

Tanggung Jawab Platform Digital dalam Pencegahan Penyebaran Konten *Deepfake* sebagai Bentuk Kejahatan Siber

Dicky Priambodo

Universitas Terbuka, Indonesia

048419863@ecampus.ut.ac.id

Abstract

The development of artificial intelligence has given birth to deepfake technology which is capable of realistically manipulating audio, video and images so that it has the potential to be misused as a form of cybercrime such as fraud, defamation and disinformation. This research aims to analyze the responsibility of digital platforms in preventing the spread of deepfake content and evaluate their effectiveness from an Indonesian legal perspective. The approach used is a normative juridical approach, with a statutory and conceptual approach, through a study of the Informasi dan Transaksi Elektronik Law, criminal law and personal data protection law. The research results show that digital platform responsibilities include preventive, repressive and corrective aspects. However, implementation is still not optimal due to the lack of specific regulations against deepfakes, limited detection technology, and lack of coordination between platforms and governments. In addition, the challenges posed by rapid technological developments and large amounts of content also affect the effectiveness of control. Therefore, it is necessary to strengthen supervision, increase technical capabilities, and collaborate between the government, digital platforms and society to create a safe and responsible digital space.

Keywords: *Cybercrime, Digital Platforms, Deepfake, Regulation, Responsibility.*

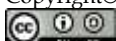
Abstrak

Perkembangan kecerdasan buatan telah melahirkan teknologi *deepfake* yang mampu memanipulasi audio, video, dan gambar secara realistis sehingga berpotensi disalahgunakan sebagai bentuk kejahatan dunia maya seperti penipuan, pencemaran nama baik, dan disinformasi. Penelitian ini bertujuan untuk menganalisis tanggung jawab platform digital dalam mencegah penyebaran konten *deepfake* dan mengevaluasi efektivitasnya dari perspektif hukum Indonesia. Pendekatan yang digunakan adalah pendekatan yuridis normatif, dengan pendekatan perundang-undangan dan konseptual, melalui kajian terhadap Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), hukum pidana, dan hukum perlindungan data pribadi. Hasil penelitian menunjukkan bahwa tanggung jawab platform digital mencakup aspek preventif, represif, dan korektif. Namun, penerapannya masih belum optimal karena kurangnya peraturan khusus terhadap *deepfake*, terbatasnya teknologi deteksi, dan kurangnya koordinasi antara platform dan pemerintah. Selain itu, tantangan yang ditimbulkan oleh perkembangan teknologi yang pesat dan konten dalam jumlah besar juga mempengaruhi efektivitas pengendalian. Oleh karena itu, perlu dilakukan penguatan pengawasan, peningkatan kemampuan teknis, dan kerja sama antara pemerintah, platform digital, dan masyarakat untuk menciptakan ruang digital yang aman dan bertanggung jawab.

Keywords: *Deepfake, Platform Digital, Kejahatan Siber, Regulasi, Tanggung Jawab.*

Pendahuluan

Perkembangan pesat teknologi informasi dan komunikasi telah membawa perubahan signifikan pada semua aspek kehidupan manusia. Kemajuan dalam kecerdasan buatan,



khususnya teknologi *deepfake*, telah memungkinkan pembuatan konten digital seperti video, audio, dan gambar yang tampak nyata tetapi sebenarnya telah diubah (Respati *et al.*, 2024). Teknologi ini awalnya dikembangkan untuk bidang kreatif dan inovatif seperti hiburan dan pendidikan, namun dalam aplikasi praktis, teknologi ini juga membawa risiko penyalahgunaan.

Menurut Haida dan Nuriyatman (2024) menyebutkan bahwa penyalahgunaan teknologi *deepfake* telah menyebabkan berbagai masalah, seperti penyebaran informasi yang salah, pencemaran nama baik, dan bahkan penipuan serta eksploitasi individu. Dengan latar belakang ini, seiring dengan perkembangan teknologi digital yang pesat, teknologi *deepfake* dapat dikategorikan sebagai bentuk kejahatan dalam ranah kejahatan siber yang semakin kompleks.

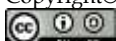
Kejahatan siber merujuk pada segala aktivitas ilegal yang menyalahgunakan jaringan komputer, sistem informasi, dan teknologi digital untuk menyebabkan kerugian bagi individu, kelompok, organisasi, atau negara (Krisyenda *et al.*, 2025). Kejahatan ini mencakup berbagai aktivitas ilegal yang memanfaatkan kemajuan teknologi informasi dan komunikasi, dengan tujuan keuntungan pribadi, seperti mencari keuntungan, mengganggu sistem, mencuri data, atau merusak infrastruktur digital.

Dalam praktiknya, kejahatan siber dapat bersifat tradisional, seperti penipuan digital dan pencurian identitas, atau lebih canggih, seperti menggunakan peretasan, *malware*, *phishing*, dan teknik rekayasa sosial untuk merugikan korban. Lebih jauh lagi, kejahatan siber mencakup tindakan yang bertujuan untuk merusak reputasi pribadi, menyebarkan informasi yang berbahaya, dan bahkan mengancam keamanan nasional melalui serangan siber terorganisir.

