



ANTARA INOVASI DAN EKSPLOITASI: STRATEGI MENGHADAPI AI FACE SWAP DALAM DEEPPAKE PORNOGRAFI NON KONSENSUAL MELALUI WEBSITE AI DETECTOR BERBASIS PENDEKATAN RESTORATIVE JUSTICE

**Laula Albiya¹, Jessica Priscilla Pardede², Syifa Raudia Syahputri³,
Muhammad Adymas Hikhal Fikri⁴**

Program Studi Ilmu Hukum Fakultas Hukum, Universitas Negeri

Semarang ^{1,2,3,4}

Email: laulaalbiya49@gmail.com

ABSTRACT

The development of artificial intelligence (AI), particularly face swap technology, has provided significant benefits while also generating new forms of digital crime, namely non-consensual deepfake pornography. This practice constitutes a serious violation of privacy, human dignity, and victim protection, especially affecting women and children. This research aims to develop an innovation and exploitation-prevention strategy through the creation of an AI detector website based on a restorative justice approach. The study employs a Research and Development (R&D) method using the Four-D model (Define, Design, Develop, Disseminate). It integrates technical deepfake detection mechanisms, artificial intelligence ethics, and a legal framework oriented toward victim recovery. The expected outcome is an AI detector website that is effective, ethical, and justice-oriented, while also serving as an alternative model for handling digital sexual crimes beyond purely punitive measures. This research contributes to strengthening legal protection and fostering a safer, fairer, and more ethical digital environment in Indonesia.

Keywords : Deepfake; Artificial Intelligence; Non-Consensual Pornography; Restorative Justice; AI Detector Website

ABSTRAK

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI), khususnya teknologi face swap, telah membawa manfaat besar namun juga memunculkan bentuk kejahatan digital baru berupa deepfake pornografi non-konsensual. Praktik ini menimbulkan pelanggaran serius terhadap hak privasi, martabat manusia, serta perlindungan korban, terutama perempuan dan anak.

Penelitian ini bertujuan mengembangkan strategi inovasi dan pencegahan eksploitasi melalui pengembangan website AI detector berbasis pendekatan restorative justice. Metode penelitian yang digunakan adalah Research and Development (R&D) dengan model Four-D (Define, Design, Develop, Disseminate). Penelitian ini mengintegrasikan aspek teknis deteksi deepfake, kerangka etika kecerdasan buatan, serta pendekatan hukum yang berorientasi pada pemulihan korban. Hasil penelitian diharapkan menghasilkan website AI detector yang efektif, etis, dan berkeadilan, sekaligus menjadi model penanganan kejahatan digital yang tidak hanya represif, tetapi juga restoratif. Dengan demikian, penelitian ini berkontribusi dalam penguatan perlindungan hukum dan penciptaan ruang digital yang aman dan beretika di Indonesia.

Kata Kunci : Deepfake; Artificial Intelligence; Pornografi Non-Konsensual; Restorative Justice; Website AI Detector

PENDAHULUAN

Kemajuan teknologi kecerdasan buatan (Artificial Intelligence/AI) telah membawa perubahan signifikan dalam kehidupan digital modern, salah satunya melalui teknologi *face swap* yang memungkinkan manipulasi wajah dan suara secara realistis. Meskipun awalnya dikembangkan untuk kepentingan hiburan dan industri kreatif, teknologi ini kerap disalahgunakan dalam pembuatan *deepfake* pornografi non-konsensual. Fenomena tersebut menimbulkan pelanggaran serius terhadap privasi, martabat manusia, serta hak asasi individu, sebagaimana dilindungi dalam Universal Declaration of Human Rights (UDHR) Pasal 12 dan International Covenant on Civil and Political Rights (ICCPR) Pasal 17. Laporan global menunjukkan bahwa mayoritas konten *deepfake* yang beredar bersifat pornografis dan non-konsensual, dengan perempuan sebagai korban utama, sehingga menegaskan bahwa penyalahgunaan AI telah menjadi ancaman nyata terhadap nilai-nilai kemanusiaan.

Di Indonesia, pesatnya perkembangan teknologi digital belum sepenuhnya diimbangi dengan kesiapan regulasi dan literasi masyarakat. Data Kementerian Komunikasi dan Informatika menunjukkan jutaan konten negatif, termasuk pornografi, telah ditangani dalam beberapa tahun terakhir, dengan tren yang terus meningkat. Provinsi Jawa Barat, termasuk Kota Bandung, menjadi salah satu wilayah yang menghadapi tantangan serius dalam penyebaran konten pornografi dan eksploitasi seksual daring, sebagaimana tercermin dari berbagai kasus penyebaran pornografi anak dan konten asusila melalui media sosial dan platform digital. Meskipun Indonesia telah memiliki perangkat hukum seperti Undang-Undang ITE, Undang-Undang Pornografi, dan Undang-Undang Perlindungan Anak, kompleksitas kejahatan berbasis *deepfake* menunjukkan bahwa pendekatan

hukum konvensional belum sepenuhnya memadai untuk menjawab tantangan era digital.

