Nasional Tahun 2023 sebagaimana disesuaikan melalui UU Nomor 1 Tahun 2026, UU Pelindungan Data Pribadi, dan UU Tindak Pidana Kekerasan Seksual. Hasil penelitian menunjukkan bahwa hukum Indonesia dapat menjangkau akibat tertentu dari deepfake melalui ketentuan manipulasi data elektronik, penipuan, pencemaran, pornografi, kekerasan seksual berbasis elektronik, dan penggunaan data pribadi secara melawan hukum. Namun, pengaturannya masih tersebar, berorientasi pada akibat, dan belum memadai dalam merespons manipulasi identitas sintetis, penghapusan konten secara cepat, pemulihan korban, atribusi pelaku, serta tanggung jawab platform. Hukum Islam mengategorikan deepfake yang merugikan sebagai jarimah ta'zir karena mengandung kebohongan, penipuan, kecurangan, fitnah, dan kemudharatan. Penelitian ini menawarkan model rekonstruksi komplementer berupa definisi hukum yang presisi, kriminalisasi bertingkat berdasarkan niat dan akibat, pemulihan korban yang cepat, standar pembuktian digital, serta kewajiban pencegahan berbasis tabayun dan perlindungan akal, jiwa, kehormatan, keturunan, dan harta.

Kata Kunci: Deepfake; Kejahatan Siber; Hukum Pidana; Hukum Islam; Maqāṣid al-Syarī'ah.

PENDAHULUAN

Transformasi digital telah mengubah cara manusia membentuk identitas, berkomunikasi, melakukan transaksi, dan memproduksi pengetahuan. Pada tahap mutakhir, kecerdasan buatan generatif tidak lagi sekadar mengolah informasi, tetapi mampu menghasilkan media sintetis yang menyerupai realitas. Salah satu bentuknya adalah *deepfake*, yaitu konten audio-visual yang dibuat atau dimanipulasi dengan teknik pembelajaran mesin sehingga wajah, suara, gerak, atau ekspresi seseorang tampak melakukan atau menyampaikan sesuatu yang sebenarnya tidak pernah terjadi. Istilah ini mula-mula diasosiasikan dengan penggunaan *Generative Adversarial Networks* (GAN), meskipun perkembangan terbaru juga memanfaatkan *autoencoder*, model difusi, dan sistem kloning suara (Goodfellow et al., 2014; Mirsky & Lee, 2021).

Deepfake bersifat *dual use*. Teknologi yang sama dapat digunakan untuk pendidikan, industri kreatif, pelestarian budaya, aksesibilitas, dan hiburan, tetapi juga dapat dipakai untuk menyamarkan identitas, memalsukan persetujuan, membuat pornografi non-konsensual, melakukan pemerasan, menipu korban melalui rekaman suara sintetis, mencemarkan nama baik, serta memproduksi disinformasi politik. Realisme yang tinggi, biaya produksi yang terus menurun, skala distribusi digital, dan anonimitas pelaku membuat deepfake berbeda dari manipulasi media konvensional. Korban tidak hanya kehilangan kendali atas data atau citranya, tetapi juga menghadapi kesulitan membuktikan bahwa konten yang beredar tidak autentik (Chesney & Citron, 2019; Kietzmann et al., 2020). Ancaman tersebut melahirkan dua krisis sekaligus. Pertama, krisis keaslian, karena masyarakat sulit membedakan konten autentik dan sintetis. Kedua, krisis kepercayaan, karena keberadaan deepfake dapat digunakan untuk menyangkal bukti yang sebenarnya autentik, suatu keadaan yang dikenal sebagai *liar's dividend*. Dengan demikian, bahaya deepfake bukan hanya terletak pada beredarnya konten palsu, melainkan juga pada melemahnya kemampuan hukum, media, dan publik untuk mempercayai bukti audio-visual (Chesney & Citron, 2019).

Dalam konteks Indonesia, penanggulangan deepfake belum bertumpu pada satu undang-undang khusus. Perbuatan dan akibatnya tersebar dalam rezim hukum informasi elektronik, KUHP Nasional, pelindungan data pribadi, pornografi, dan tindak pidana kekerasan seksual. Pasal 35 UU ITE, misalnya, melarang manipulasi, penciptaan, perubahan, penghilangan, atau kerusakan informasi elektronik dengan tujuan agar dianggap seolah-olah data autentik. Ketentuan ini merupakan norma yang paling dekat dengan proses pembuatan deepfake, tetapi tidak menggunakan istilah deepfake, tidak membedakan penggunaan yang sah dan yang merugikan, serta belum mengatur secara komprehensif persetujuan subjek, pelabelan konten sintetis, dan pemulihan korban. Lanskap hukum tersebut menjadi lebih kompleks setelah berlakunya KUHP Nasional pada 2 Januari 2026. UU Nomor 1 Tahun 2024 telah menentukan bahwa sejumlah

ketentuan UU ITE hanya berlaku sampai KUHP Nasional diberlakukan. Selanjutnya, UU Nomor 1 Tahun 2026 tentang Penyesuaian Pidana menata kembali keterkaitan ketentuan pidana di luar KUHP dan mengalihkan sejumlah rujukan tindak pidana informatika, kesusilaan, penghinaan, dan pornografi ke dalam KUHP. Perubahan ini penting bagi analisis deepfake karena norma yang sebelumnya dibaca secara terpisah kini harus ditafsirkan dalam relasi antara UU ITE, KUHP Nasional, UU Pelindungan Data Pribadi, dan undang-undang sektoral lainnya.