Pada tahun 2025, Indonesia menyaksikan kasus penyalahgunaan teknologi *deepfake*. Polisi membongkar jaringan penipuan yang menggunakan kecerdasan buatan (*artificial intelligence*) untuk mengubah video pejabat publik. Dalam kasus ini, para pelaku kejahatan membuat video *deepfake* yang membuat para pemimpin daerah tampak memberikan instruksi atau pesan kepada publik, dan menggunakan video-video ini untuk menipu para korban. Dilaporkan bahwa para pelaku kejahatan memperoleh keuntungan puluhan juta rupiah dari penipuan ini. Menurut Meilina (2025) bahwa awal tahun 2025, polisi Indonesia juga membongkar kasus penipuan digital yang melibatkan video *deepfake*. Dalam kasus ini, presiden Indonesia dipalsukan untuk mengumumkan bantuan keuangan kepada masyarakat. Video-video ini digunakan untuk menipu para korban agar mentransfer uang, sehingga menyebabkan kerugian finansial bagi mereka kurang lebih enam puluh juta rupiah.

Fenomena ini menunjukkan urgensi pengaturan hukum yang lebih spesifik terkait penyalahgunaan teknologi *deepfake* di Indonesia, karena kerangka hukum yang ada belum mampu mengaturnya secara komprehensif sehingga penegakan hukum masih menghadapi kendala. Menurut Wahyudi (2025) diperlukan pembaruan undang-undang atau pembentukan regulasi baru yang khusus mengatur *deepfake* dan implikasinya. Di Indonesia, regulasi terkait tanggung jawab *platform digital* masih terus berkembang, khususnya dalam menangani fenomena baru seperti *deepfake*. Regulasi yang ada, seperti Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), tidak secara eksplisit membahas *deepfake*. Hal ini menghadirkan tantangan dalam mendefinisikan batasan tanggung jawab platform digital, khususnya dalam menyeimbangkan kebebasan berekspresi dengan pencegahan penyalahgunaan teknologi.

Platform digital seperti media sosial, layanan *streaming* video, dan aplikasi pesan instan memainkan peran penting dalam penyebaran informasi. Sebagai kunci distribusi



konten, platform digital seringkali menjadi saluran utama untuk penyebaran konten orisinal yang mendalam secara luas dan cepat. Hal ini menimbulkan pertanyaan tentang tanggung jawab platform digital dalam mencegah dan mengatasi penyebaran konten yang berpotensi bermasalah tersebut.

Oleh karena itu, diperlukan studi yang lebih mendalam tentang tanggung jawab platform digital dalam mencegah penyebaran konten *deepfake* (suatu bentuk kejahatan siber). Berdasarkan hal tersebut, penelitian ini membahas pertanyaan-pertanyaan terkait apa saja karakteristik konten *deepfake* sebagai kejahatan siber? Apa peran dan tanggung jawab platform digital dalam mencegah penyebaran konten *deepfake*? Dan apa saja hukum Indonesia yang relevan mengenai tanggung jawab platform digital atas penyebaran konten tersebut?

Penelitian ini bertujuan untuk menganalisis karakteristik konten *deepfake* sebagai kejahatan siber, di mana membahas terkait peran dan tanggung jawab platform digital dalam mencegah penyebaran konten *deepfake*, dan menganalisis hukum Indonesia yang relevan mengenai tanggung jawab platform digital atas penyebaran konten tersebut.

Metode Penelitian

Penelitian ini menggunakan pendekatan hukum normatif, dengan menggunakan tinjauan pustaka sebagai titik awal untuk mengeksplorasi norma dan prinsip hukum. Materi penelitian meliputi dokumen hukum primer seperti undang-undang dan peraturan, dokumen hukum sekunder seperti buku, jurnal, dan literatur terkait, serta dokumen hukum tersier seperti kamus dan ensiklopedia. Pengumpulan data dilakukan melalui tinjauan pustaka, pencarian basis data seperti *Google Scholar* dan *Scopus*, serta berbagai sumber akademik di jurnal nasional yang berotoritas. Dari perspektif hukum siber, penelitian ini menganalisis fenomena kejahatan *deepfake* dan upaya pencegahannya di Indonesia. Berdasarkan hal tersebut, penelitian ini mengeksplorasi karakteristik konten *deepfake* sebagai kejahatan siber, peran dan tanggung jawab platform digital dalam mencegah penyebarannya, dan hukum Indonesia yang terkait dengan tanggung jawab tersebut, dengan tujuan untuk menganalisis secara komprehensif dampak hukum, teknologi, dan sosial dari penyebaran konten *deepfake*.

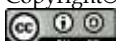
Hasil dan Pembahasan

Peran dan tanggung jawab platform digital dalam ekosistem informasi

Perkembangan teknologi digital menjadikan platform digital sebagai andalan penyebaran informasi. Platform seperti media sosial, layanan berbagi video, dan forum online tidak hanya berfungsi sebagai sarana komunikasi namun juga sebagai penyebar utama konten yang dikonsumsi oleh masyarakat luas. Dalam konteks ini, platform digital memikul tanggung jawab besar atas konten yang beredar di sistem mereka, termasuk konten berbasis teknologi *deepfake*.

Deepfake merupakan hasil manipulasi media berbasis kecerdasan buatan, yang menciptakan konten visual dan audio yang tampak nyata namun sebenarnya palsu (Prayoga dan Tuaksikal, 2025). Karakteristik ini menjadikan *deepfake* sangat berbahaya jika disebarkan melalui platform digital, karena dapat dengan mudah memengaruhi persepsi publik.

Tanggung jawab platform digital dalam mencegah penyebaran konten *deepfake* dapat dipahami melalui beberapa aspek utama yang saling terkait dan menunjukkan bahwa platform tidak lagi menjadi media pasif atau sekadar perantara. Pertama, platform digital berfungsi sebagai perantara (*intermediaries*) distribusi konten, berperan sebagai jembatan antara pembuat konten dan pengguna sebagai konsumen informasi (Putranto dan Andari,



2025). Dalam posisi ini, platform berperan strategis dalam menentukan bagaimana suatu konten diakses, didistribusikan, bahkan diviralkan di ruang publik digital. Oleh karena itu, tanggung jawab yang melekat tidak hanya bersifat teknis tetapi juga moral dan hukum. Menurut Librianti (2025) menyebutkan bahwa platform tidak bisa sepenuhnya lepas dari konsekuensi penyebaran konten, apalagi jika konten tersebut mengandung unsur manipulatif seperti deepfake yang dapat merugikan individu dan masyarakat secara keseluruhan. Kewajiban ini menjadi semakin penting mengingat algoritma distribusi konten yang digunakan oleh platform seringkali mempercepat penyebaran informasi tanpa verifikasi yang memadai.

Kedua, sebagai pengelola sistem (tata kelola konten), platform digital mempunyai kendali penuh terhadap mekanisme internal pengaturan peredaran konten, termasuk penggunaan algoritma, sistem moderasi, dan kebijakan komunitas (Jaya dan Khalid, 2025). Hal ini menunjukkan bahwa platform sebenarnya memiliki kekuatan besar untuk memungkinkan deteksi dini konten deepfake melalui teknologi kecerdasan buatan dan mekanisme pelaporan pengguna. Selain itu, kebijakan komunitas yang dirancang secara komprehensif dapat menjadi alat penting untuk menetapkan batasan yang jelas seputar jenis konten yang diperbolehkan dan dilarang. Dengan cara ini, platform tidak hanya dapat bereaksi terhadap pelanggaran namun juga secara proaktif menciptakan lingkungan digital yang lebih aman. Namun, efektivitas peran ini sangat bergantung pada konsistensi penerapan kebijakan dan transparansi proses moderasi konten.

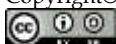
Ketiga, platform digital sebagai entitas yang bertanggung jawab secara sosial tidak bisa hanya berorientasi pada kepentingan ekonomi tanpa mempertimbangkan dampak sosial dari aktivitasnya (Ilham et al., 2025). Di era digital yang sangat terhubung, platform mempunyai dampak yang signifikan terhadap opini publik, stabilitas sosial, dan bahkan keamanan nasional. Oleh karena itu, tanggung jawab sosial platform mencakup upaya untuk melindungi pengguna dari segala bentuk kejahatan dunia maya, termasuk penyalahgunaan teknologi deepfake yang dapat digunakan untuk menipu, mencemarkan nama baik, atau memanipulasi informasi. Bentuk akuntabilitas ini dapat dicapai dengan meningkatkan literasi digital, bekerja sama dengan pemerintah dan lembaga penegak hukum, serta menyediakan mekanisme pengaduan yang cepat dan efektif kepada para korban.

Dari ketiga aspek tersebut, posisi platform digital dalam ekosistem informasi modern sudah tidak netral lagi. Platform telah berkembang menjadi aktor aktif yang memiliki kekuasaan dan tanggung jawab dalam mengelola arus informasi. Oleh karena itu, dalam menghadapi ancaman konten palsu yang mendalam, platform digital perlu memainkan peran yang lebih sensitif, bertanggung jawab, dan berorientasi untuk melindungi pengguna dan kepentingan masyarakat luas.

Bentuk tanggung jawab platform digital dalam pencegahan konten *deepfake*

Dalam praktiknya, tanggung jawab platform digital dalam mencegah penyebaran konten deepfake tidak bersifat tunggal, melainkan mencakup serangkaian upaya terstruktur dan berkelanjutan yang dapat dibagi menjadi tiga bentuk utama, yaitu preventif, represif, dan korektif (Wafi et al., 2025). Ketiga bentuk pertanggungjawaban tersebut menunjukkan bahwa platform digital tidak hanya perlu mengambil tindakan setelah terjadi pelanggaran, namun juga harus mampu mengantisipasi dan meminimalkan dampak penyebaran konten manipulatif berbasis teknologi.