Kajian hukum terhadap deepfake di Indonesia umumnya berfokus pada kemungkinan penerapan UU ITE, urgensi kriminalisasi khusus, atau pertanggungjawaban pidana pelaku. Pendekatan tersebut penting, tetapi sering berhenti pada pemetaan pasal dan belum merumuskan desain regulasi yang membedakan tahap pembuatan, penggunaan, penyebaran, dan akibat; belum mengintegrasikan mekanisme pemulihan korban; serta belum menempatkan pembuktian dan tata kelola platform sebagai bagian dari kebijakan hukum pidana. Di sisi lain, pendekatan hukum Islam sering kali terbatas pada pelabelan deepfake sebagai kebohongan atau penipuan tanpa menghubungkannya secara operasional dengan kebutuhan rekonstruksi hukum nasional.

Hukum Islam menyediakan kerangka etik yang relevan melalui larangan *al-kadzib* (kebohongan), *tadlis* (penyamaran atau penipuan), *ghisy* (kecurangan), *ifk* dan *buhtān* (dusta atau tuduhan palsu), serta prinsip *lā ḍarara wa lā ḍirār* yang menolak perbuatan merugikan. Prinsip *tabayyun* dalam QS. al-Hujurat [49]: 6 juga memberikan dasar epistemik untuk memverifikasi informasi sebelum menerimanya atau menyebarkannya. Melalui maqāṣid al-syarī'ah, dampak deepfake dapat dianalisis terhadap perlindungan akal, jiwa, kehormatan, keturunan, dan harta, sehingga hukum tidak hanya berorientasi pada penghukuman, tetapi juga pada pencegahan kemudaratatan dan pemulihan martabat korban (al-Syāṭibī, 1997; Kamali, 2008).

Kebaruan penelitian ini terletak pada rekonstruksi model penanggulangan deepfake yang menghubungkan empat lapis analisis: karakter teknologi, pemetaan hukum pidana Indonesia setelah berlakunya KUHP Nasional dan UU Nomor 1 Tahun 2026, klasifikasi perbuatan menurut hukum Islam, serta desain kebijakan yang berorientasi pada korban dan pembuktian digital. Berdasarkan kerangka tersebut, penelitian menjawab empat persoalan: bagaimana karakteristik dan tipologi penyalahgunaan deepfake; sejauh mana hukum pidana Indonesia dapat menjangkanya; bagaimana hukum Islam dan maqāṣid al-syarī'ah menilainya; serta bagaimana model regulasi yang lebih koheren dapat dirumuskan.

METODE PENELITIAN

Penelitian ini merupakan penelitian hukum normatif yang menempatkan hukum sebagai norma, asas, konsep, dan sistem argumentasi. Pendekatan yang digunakan meliputi pendekatan perundang-undangan (*statute approach*), pendekatan konseptual (*conceptual approach*), dan pendekatan hukum Islam. Pendekatan perundang-undangan digunakan untuk menelaah hubungan antara UU ITE, KUHP Nasional, UU Pelindungan Data Pribadi, UU Tindak Pidana Kekerasan Seksual, serta UU Nomor 1 Tahun 2026 tentang Penyesuaian Pidana. Pendekatan konseptual digunakan untuk menganalisis deepfake, identitas digital, autentisitas bukti elektronik, pertanggungjawaban pidana, dan pemulihan korban. Pendekatan hukum Islam digunakan untuk mengkaji jarīmah ta'zīr, prinsip tabayyun, kaidah pencegahan kemudaratatan, dan maqāṣid al-syarī'ah.

Bahan hukum primer terdiri atas peraturan perundang-undangan Indonesia, Al-Qur'an, hadis dan kaidah fikih yang relevan. Bahan hukum sekunder meliputi buku metodologi hukum, literatur maqāṣid al-syarī'ah, artikel ilmiah mengenai deepfake, disinformasi, keamanan siber, dan pembuktian digital. Pengumpulan bahan hukum dilakukan melalui studi kepustakaan dan dokumentasi. Analisis dilakukan secara kualitatif dengan teknik interpretasi gramatikal, sistematis,

dan teleologis, kemudian dilanjutkan dengan evaluasi preskriptif untuk merumuskan model regulasi. Konsistensi argumentasi diuji dengan membedakan norma yang mengatur proses pembuatan, norma yang mengatur penyebaran, norma yang mensyaratkan akibat tertentu, dan norma yang berorientasi pada perlindungan korban (Marzuki, 2016).

HASIL DAN PEMBAHASAN

1. Deepfake sebagai Bentuk Baru Kejahatan Siber Berbasis Kecerdasan Buatan

Secara konseptual, deepfake merupakan bagian dari *synthetic media*, yaitu media yang keseluruhan atau sebagian unsurnya dihasilkan, diganti, atau dimodifikasi oleh sistem kecerdasan buatan. Fokus hukum tidak seharusnya hanya pada teknologi tertentu, sebab teknik pembuatan akan terus berubah. Definisi hukum yang terlalu sempit pada GAN akan cepat tertinggal ketika pelaku menggunakan model difusi, kloning suara, atau sistem multimodal. Oleh karena itu, unsur yang lebih stabil adalah adanya representasi sintetis atau manipulatif yang secara material menyerupai orang, peristiwa, atau pernyataan nyata dan berpotensi menyesatkan pihak lain mengenai autentisitasnya.