Menurut Lestari et al. (2026) menyebutkan bahwa tanggung jawab preventif adalah langkah pertama yang penting dalam mencegah penyebaran deepfake sebelum konten



menjangkau khalayak luas. Dalam konteks ini, platform digital perlu mengoptimalkan penggunaan teknologi, khususnya sistem berbasis kecerdasan buatan (artificial intelligence) yang mampu mendeteksi pola manipulasi audio, video, dan gambar secara otomatis (Salim, 2025). Selain itu, kebijakan masyarakat (Pedoman Komunitas) yang jelas dan tegas juga merupakan alat penting dalam upaya pencegahan (Jufri dan Putra, 2021). Platform perlu menetapkan standar yang secara eksplisit melarang konten manipulatif dan berbahaya, seperti pencemaran nama baik, penipuan, dan pornografi non-konsensual berdasarkan deepfake. Tidak kalah penting, platform juga mempunyai tanggung jawab untuk meningkatkan literasi digital pengguna melalui edukasi berkelanjutan, sehingga masyarakat bisa lebih pilih-pilih dalam menerima dan menyebarkan informasi, serta tidak mudah tersinggung dengan konten yang belum terverifikasi (Fauzi et al., 2023).

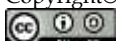
Selain itu, ketika konten deepfake ditemukan dan disebar di platform, akan ada tanggung jawab atas hal tersebut. Tanggung jawab ini di sebut dengan tanggung jawab represif. Dalam konteks ini, platform mempunyai kewajiban untuk mengambil tindakan tegas dalam mencegah penyebaran epidemi dan meminimalkan dampak negatifnya. Menurut Margono (2026) menyebutkan bahwa salah satu bentuk tindakan keras yang utama adalah penghapusan konten yang terbukti melanggar hukum atau kebijakan internal platform. Selain itu, platform ini dapat menjatuhkan sanksi terhadap akun-akun yang terlibat dalam penyebaran deepfake, mulai dari pembatasan aktivitas hingga penutupan akun secara permanen. Dalam kasus-kasus tertentu yang berdampak besar, seperti penipuan atau pencemaran nama baik yang meluas, platform juga harus bekerja sama dengan aparat penegak hukum melalui mekanisme pelaporan yang transparan dan akuntabel. Hal ini menunjukkan bahwa platform ini tidak hanya bertindak sebagai pengelola sistem tetapi juga sebagai mitra penegakan hukum di ruang digital.

Adapun tanggung jawab korektif merupakan tindakan pemulihan yang diambil setelah konten deepfake menyebar (Abdurrachman, 2026). Tanggung jawab ini penting mengingat dampak deepfake tidak hanya bersifat teknis, tetapi juga psikologis, sosial, dan reputasi bagi para korbannya. Dalam hal ini, platform dapat memberikan klarifikasi atau pelabelan khusus terhadap konten yang terindikasi manipulatif guna memberikan informasi yang benar kepada masyarakat (Rizal et al., 2025). Selain itu, platform perlu mendukung pemulihan reputasi korban, seperti dengan memprioritaskan laporan atau membantu memberikan klarifikasi. Yang terpenting adalah menyediakan mekanisme pengaduan yang cepat, mudah diakses, dan responsif merupakan bagian dari tanggung jawab perbaikan yang harus dipenuhi. Dengan adanya sistem pengaduan yang efektif, para korban dapat segera melaporkan konten berbahaya dan menerima perlindungan yang diperlukan.

Secara keseluruhan, ketiga bentuk pertanggungjawaban ini menunjukkan bahwa peran platform digital dalam mencegah konten deepfake bersifat komprehensif dan berkelanjutan. Platform tidak hanya harus mencegah dan mengambil tindakan, namun juga bertanggung jawab untuk memulihkan dampak dan menciptakan ekosistem digital yang lebih aman, adil, dan dapat dipercaya.

Keterkaitan regulasi hukum dengan tanggung jawab platform digital

Di Indonesia, tanggung jawab platform digital dalam mencegah penyebaran konten deepfake tidak dapat dilepaskan dari kerangka hukum yang berlaku, khususnya melalui Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), Kitab Undang-Undang Hukum Pidana (KUHP), serta Undang-Undang Perlindungan Data Pribadi (UU PDP) (Arvitto, 2025). Ketiga instrumen hukum tersebut pada dasarnya telah memberikan landasan

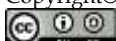


normatif untuk menindak berbagai bentuk pelanggaran di ruang digital, seperti pencemaran nama baik, penyebaran konten bermuatan asusila, penipuan, hingga pelanggaran privasi. Meskipun demikian, regulasi yang ada saat ini belum secara eksplisit mengatur mengenai fenomena deepfake sebagai bentuk kejahatan siber yang berbasis teknologi kecerdasan buatan. Kondisi ini menimbulkan tantangan tersendiri dalam praktik penegakan hukum, khususnya dalam menentukan batasan tanggung jawab platform digital. Ketidadaan definisi hukum yang spesifik mengenai deepfake menyebabkan terjadinya kekaburan norma, sehingga aparat penegak hukum maupun pihak platform harus melakukan penafsiran terhadap ketentuan yang ada. Hal ini berpotensi menimbulkan inkonsistensi dalam penerapan hukum serta ketidakpastian bagi para pihak yang terlibat.

Menurut Oktavia dan Prabowo (2025) menyebutkan bahwa pada tahun 2025 terdapat fenomena dari beberapa kasus nyata yang terjadi di Indonesia. Salah satu kasus yang cukup menonjol adalah penipuan berbasis deepfake yang memanipulasi video tokoh publik, seperti Presiden dan pejabat negara, seolah-olah memberikan bantuan kepada masyarakat. Dalam kasus tersebut, pelaku menggunakan teknologi deepfake untuk membuat video yang tampak autentik, kemudian menyebarkannya melalui media sosial disertai nomor kontak guna menipu korban. Aparat penegak hukum melalui Direktorat Tindak Pidana Siber berhasil menangkap pelaku yang memproduksi dan menyebarkan konten tersebut.

Dari perspektif hukum, pelaku dalam kasus ini dapat dijerat dengan beberapa ketentuan, antara lain Pasal 28 ayat (1) Undang-Undang Informasi dan Transaksi Elektronik terkait penyebaran berita bohong dan menyesatkan yang merugikan konsumen. Selain itu, Pelaku tersebut dikenakan Pasal 378 Kitab Undang-Undang Hukum Pidana. Ketentuan dalam Pasal 51 ayat (1) Jo. Pasal 35 Undang-Undang Informasi dan Transaksi Elektronik mengatur mengenai perbuatan manipulasi informasi elektronik yang dilakukan secara sengaja dan berpotensi merugikan pihak lain atau masyarakat. Dalam hal ini, pelaku dinilai telah dengan sengaja mengubah atau merekayasa informasi elektronik untuk tujuan penipuan. Di sisi lain, Pasal 378 Kitab Undang-Undang Hukum Pidana mengatur mengenai tindak pidana penipuan, yaitu perbuatan untuk memperoleh keuntungan secara melawan hukum, baik untuk diri sendiri maupun orang lain, melalui cara-cara yang bersifat menyesatkan atau tipu daya. Berdasarkan kedua ketentuan tersebut, pelaku dapat dikenai ancaman pidana penjara mulai dari empat hingga dua belas tahun serta denda dengan jumlah maksimal mencapai dua belas miliar rupiah (Kristiyenda et al., 2025). Penerapan pasal-pasal ini mencerminkan keseriusan pemerintah dalam menegakkan hukum terhadap kejahatan siber, khususnya yang memanfaatkan teknologi canggih seperti deepfake.

Tidak hanya itu, terdapat kasus deepfake jenis pornografi juga mulai marak di Indonesia, bahkan melibatkan tokoh publik dan figur terkenal. Pemerintah melalui Kementerian Komunikasi dan Digital menyatakan bahwa penyalahgunaan deepfake dalam bentuk pornografi semakin marak, meskipun belum ada regulasi khusus yang mengaturnya secara spesifik. Menurut Natasya dan Andriasari (2023) menyebutkan dalam praktiknya, tindak pidana yang berkaitan dengan pornografi pada dasarnya telah diatur sejak lama dalam Undang-Undang Republik Indonesia Nomor 1 Tahun 1946 tentang Peraturan Hukum Pidana, khususnya dalam Buku II Bab XIV yang mengatur kejahatan terhadap kesusilaan. Pengaturan tersebut kemudian dipertegas dan diperluas melalui Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi (selanjutnya disebut UU Pornografi). Namun, seiring dengan perkembangan teknologi dan perubahan modus penyebaran yang kini banyak memanfaatkan media sosial dan platform digital, ketentuan mengenai pornografi juga



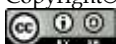
diakomodasi dalam Undang-Undang Informasi dan Transaksi Elektronik (UU ITE) sebagai upaya untuk menjangkau bentuk-bentuk pelanggaran yang terjadi di ruang siber.

Dalam konteks keamanan siber, Rancangan Undang-Undang Keamanan dan Ketahanan Siber (RUU KKS) menjadi penting sebagai upaya memperkuat sistem pertahanan nasional terhadap ancaman teknologi seperti deepfake. Menurut Fadillah et al., (2025), RUU ini bertujuan meningkatkan kemampuan negara dalam menghadapi serangan siber dan penyebaran informasi palsu yang berpotensi mengganggu stabilitas sosial, politik, dan ekonomi. Mengingat deepfake semakin sering digunakan dalam disinformasi politik, terutama menjelang momen penting seperti pemilu, maka urgensi pengaturannya semakin tinggi. Oleh karena itu, RUU KKS perlu memuat ketentuan khusus terkait pencegahan dan penindakan penyalahgunaan deepfake, termasuk melalui kerja sama dengan platform digital dalam mendeteksi dan menghapus konten yang merugikan masyarakat.