Deepfake memiliki sekurang-kurangnya enam karakteristik kriminologis. Pertama, tingkat realisme tinggi yang membuat korban dan publik sulit melakukan verifikasi kasatmata. Kedua, skalabilitas, karena satu model dapat menghasilkan banyak konten dengan biaya rendah. Ketiga, kecepatan penyebaran yang membuat kerugian reputasi terjadi sebelum proses klarifikasi atau hukum berjalan. Keempat, *attribution gap*, yaitu kesenjangan antara kemampuan membuat konten dan kemampuan mengidentifikasi pembuatnya. Kelima, sifat lintas batas, sebab model, server, platform, pelaku, dan korban dapat berada di yurisdiksi berbeda. Keenam, ambiguitas penggunaan, karena teknologi yang sah dapat berubah menjadi sarana tindak pidana ketika digunakan tanpa persetujuan atau dengan maksud merugikan.

Berdasarkan tahap perbuatannya, penyalahgunaan deepfake dapat dibedakan menjadi pembuatan, penggunaan, distribusi, dan monetisasi. Pembuatan mencakup pengumpulan data wajah atau suara, pelatihan model, dan penciptaan media sintetis. Penggunaan terjadi ketika konten dipakai untuk menipu, memperoleh akses, memengaruhi keputusan, atau mengancam korban. Distribusi mencakup pengunggahan, transmisi, atau penyediaan akses. Monetisasi terjadi ketika pelaku memperoleh keuntungan langsung melalui penipuan, pemerasan, perdagangan konten, atau peningkatan trafik. Perbedaan ini penting karena tidak semua tahap menghasilkan akibat yang sama dan tidak semua pembuat memiliki niat yang identik dengan penyebar.

Berdasarkan objek serangan dan tujuan pelaku, tipologi utama deepfake meliputi: (a) pemalsuan identitas untuk penipuan finansial atau akses sistem; (b) pornografi non-konsensual dan kekerasan seksual berbasis elektronik; (c) pencemaran nama baik dan pemerasan; (d) disinformasi politik atau sosial; (e) pemalsuan bukti dan rekayasa narasi peristiwa; serta (f) penipuan korporasi melalui imitasi suara atau wajah pejabat. Setiap tipologi memiliki korban, unsur kesalahan, akibat, dan kebutuhan pemulihan yang berbeda. Karena itu, kriminalisasi tunggal yang menghukum semua deepfake dengan cara yang sama berisiko melanggar prinsip proporsionalitas.

Dampak deepfake bersifat multidimensional. Pada tingkat individual, korban dapat mengalami kehilangan kendali atas identitas, trauma, stigma, kehilangan pekerjaan, atau kerugian finansial. Pada tingkat sosial, deepfake mengikis kepercayaan terhadap media dan memperbesar beban verifikasi. Pada tingkat kelembagaan, konten sintetis dapat mengganggu proses pemilu, pasar, layanan keuangan, dan penegakan hukum. Pada tingkat pembuktian, pihak yang berkepentingan dapat mempersoalkan autentisitas setiap rekaman, sehingga penegak hukum harus

membangun standar integritas bukti yang lebih kuat daripada sekadar pemeriksaan visual (Vaccari & Chadwick, 2020; Mirsky & Lee, 2021).

2. Konstruksi dan Keterbatasan Hukum Pidana Indonesia terhadap Deepfake

Hukum pidana Indonesia tidak mengenal delik bernama “deepfake”. Penegakan hukum dilakukan dengan menghubungkan perbuatan konkret dengan unsur delik yang telah tersedia. Pendekatan ini memungkinkan respons awal, tetapi menghasilkan fragmentasi karena satu rangkaian deepfake dapat menyentuh beberapa undang-undang sekaligus. Misalnya, pelaku dapat mengumpulkan data biometrik tanpa hak, menciptakan video seolah-olah autentik, menyebarkannya untuk merusak reputasi, lalu meminta uang agar konten dihapus. Setiap tahap tunduk pada norma yang berbeda dan menuntut pembuktian unsur yang berbeda pula.

Pasal 35 UU ITE merupakan norma paling langsung terhadap proses penciptaan deepfake. Pasal tersebut melarang manipulasi, penciptaan, perubahan, penghilangan, atau perusakan Informasi Elektronik dan/atau Dokumen Elektronik dengan tujuan agar dianggap seolah-olah data yang autentik. Dalam perkara deepfake, unsur “manipulasi atau penciptaan” dapat meliputi pertukaran wajah, kloning suara, sinkronisasi bibir, atau penyusunan adegan sintetis. Unsur tujuan agar dianggap autentik membedakan deepfake menyesatkan dari parodi yang jelas diberi label. Namun, norma ini belum secara eksplisit menempatkan ketiadaan persetujuan, penggunaan identitas, atau risiko kerugian sebagai unsur pemberat.