Regulasi ini perlu menetapkan sanksi tegas dan proporsional bagi penyalahgunaan teknologi deepfake untuk tujuan kriminal, seperti penipuan identitas dan eksploitasi seksual digital, termasuk hukuman penjara, denda besar, serta kewajiban memulihkan reputasi korban. Selain itu, platform digital juga harus bertanggung jawab dengan menyediakan sistem deteksi dan moderasi konten deepfake. Jika terjadi kelalaian dapat dikenai sanksi administratif seperti denda hingga pencabutan izin. Di sisi lain, regulasi harus melindungi korban melalui bantuan hukum, dukungan psikologis, dan hak atas ganti rugi. Sinergi antara RUU KKS dan UU Artificial Intelligence diperlukan, di mana RUU KKS berfokus pada penegakan hukum siber, sementara UU Artificial Intelligence mengatur etika, transparansi, dan tanggung jawab pengembang. Menurut Setiarma (2023) yang mengacu pada praktik global, seperti Uni Eropa dan Amerika Serikat, Indonesia perlu segera membentuk regulasi artificial intelligence yang komprehensif. Tanpa itu, Indonesia berisiko tertinggal dalam perlindungan hukum. Dengan regulasi yang tepat, artificial intelligence dapat dimanfaatkan secara aman sekaligus meminimalkan dampak negatifnya bagi masyarakat.

Dalam hal ini terdapat ambiguitas mengenai tanggung jawab platform, apakah mereka hanya bertindak sebagai perantara atau juga bertanggung jawab atas sirkulasi konten, terutama mengingat platform memiliki kendali atas algoritma distribusi konten. Selain itu, tantangan yurisdiksi merupakan hambatan yang serius karena banyak platform digital beroperasi lintas negara, sehingga proses penegakan hukum sering kali terhambat oleh perbedaan sistem hukum. Keterbatasan sistem moderasi juga menjadi kendala karena tidak semua konten deepfake dapat terdeteksi secara otomatis, terutama konten yang memiliki tingkat keaslian tinggi.

Di sisi lain, platform digital juga menghadapi berbagai tantangan dalam menjalankan tanggung jawabnya. Perkembangan teknologi deepfake yang sangat cepat membuat metode deteksi menjadi terus tertinggal, sehingga memerlukan inovasi berkelanjutan dalam sistem keamanan digital. Selain itu, volume konten yang sangat besar yang diunggah setiap harinya menjadikan proses moderasi semakin kompleks dan sulit dikendalikan secara menyeluruh. Upaya pembatasan konten juga sering kali menimbulkan konflik dengan prinsip kebebasan berekspresi, sehingga platform harus berhati-hati dalam mengambil kebijakan agar tidak dianggap melakukan sensor berlebihan. Lebih lanjut, belum adanya standar global yang seragam dalam menangani konten deepfake menyebabkan perbedaan kebijakan antar platform, sehingga efektivitas penanganannya menjadi tidak konsisten. Oleh karena itu, diperlukan penguatan regulasi yang lebih spesifik serta kolaborasi antara pemerintah, platform digital, dan masyarakat guna menghadapi ancaman deepfake secara komprehensif sebagai bentuk kejahatan siber modern.



Tantangan platform digital dalam mengatasi penyebaran *deepfake*

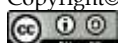
Deepfake sendiri memiliki potensi besar untuk disalahgunakan dalam berbagai bidang, seperti penipuan, pencemaran nama baik, disinformasi politik, eksploitasi seksual digital, hingga ancaman terhadap keamanan siber dan nasional. Oleh karena itu, diperlukan regulasi yang mampu memberikan definisi yang jelas mengenai *deepfake*, termasuk membedakan antara penggunaan yang sah, misalnya dalam kepentingan hiburan, seni, atau pendidikan dalam penyalahgunaan yang merugikan individu maupun masyarakat.

Untuk memenuhi tanggung jawabnya mencegah dan mengatasi penyebaran konten palsu, platform digital menghadapi berbagai tantangan kompleks dari sudut pandang teknis, hukum, dan sosial. Menurut Prayoga dan Tuasikal (2025) menyebutkan salah satu tantangan utamanya adalah pesatnya perkembangan teknologi *deepfake* itu sendiri, di mana mengalami peningkatan kualitas visual dan audio sehingga semakin sulit membedakannya dengan konten aslinya. Akibatnya, sistem deteksi platform sering kali lambat dan memerlukan pembaruan terus-menerus agar dapat mengidentifikasi konten yang dimanipulasi secara efektif. Selain itu, platform digital harus menangani konten dalam jumlah besar, dengan jutaan atau bahkan miliaran konten diunggah setiap hari. Situasi ini membuat proses moderasi konten tidak hanya rumit, tetapi juga memerlukan kombinasi teknologi otomatis dan campur tangan manusia, yang tentunya membutuhkan banyak sumber daya dan banyak waktu (Immanuel, 2023).

Di sisi lain, tantangan juga muncul dari sisi normatif, khususnya terkait kebebasan berekspresi. Upaya platform untuk membatasi atau menghapus konten *deepfake* sering kali menuai kritik dari pihak-pihak yang memandang tindakan tersebut sebagai bentuk sensor atau membatasi hak berpendapat. Hal ini menempatkan platform dalam dilema, di mana mereka harus menyeimbangkan kewajiban mereka untuk melindungi pengguna dari konten berbahaya dengan menghormati hak atas kebebasan berpendapat. Yang tak kalah penting, situasi ini semakin diperumit dengan tidak adanya standar global yang terpadu dalam menangani konten *deepfake*. Setiap platform memiliki kebijakan dan pendekatan yang berbeda, dan peraturan di setiap negara tidak selalu konsisten. Akibatnya penanganan *deepfake* menjadi tidak konsisten dan berpotensi pelaku kejahatan memanfaatkan celah hukum yang ada. Oleh karena itu, kerja sama lintas industri dan transnasional diperlukan untuk mengembangkan standar yang lebih terpadu guna menghadapi ancaman pemalsuan besar-besaran di era digital.