UU ITE juga dapat menjangkau akibat tertentu. Pasal 28 ayat (1) relevan apabila informasi bohong atau menyesatkan mengakibatkan kerugian materiel bagi konsumen dalam transaksi elektronik. Cakupannya terbatas karena mensyaratkan hubungan konsumen dan transaksi elektronik. Pasal 27B dapat digunakan terhadap pemerasan atau pengancaman melalui informasi elektronik. Sementara itu, perbuatan akses ilegal dan gangguan sistem tetap dapat berlaku ketika deepfake digunakan untuk melewati autentikasi atau mengambil alih akun. Dengan demikian, UU ITE tidak menyediakan satu delik menyeluruh, melainkan kumpulan norma berdasarkan modus dan akibat.

Sejak 2 Januari 2026, KUHP Nasional menjadi rujukan utama bagi sejumlah perbuatan yang sebelumnya diatur dalam UU ITE atau undang-undang lain. Pencemaran dan fitnah diatur dalam Pasal 433 dan Pasal 434, dengan pemberatan apabila dilakukan melalui sarana teknologi informasi berdasarkan Pasal 441. Penipuan diatur dalam Pasal 492, termasuk penggunaan nama palsu, kedudukan palsu, tipu muslihat, atau rangkaian kata bohong untuk menggerakkan korban menyerahkan barang atau membuat pengakuan utang. Berita bohong yang mengakibatkan kerusakan diatur dalam Pasal 263. Pornografi diatur dalam Pasal 407 sebagaimana disesuaikan melalui UU Nomor 1 Tahun 2026. Ketentuan tersebut dapat menjangkau akibat deepfake, tetapi tetap tidak menyebut manipulasi identitas sintetis sebagai bentuk khusus.

Peralihan norma harus dibaca bersama Pasal II UU Nomor 1 Tahun 2024 dan UU Nomor 1 Tahun 2026. Sejumlah ketentuan UU ITE tentang kesusilaan, pencemaran, kebencian, dan kerusakan memang dirancang berlaku sampai KUHP Nasional efektif. UU Nomor 1 Tahun 2026 kemudian menyesuaikan rujukan dan mencabut beberapa ketentuan pidana di luar KUHP untuk menghindari duplikasi. Konsekuensinya, analisis yuridis setelah 2026 tidak tepat apabila masih semata-mata mengutip UU Nomor 19 Tahun 2016 atau memperlakukan seluruh Pasal 27 dan Pasal 28 UU ITE sebagai norma pidana yang berdiri sendiri.

UU Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi menambah lapisan perlindungan ketika deepfake dibuat dari data yang dapat mengidentifikasi seseorang. Data biometrik merupakan data pribadi spesifik. Representasi wajah, suara, pola gerak, atau karakteristik tubuh yang diproses untuk mengidentifikasi individu dapat menjadi bagian dari pemrosesan data pribadi. Pasal 65 melarang perolehan, pengumpulan, pengungkapan, dan

penggunaan data pribadi orang lain secara melawan hukum, sedangkan Pasal 67 memuat ketentuan pidananya. Keterbatasannya, tidak setiap foto atau rekaman otomatis merupakan data biometrik; penegak hukum harus membuktikan konteks pemrosesan dan tujuan identifikasi.

Untuk deepfake seksual, UU Nomor 12 Tahun 2022 tentang Tindak Pidana Kekerasan Seksual memberikan orientasi korban yang lebih kuat. Pasal 14 mengatur perekaman atau pengambilan gambar bermuatan seksual tanpa persetujuan, transmisi konten seksual di luar kehendak penerima, serta penguntitan elektronik untuk tujuan seksual. Norma ini dapat mendukung penanganan distribusi dan pemerasan seksual, tetapi rumusannya belum secara eksplisit menyebut penciptaan gambar seksual sintetis dengan memasukkan wajah korban ke tubuh orang lain. Kekosongan inilah yang membuat kriminalisasi spesifik terhadap *non-consensual sexual deepfake* tetap diperlukan.

Selain persoalan substansi, penegakan hukum menghadapi masalah prosedural. Pembuktian keaslian memerlukan pemeriksaan metadata, jejak kompresi, artefak model, riwayat unggahan, *hash value*, serta rantai penguasaan bukti. Detektor deepfake tidak boleh diperlakukan sebagai bukti tunggal karena akurasi dapat menurun terhadap model baru atau konten yang telah dikompresi. Atribusi pelaku juga memerlukan korelasi antara perangkat, akun, alamat jaringan, transaksi, dan motif. Tanpa protokol forensik yang terstandar, perkara dapat gagal bukan karena tidak ada kerugian, tetapi karena hubungan antara konten, pelaku, dan akibat tidak terbukti secara meyakinkan.

Kelemahan lain adalah ketidakseimbangan waktu. Konten dapat menyebar dalam hitungan menit, sedangkan permintaan penghapusan, penyidikan, dan putusan pengadilan memerlukan waktu jauh lebih lama. Pendekatan pidana yang hanya menghukum pelaku setelah kerugian meluas tidak cukup. Korban membutuhkan mekanisme pemutusan akses sementara, pelestarian bukti, klarifikasi autentisitas, penghapusan salinan, restitusi, dan pemulihan reputasi. Hal ini menuntut koordinasi antara penyidik, platform, penyelenggara sistem elektronik, lembaga perlindungan data, dan penyedia layanan keuangan.

Tabel 1. Pemetaan Norma Hukum Indonesia terhadap Penyalahgunaan Deepfake

Bentuk Perbuatan	Norma Relevan	Unsur Kunci	Kekuatan	Keterbatasan
Penciptaan konten seolah autentik	Pasal 35 UU ITE jo. Pasal 51 ayat (1)	Manipulasi/penciptaan dengan tujuan dianggap autentik	Paling dekat dengan proses deepfake	Tidak mengatur persetujuan, label, dan tingkat bahaya secara khusus
Penipuan dan kerugian transaksi	Pasal 28 ayat (1) UU ITE; Pasal 492 KUHP	Kebohongan/tipu muslihat dan penyerahan barang atau kerugian materiel	Dapat menjangkau kloning suara atau video untuk penipuan	Pasal 28 ayat (1) terbatas pada konsumen dan transaksi elektronik
Pencemaran atau fitnah digital	Pasal 433–434 jo. Pasal 441 KUHP	Serangan kehormatan melalui tuduhan yang diketahui umum	Mencakup tulisan, gambar, dan	Delik aduan; fokus pada tuduhan dan reputasi,

Bentuk Perbuatan	Norma Relevan	Unsur Kunci	Kekuatan	Keterbatasan
			sarana teknologi informasi	bukan manipulasi identitas
Deepfake seksual	Pasal 407 KUHP; Pasal 14 UU TPKS; UU PDP	Pornografi, transmisi seksual, ketiadaan persetujuan, penggunaan data	Menyediakan pidana dan kerangka hak korban	Penciptaan seksual sintetis belum dirumuskan secara eksplisit
Penyalahgunaan wajah/suara	Pasal 4, 65, dan 67 UU PDP	Pemrosesan data pribadi spesifik secara melawan hukum	Melindungi kontrol subjek atas data pribadi	Perlu pembuktian bahwa data diproses untuk identifikasi individu
Disinformasi yang memicu gangguan publik	Pasal 263 KUHP	Berita bohong yang mengakibatkan kerusuhan	Dapat menjangkau akibat sosial yang serius	Tidak berlaku untuk semua disinformasi; mensyaratkan akibat kerusuhan

Sumber: Diolah penulis berdasarkan peraturan perundang-undangan yang berlaku pada Juli 2026.

3. Perspektif Hukum Islam dan Analisis Maqāṣid al-Syarī'ah

Dalam hukum pidana Islam, tidak setiap perbuatan terlarang harus disebut secara literal dalam Al-Qur'an atau hadis. Perbuatan baru yang mengandung kemudharatan dapat dimasukkan ke dalam jarīmah ta'zīr, yaitu tindak pelanggaran yang jenis dan sanksinya ditetapkan oleh otoritas berdasarkan kemaslahatan, tingkat bahaya, dan kondisi sosial. Deepfake yang digunakan secara sah, diberi persetujuan, dan tidak menyesatkan tidak otomatis merupakan jarīmah. Yang menjadi objek larangan adalah penggunaan yang mengandung kebohongan, penipuan, pelanggaran kehormatan, pemerasan, atau kerugian.

Unsur *al-kadzib* muncul ketika pelaku secara sadar menghasilkan atau menyebarkan representasi palsu sebagai kebenaran. Unsur *tadlīs* terdapat pada menyembunyikan fakta atau pencampuran yang benar dengan yang palsu sehingga pihak lain keliru. Unsur *ghisy* tampak pada kecurangan untuk memperoleh keuntungan. Deepfake yang menuduhkan perbuatan tercela kepada korban dapat mendekati *buhtān* atau *ifk*, terutama apabila konten dibuat untuk merusak kehormatan. Ketika deepfake digunakan untuk mengambil harta, perbuatannya juga bertentangan dengan larangan memakan harta secara batil.

QS. al-Ḥujurāt [49]: 6 memerintahkan tabayyun ketika menerima berita yang berpotensi merugikan orang lain. Prinsip ini bukan sekadar anjuran etis kepada penerima informasi, tetapi dapat dikembangkan sebagai norma pencegahan bagi platform, media, dan pejabat publik. Verifikasi sumber, pelabelan konten sintetis, pemeriksaan silang, dan penundaan distribusi konten

berisiko tinggi merupakan bentuk operasional tabayyun di ruang digital. QS. al-Isrā' [17]: 36 juga melarang mengikuti sesuatu tanpa pengetahuan, sehingga penyebar tidak dapat melepaskan tanggung jawab hanya dengan alasan menerima konten dari pihak lain.

Prinsip *lā ḍarara wa lā ḍirār* menegaskan bahwa bahaya tidak boleh ditimbulkan dan tidak boleh dibalas dengan bahaya yang lain. Dalam kebijakan deepfake, prinsip ini mendukung pencegahan, penghentian distribusi, dan pemulihan korban. Kaidah *al-ḍarar yuzāl* (kemudaratan harus dihilangkan) mengarahkan negara untuk menyediakan mekanisme penghapusan konten dan koreksi publik. Kaidah *dar'u al-maḥṣid muqaddam 'alā jalb al-maṣāliḥ* menuntut pembatasan terhadap penggunaan deepfake yang berisiko tinggi, tanpa menutup manfaat teknologi secara keseluruhan.