Analisis efektivitas tanggung jawab platform digital

Berdasarkan pembahasan yang telah diuraikan sebelumnya, dapat dianalisis bahwa efektivitas akuntabilitas platform digital dalam mencegah penyebaran konten *deepfake* masih kurang optimal. Hal ini dipengaruhi oleh beberapa faktor utama yang saling berkaitan. Pertama-tama, dari sudut pandang peraturan, kurangnya peraturan yang secara khusus mengatur pemalsuan besar-besaran, mengakibatkan kesenjangan hukum dan tidak jelasnya batasan tanggung jawab platform. Akibatnya, platform digital sering kali hanya mengandalkan kebijakan internal tanpa dukungan dasar hukum yang kuat dan mengikat. Kedua, dari segi teknis, keterbatasan teknologi pendeteksian merupakan kendala yang signifikan. Meskipun platform telah mengembangkan sistem berbasis artificial intelligence untuk mendeteksi konten yang dimanipulasi, perkembangan teknologi *deepfake* yang semakin canggih membuat sistem ini tidak selalu mampu mengidentifikasi konten secara akurat, terutama konten yang memiliki tingkat keaslian tinggi. Ketiga, kurangnya koordinasi



antara platform digital dan pemerintah juga menjadi hambatan dalam membangun mekanisme pemrosesan yang efektif. Kurangnya kerja sama dalam pertukaran data, pelaporan, dan penegakan hukum membuat upaya pencegahan dan penindakan terhadap penyebaran *deepfake* belum terintegrasi secara optimal.

Namun berbagai langkah yang diambil oleh platform digital, seperti penerapan moderasi konten, penguatan kebijakan komunitas, dan pengembangan teknologi deteksi, menunjukkan bahwa platform memiliki potensi besar untuk mengendalikan penyebaran *deepfake*. Efektivitas tanggung jawab ini dapat ditingkatkan secara signifikan, didukung oleh peraturan yang lebih jelas, peningkatan kemampuan teknis, dan sinergi yang kuat antara pemerintah dan platform digital. Oleh karena itu, diperlukan pendekatan kolaboratif yang komprehensif agar upaya pencegahan penyebaran *deepfake* dapat berjalan lebih optimal dan berkelanjutan.

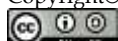
Kesimpulan

Teknologi *deepfake* telah menjadi bentuk kejahatan dunia maya yang semakin canggih dan berpotensi menimbulkan konsekuensi serius bagi individu dan masyarakat secara keseluruhan. Dalam konteks ini, platform digital berperan strategis sebagai pihak yang tidak hanya berperan sebagai perantara namun juga sebagai pengendali distribusi konten di ruang digital. Tanggung jawab platform ini mencakup pencegahan, pemberantasan, dan koreksi, namun implementasinya masih belum memuaskan karena kurangnya peraturan khusus, terbatasnya teknologi deteksi, dan kurangnya koordinasi antara platform dan pemerintah. Namun berbagai upaya seperti moderasi konten dan penerapan kebijakan komunitas menunjukkan bahwa platform digital memiliki potensi besar untuk mengekang penyebaran konten palsu jika didukung oleh sistem hukum dan kebijakan yang lebih jelas.

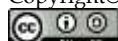
Terkait dengan hal tersebut, disarankan agar pemerintah segera menyusun peraturan khusus yang mengatur *deepfake* secara komprehensif, termasuk definisi, batasan penggunaan, dan sanksi tegas. Platform digital juga perlu meningkatkan kemampuan teknologi deteksi, memperkuat sistem audit, dan menyediakan mekanisme pelaporan yang cepat dan transparan. Selain itu, diperlukan kerja sama yang lebih besar antara pemerintah, platform digital, dan aparat penegak hukum untuk menangani kasus-kasus *deepfake*, termasuk kasus lintas negara. Di sisi lain, masyarakat juga perlu meningkatkan literasi digitalnya agar bisa lebih kritis menyikapi informasi yang beredar. Diharapkan melalui upaya bersama semua pihak, pencegahan dan pengendalian penyebaran "*deepfakes*" dapat lebih efektif dan berkelanjutan.