Maqāṣid al-syarī'ah memberi dasar untuk menilai jenis kerugian. Perlindungan akal (*ḥifẓ al-'aql*) terganggu oleh disinformasi yang merusak kemampuan publik mengambil keputusan rasional. Perlindungan jiwa (*ḥifẓ al-nafs*) terancam ketika korban mengalami intimidasi, tekanan psikologis, atau risiko kekerasan. Perlindungan harta (*ḥifẓ al-māl*) dilanggar melalui penipuan dan pemerasan. Perlindungan keturunan dan keluarga (*ḥifẓ al-nasab*) terganggu oleh konten seksual sintesis yang merusak relasi dan martabat keluarga. Perlindungan kehormatan (*ḥifẓ al-'ird*), yang dalam pemikiran kontemporer dipandang berkaitan erat dengan jiwa dan keturunan, menjadi dimensi paling langsung pada pencemaran dan pornografi non-konsensual (al-Syātibī, 1997; Ibn 'Āshūr, 2006).

Analisis maqāṣid juga mencegah pendekatan yang semata-mata represif. Tujuan hukum tidak tercapai apabila pelaku dipidana tetapi konten tetap beredar, korban tetap distigma, dan platform tidak memperbaiki sistem. Kemaslahatan menuntut keseimbangan antara kepastian hukum, kebebasan berekspresi, inovasi, hak atas privasi, dan perlindungan korban. Karena itu, pengecualian untuk satire, seni, penelitian, pendidikan, dan kepentingan publik dapat diakui sepanjang terdapat transparansi yang memadai dan tidak menimbulkan penyesatan atau eksploitasi.

Integrasi hukum Islam ke dalam hukum nasional tidak berarti membentuk sistem penghukuman paralel atau menerapkan sanksi agama secara langsung. Nilai maqāṣid berfungsi sebagai dasar filosofis dan parameter evaluasi yang dapat memperkuat asas kemanusiaan, keadilan, kemanfaatan, dan perlindungan hak konstitusional. Pendekatan komplementer ini sesuai dengan karakter hukum nasional yang plural dan memungkinkan nilai etis Islam berkontribusi tanpa menghilangkan prinsip legalitas, pembuktian, dan prosedur hukum positif.

4. Rekonstruksi Model Penanggulangan Deepfake di Indonesia

Rekonstruksi regulasi harus dimulai dari definisi yang netral teknologi. Deepfake perlu dirumuskan sebagai informasi elektronik atau dokumen elektronik berupa audio, gambar, video, atau kombinasi media yang secara material dibuat atau dimanipulasi dengan sistem otomatis untuk menampilkan identitas, ucapan, tindakan, atau peristiwa seolah-olah autentik. Definisi harus membedakan konten sintesis yang sah dari deepfake berbahaya. Unsur pidana sebaiknya mensyaratkan kesengajaan atau pengetahuan mengenai sifat sintesis, ketiadaan hak atau persetujuan, tujuan menyesatkan atau mengeksploitasi, dan/atau timbulnya risiko kerugian yang nyata.

Kriminalisasi perlu disusun secara bertingkat. Tingkat pertama berupa pelanggaran administratif atau kewajiban label bagi konten sintesis yang digunakan secara komersial atau untuk komunikasi publik. Tingkat kedua berupa delik terhadap pembuatan atau penyebaran tanpa persetujuan yang menimbulkan kerugian reputasi atau privasi. Tingkat ketiga berupa delik berat

apabila deepfake digunakan untuk pornografi non-konsensual, penipuan, pemerasan, manipulasi pemilu, pemalsuan bukti, atau menimbulkan kekerasan. Pemberatan dapat dikenakan apabila korban anak, penyandang disabilitas, pejabat yang memiliki kewenangan transaksi, atau apabila distribusi dilakukan secara massal dan terorganisasi.

Model tersebut harus mempertahankan asas legalitas dan menghindari kriminalisasi berlebihan. Satire, parodi, karya seni, penelitian, dan pendidikan tidak seharusnya dipidana sepanjang konteks sintetisnya dapat dikenali, tidak menggunakan data secara melawan hukum, dan tidak dimaksudkan untuk menipu atau merugikan. Persetujuan juga harus bersifat spesifik, sadar, dan dapat ditarik, terutama untuk penggunaan wajah dan suara. Persetujuan untuk satu produksi tidak otomatis menjadi izin untuk melatih model, membuat variasi baru, atau mendistribusikannya tanpa batas.

Rekonstruksi kedua adalah pemulihan korban. Korban harus memiliki jalur cepat untuk meminta preservasi bukti dan keputusan akses sementara tanpa harus menunggu putusan pidana berkekuatan hukum tetap. Setelah penilaian awal yang akuntabel, platform wajib membatasi distribusi konten yang secara prima facie merupakan deepfake seksual, pemerasan, atau penipuan identitas. Mekanisme ini harus disertai hak banding agar tidak berubah menjadi sensor sewenang-wenang. Pemulihan dapat mencakup penghapusan, deindeksasi, klarifikasi, restitusi, kompensasi, bantuan psikologis, dan pemulihan nama baik.

Rekonstruksi ketiga adalah standar pembuktian. Pemerintah dan aparat penegak hukum perlu menyusun protokol nasional forensik media sintetis yang memuat prosedur akuisisi, pembuatan nilai hash, pencatatan rantai penguasaan, analisis metadata, pemeriksaan model dan perangkat, serta penggunaan lebih dari satu metode deteksi. Kesimpulan ahli harus menjelaskan tingkat keyakinan dan keterbatasan metode. Platform juga harus menyimpan informasi asal-usul konten, catatan modifikasi, dan identitas akun sesuai prinsip proporsionalitas dan perlindungan data pribadi.

Rekonstruksi keempat adalah tanggung jawab platform dan ekosistem teknologi. Penyelenggara sistem elektronik yang menyediakan alat generatif berisiko tinggi perlu menerapkan langkah *safety by design*, antara lain pencegahan penggunaan identitas tanpa hak, tanda asal-usul konten, watermark yang tahan manipulasi, kanal pelaporan prioritas, dan respons cepat terhadap korban. Kewajiban tidak boleh disamakan untuk semua layanan; tingkat tanggung jawab harus mempertimbangkan ukuran, fungsi, kemampuan teknis, dan risiko sistem. Platform tidak serta-merta bertanggung jawab pidana atas seluruh unggahan pengguna, tetapi dapat dimintai pertanggungjawaban administratif apabila mengabaikan risiko yang diketahui atau tidak menjalankan perintah yang sah.

Rekonstruksi kelima adalah pencegahan berbasis tabayyun. Literasi digital perlu bergeser dari sekadar imbauan “jangan percaya hoaks” menuju keterampilan memeriksa sumber, konteks, jejak publikasi, dan indikator media sintetis. Institusi keuangan perlu menerapkan verifikasi berlapis untuk instruksi transaksi berbasis suara atau video. Lembaga pemerintah dan media perlu menyiapkan kanal autentikasi resmi untuk pernyataan pejabat. Tokoh agama dan lembaga pendidikan dapat menerjemahkan tabayyun sebagai etika verifikasi, tanggung jawab penyebaran, dan penghormatan terhadap kehormatan orang lain.

Secara kelembagaan, rekonstruksi ideal tidak selalu harus diwujudkan melalui undang-undang baru yang sepenuhnya terpisah. Pilihan yang lebih realistis ialah membentuk bab khusus mengenai media sintetis dalam revisi UU ITE atau undang-undang kecerdasan buatan, kemudian mengharmonisasikannya dengan KUHP, UU PDP, dan UU TPKS. Norma baru harus menegaskan relasi antara delik pembuatan, delik akibat, dan delik sektoral untuk menghindari pembedaan ganda serta memberikan pedoman mengenai perbarengan tindak pidana.

Tabel 2. Model Integrasi Rekonstruksi Penanggulangan Deepfake

Dimensi	Rekonstruksi Hukum Positif	Nilai Hukum Islam	Keluaran yang Diharapkan
Definisi dan ruang lingkup	Definisi netral teknologi; unsur niat, persetujuan, penyesatan, dan bahaya	Kepastian hukum, maṣlaḥah, sadd al-dharī'ah	Norma adaptif tanpa menghambat penggunaan yang sah
Kriminalisasi	Delik bertingkat untuk pembuatan, penggunaan, distribusi, dan monetisasi	Ta'zīr proporsional; larangan kadzib, tadlīs, dan ghisī	Sanksi sesuai tingkat kesalahan dan kerugian
Pemulihan korban	Takedown cepat, deindeksasi, restitusi, klarifikasi, dan dukungan psikologis	Al-ḍarar yuzāl; ḥifẓ al-nafs dan ḥifẓ al-'ird	Kerugian dihentikan dan martabat dipulihkan
Pembuktian	Protokol forensik, rantai penguasaan, multi-metode deteksi, preservasi data	Tabayyun dan mas'ūliyyah	Bukti andal dan putusan yang dapat dipertanggungjawabkan
Platform dan pencegahan	Safety by design, provenance, watermark, kanal prioritas, literasi digital	Ḥifẓ al-'aql dan pencegahan mafsadah	Ekosistem digital yang lebih aman dan terpercaya

Sumber: Konstruksi penulis.

5. Implikasi Kebijakan dan Agenda Implementasi

Implementasi model rekonstruksi membutuhkan prioritas yang terukur. Dalam jangka pendek, aparat penegak hukum perlu menggunakan konstruksi pasal yang berlaku secara sistematis dan tidak lagi merujuk pada ketentuan yang telah dialihkan ke KUHP. Pedoman bersama antara kepolisian, kejaksaan, kementerian yang membidangi komunikasi digital, dan lembaga perlindungan data diperlukan untuk menentukan jalur penanganan berdasarkan jenis kerugian. Perkara deepfake tidak boleh dipaksakan hanya ke satu undang-undang apabila rangkaian perbuatannya mencakup manipulasi data, penipuan, pemerasan, dan pelanggaran data pribadi.

Dalam jangka menengah, pemerintah perlu membangun pusat kompetensi forensik media sintetis yang menyediakan laboratorium, standar metodologi, pelatihan ahli, dan basis pengetahuan mengenai perkembangan model. Pusat tersebut tidak menggantikan independensi ahli, tetapi memastikan bahwa metode yang digunakan dapat diuji dan direplikasi. Kerja sama internasional diperlukan untuk pelacakan akun, preservasi data lintas yurisdiksi, dan pengembangan standar provenance.