Daftar Pustaka

- Fadillah, M. F., Nazar, dan Setiawan, H. (2025). Dampak teknologi *deepfake* terhadap kepercayaan publik dan penyebaran informasi di media sosial. *ResearchGate*, e-ISSN 2962-889X; p-ISSN 2962-8903: 149-164.
- Fauzi, A. M., Wahyuni, A. T., Chintia, G., Nenci, I. S., Nurwahidah, N., dan Sari, P. N. (2023). Edukasi pencegahan penipuan online berbasis sosial media di Desa Mekarwangi. *Jurnal Pengabdian Kepada Masyarakat*, 3(2): 60-73.
- Haida, R. S. N., dan Nuriyatman, E. (2024). Urgensi pengaturan perlindungan hukum terhadap korban *deepfake* melalui artificial intelligence (AI) dari perspektif hukum pidana Indonesia. *Jurnal Hukum Republica*, 24(1): 2-3.
- Hukmana, S. Y. (2025). Penipuan dengan AI *deepfake* wajah Prabowo, warga tertipu tawaran bantuan pemerintah. *Media Indonesia*. Diakses 26 April 2026 dari:



- <https://mediaindonesia.com/politik-dan-hukum/741824/penipuan-dengan-ai-deepfake-wajah-prabowo-warga-tertipu-tawaran-bantuan-pemerintah>
- Jaya, Alatas, M., S. M., & Khalid, I. (2025). Konstruksi etika publik dalam tata kelola algoritma media sosial di masyarakat demokratis. *Perspektif: Jurnal Ilmu Sosial dan Humaniora*, 3(2): 86–95.
- Jufri, M. A. A., Putra, A. K. (2021). Aspek hukum internasional dalam pemanfaatan deepfake technology terhadap perlindungan data pribadi. *Uti Possidetis: Journal of International Law*, 1: 31–57.
- Kristiyenda, Y. S., Faradila, J., Basanova, C. (2025). Pencegahan kejahatan deepfake: studi kasus terhadap modus penipuan deepfake Prabowo Subianto dalam tawaran bantuan uang. *Aladalah: Jurnal Politik, Sosial, Hukum dan Humaniora*, 3(2): 149–164. <https://doi.org/10.59246/aladalah.v2i4>
- Librianti, E. O. I. (2025). Deepfake dalam komunikasi politik: tantangan etika dan aspek hukum dalam era artificial intelligence. *Siyasah*, 5(2): 221–238. <https://doi.org/10.32332/siyasah.v4i1>
- Meilina, K. (2025). Penipuan bansos pakai AI wajah Prabowo, ratusan korban rugi puluhan juta. *Katadata.co.id*. Diakses 26 April 2026 dari: <https://katadata.co.id/digital/teknologi/67a5dc910647b/penipuan-bansos-pakai-ai-wajah-prabowo-ratusan-korban-rugi-puluhan-juta>
- Natasya, D. P., Andriasari, D. (2023). Penegakan hukum terhadap tindak pidana penyebaran konten kejahatan pornografi balas dendam (revenge porn) di media sosial ditinjau dari UU ITE dan UU Pornografi. *Bandung Conference Series: Law Studies*. 3(1):115–121. <https://doi.org/10.29313/bcsls.v3i1.4922>
- Octavia, S., Prabowo, D. (2025). Wajah hingga suara Gibran dan Sri Mulyani ikut dicatut untuk tipu korban via video “deepfake”. *Kompas.com*. Diakses 26 April 2026 dari: <https://nasional.kompas.com/read/2025/01/23/21051441/wajah-hingga-suara-gibran-dan-sri-mulyani-ikut-dicatut-untuk-tipu-korban-via>
- Prayoga, H., & Tuasikal, H. (2024). Penyebaran konten deepfake sebagai tindak pidana: analisis kritis terhadap penegakan hukum dan perlindungan publik di Indonesia. *Abdurrauf Law and Sharia*, 1(2): 22–38. <https://doi.org/10.70742/arlash.v2i1.194>
- Putranto, R. B., dan Andari, L. D. (2025). Dari YouTube ke layar lebar: strategi pendapatan dan model bisnis kanal NussaOfficial dalam ekosistem media digital. *Jurnal Studi Komunikasi dan Media*. Tersedia pada: <https://jkd.komdigi.go.id>
- Respati, A. A., Setyarini, A. D., Parlagutan, D., Rafli, M., Mahendra, R. S., & Nugroho, A. A. (2024). Analisis hukum terhadap pencegahan kasus deepfake serta perlindungan hukum terhadap korban. *Media Hukum Indonesia (MHI)*, 2(2): 586. <https://doi.org/10.5281/zenodo.12508126>
- Rifauddin, M., dan Halida, A. N. (2018). Waspada cybercrime dan informasi hoax pada media sosial Facebook. *Khizanah Al-Hikmah: Jurnal Ilmu Perpustakaan, Informasi, dan Kearsipan*, 6(2).
- Rizal, Z. Z., Mujaib, S., dan Febriyanti, P. A. (2025). Tanggung jawab etika dan regulasi platform digital dalam menangani ancaman disinformasi pada media TikTok. *Integrative Perspectives of Social and Science Journal*, 2(7).
- Salim, H. J. (2025). Pelaku penipuan deepfake AI Prabowo ditangkap, begini modus kejahatannya. *Liputan6.com*. Diakses 26 April 2026 dari: <https://www.liputan6.com/cek-fakta/read/5894584/pelaku-penipuan-deepfake-ai-prabowo-ditangkap-begini-modus-kejahatannya>



Wafi, M. S., Wisnubroto, A., dan Prayudi, Y. Tahun. Kejahatan deepfake berbasis artificial intelligence: suatu konsepsi pada penggunaan asas culpabilitas sebagai pembaharuan pertanggungjawaban pidana. *Reformasi Hukum*. 29 (2): 168-183.