Dalam jangka panjang, Indonesia memerlukan tata kelola kecerdasan buatan yang menempatkan risiko terhadap manusia sebagai titik pusat. Pengaturan deepfake harus terhubung

dengan penilaian risiko sistem AI, kewajiban transparansi, audit, keamanan model, dan perlindungan kelompok rentan. Prinsip maqāṣid al-syarī'ah dapat berfungsi sebagai kerangka evaluatif: apakah inovasi menjaga akal dan martabat, mencegah kerugian, serta memberikan manfaat yang lebih besar daripada risiko yang ditimbulkan.

Keberhasilan kebijakan tidak cukup diukur dari jumlah perkara atau lamanya pidana. Indikator yang lebih relevan meliputi kecepatan penghentian distribusi, tingkat pemulihan korban, keberhasilan restitusi, jumlah konten yang diberi provenance, akurasi prosedur forensik, dan penurunan keberhasilan penipuan berbasis imitasi suara atau video. Orientasi hasil tersebut menjadikan hukum pidana sebagai bagian dari ekosistem perlindungan, bukan satu-satunya alat respons.

KESIMPULAN

Deepfake merupakan bentuk baru kejahatan siber yang tidak hanya memanipulasi data, tetapi juga merekayasa identitas dan persepsi mengenai realitas. Tipologinya mencakup pemalsuan identitas, penipuan finansial, pornografi non-konsensual, pencemaran, pemerasan, disinformasi, dan pemalsuan bukti. Hukum Indonesia dapat menjangkau sebagian perbuatan tersebut melalui Pasal 35 UU ITE, ketentuan penipuan, penghinaan, pornografi dan berita bohong dalam KUHP Nasional, larangan penggunaan data pribadi secara melawan hukum dalam UU PDP, serta kekerasan seksual berbasis elektronik dalam UU TPKS. Namun, setelah berlakunya KUHP Nasional dan UU Nomor 1 Tahun 2026, penerapannya harus dibaca secara sistematis karena sejumlah norma lama telah dialihkan atau disesuaikan. Fragmentasi, kekosongan mengenai identitas sintetis, keterlambatan penghapusan, kesulitan atribusi, dan terbatasnya pemulihan korban menunjukkan bahwa kerangka yang ada belum memadai.

Dalam hukum Islam, deepfake yang digunakan untuk menipu atau merugikan dapat dikategorikan sebagai jarīmah ta'zīr karena mengandung al-kadzib, tadlīs, ghisī, ifk atau buhtān, dan ḍarar. Maqāṣid al-syarī'ah memperlihatkan bahwa dampaknya mengancam perlindungan akal, jiwa, kehormatan, keturunan, dan harta. Oleh karena itu, rekonstruksi yang diperlukan adalah model komplementer berupa definisi netral teknologi, kriminalisasi bertingkat, pengecualian yang jelas bagi penggunaan sah, pemulihan korban secara cepat, standar pembuktian digital, tanggung jawab platform berbasis risiko, dan pendidikan tabayyun. Integrasi ini tidak menggantikan hukum positif dengan sanksi keagamaan, melainkan menjadikan nilai keadilan, kemaslahatan, dan pencegahan kemudaratatan sebagai dasar filosofis bagi hukum siber Indonesia yang responsif dan berorientasi pada martabat manusia.

DAFTAR PUSTAKA

- Al-Qur'an al-Karim.
- al-Syātibī, Abū Ishāq. (1997). *Al-Muwāfaqāt fī Uṣūl al-Sharī'ah* (Jilid 2). Beirut: Dār al-Ma'rifah.
- al-Zuhaylī, Wahbah. (2011). *Fiqh Islam wa Adillatuhu* (Jilid 6). Jakarta: Gema Insani.
- Ariwibowo, T., Ismaidar, Rahman, A. M. S., Nurdana, S. A., & Waruwu, P. R. (2026). Analisis hukum pidana terhadap penyebaran deepfake berbasis artificial intelligence dalam perspektif pembaruan hukum pidana di Indonesia. *YUSTISI*, 13(2). <https://doi.org/10.32832/yustisi.v13i2.23887>.
- Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107, 1753–1820. <https://doi.org/10.15779/Z38RVOD15J>.

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
- Ibn 'Āshūr, Muḥammad al-Ṭāhir. (2006). *Treatise on Maqāṣid al-Sharī'ah*. London: International Institute of Islamic Thought.
- Kamali, M. H. (2008). *Shari'ah Law: An Introduction*. Oxford: Oneworld Publications.
- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. <https://doi.org/10.1016/j.bushor.2019.11.006>.
- Marzuki, P. M. (2016). *Penelitian Hukum: Edisi Revisi*. Jakarta: Kencana.
- Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), Article 7. <https://doi.org/10.1145/3425780>.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52. <https://doi.org/10.22215/timreview/1282>.
- Republik Indonesia. Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana telah diubah terakhir dengan Undang-Undang Nomor 1 Tahun 2024.
- Republik Indonesia. Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi.
- Republik Indonesia. Undang-Undang Nomor 12 Tahun 2022 tentang Tindak Pidana Kekerasan Seksual.
- Republik Indonesia. Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi.
- Republik Indonesia. Undang-Undang Nomor 1 Tahun 2023 tentang Kitab Undang-Undang Hukum Pidana sebagaimana diubah dengan Undang-Undang Nomor 1 Tahun 2026 tentang Penyesuaian Pidana.
- Republik Indonesia. Undang-Undang Nomor 1 Tahun 2026 tentang Penyesuaian Pidana.