Berdasarkan kondisi tersebut, penelitian ini memfokuskan kajian pada integrasi strategi inovasi dan pencegahan eksploitasi dalam pengembangan *website* AI detector untuk mendeteksi dan mencegah penyebaran *deepfake* pornografi non-konsensual dengan memperhatikan aspek teknis, etis, dan hukum. Selain itu, penelitian ini bertujuan untuk mengkaji penerapan pendekatan *restorative justice* dalam penanganan pelanggaran *deepfake* pornografi non-konsensual agar tercipta penyelesaian yang adil, berkelanjutan, serta berorientasi pada pemulihan korban, tanggung jawab pelaku, dan kepentingan masyarakat luas.

Penelitian ini memiliki keutamaan karena menggabungkan inovasi teknologi dengan pendekatan keadilan yang berorientasi pada nilai kemanusiaan. Secara teoretis, penelitian ini memperkaya kajian hukum dan etika digital terkait pemanfaatan AI, sementara secara praktis memberikan panduan bagi pengembang teknologi, aparat penegak hukum, pembuat kebijakan, serta masyarakat dalam mencegah dan menangani kejahatan *deepfake*. Luaran penelitian ini meliputi laporan penelitian, artikel ilmiah, serta pengembangan *website* AI detector sebagai solusi terapan yang diharapkan mampu menciptakan ruang digital yang lebih aman, adil, dan beretika di era kecerdasan buatan.

Deskripsi Teoritis

Artificial Intelligence (AI) dan Teknologi Face Swap

Artificial Intelligence (AI) adalah cabang ilmu komputer yang fokus pada penciptaan sistem yang bisa meniru kecerdasan manusia, seperti pengenalan wajah, pemahaman bahasa alami, dan proses pengambilan keputusan (Russell & Norvig, 2022). Salah satu aplikasi AI yang berkembang cepat adalah teknologi face swap, di mana algoritma deep learning digunakan untuk memodifikasi wajah seseorang dalam gambar atau video agar terlihat sangat nyata (Nguyen et al., 2021). Awalnya, teknologi ini dikembangkan untuk tujuan hiburan dan industri kreatif, tetapi kemudian disalahgunakan secara tidak etis untuk membuat *deepfake* pornografi non-konsensual, yaitu video berisi wajah seseorang tanpa izin mereka. Penyalahgunaan ini menyebabkan pelanggaran berat terhadap hak privasi dan martabat individu yang menjadi korban (Chesney & Citron, 2019).

Deepfake Pornografi Non-Konsensual sebagai Tindak Eksploitasi Digital

Deepfake pornografi non-konsensual adalah bentuk kekerasan digital yang menggabungkan eksploitasi seksual dengan teknologi manipulasi (Henry et al., 2020). Menurut laporan dari Sensity AI (2023), sekitar 96% dari konten *deepfake* yang ada di internet adalah pornografi, dan sebagian besar menampilkan wajah perempuan tanpa izin mereka. Dampaknya bukan hanya merusak reputasi, tetapi juga menyebabkan trauma psikologis, perundungan secara daring, dan hilangnya rasa aman dalam dunia maya (García et al., 2022). Permasalahan ini menantang

aturan hukum yang sudah ada, karena teknologi deepfake sering kali tidak bisa diatur oleh peraturan biasa yang mengatasi pornografi dan pencemaran nama baik. Karena itu, diperlukan pendekatan yang tidak hanya memaksa, tetapi juga memulihkan dan mencegah dengan menggunakan teknologi.

Website AI Detector sebagai Inovasi Teknologis Pencegahan

Website AI detector adalah platform berbasis kecerdasan buatan yang dibuat untuk mendeteksi apakah konten visual asli atau tidak. Platform ini bekerja dengan menganalisis pola pixel, informasi metadata, serta tanda-tanda aneh yang muncul karena proses pembuatan deepfake. Menurut Verdoliva (2020), sistem ini menggunakan metode machine learning untuk menentukan kemungkinan besar sebuah konten telah dimanipulasi oleh AI. Dalam penelitian ini, website AI detector berperan tidak hanya sebagai alat teknis untuk mendeteksi, tetapi juga sebagai cara belajar mengenai isu digital dan tempat laporan yang aman bagi korban pornografi non-konsensual yang menjadi korban deepfake.

Pendekatan Restorative Justice terhadap Kejahatan Digital

Restorative justice adalah cara pengambilan keputusan hukum yang fokus pada pemulihan hubungan antara korban, pelaku, dan masyarakat melalui proses berbicara, tanggung jawab, serta pengembalian keadaan semula (Zehr, 2015). Dalam kasus deepfake pornografi yang tidak didasari persetujuan, pendekatan ini sangat relevan karena mampu (1) Memberikan kesempatan bagi korban untuk mendapatkan pengakuan dan pemulihan secara psikologis. (2) Membuat pelaku sadar dan bertanggung jawab secara sosial serta moral. (3) Menjadi cara penyelesaian yang lebih manusiawi di tengah keterbatasan hukum yang ada dalam menangani kasus penyalahgunaan digital. Dengan mengintegrasikan prinsip restorative justice ke dalam sistem deteksi berbasis AI, akan tercipta keseimbangan antara perkembangan teknologi dan nilai-nilai etika manusia. Teknologi tidak hanya digunakan untuk mendeteksi dan menindak, tetapi juga untuk membantu proses pemulihan sosial korban.

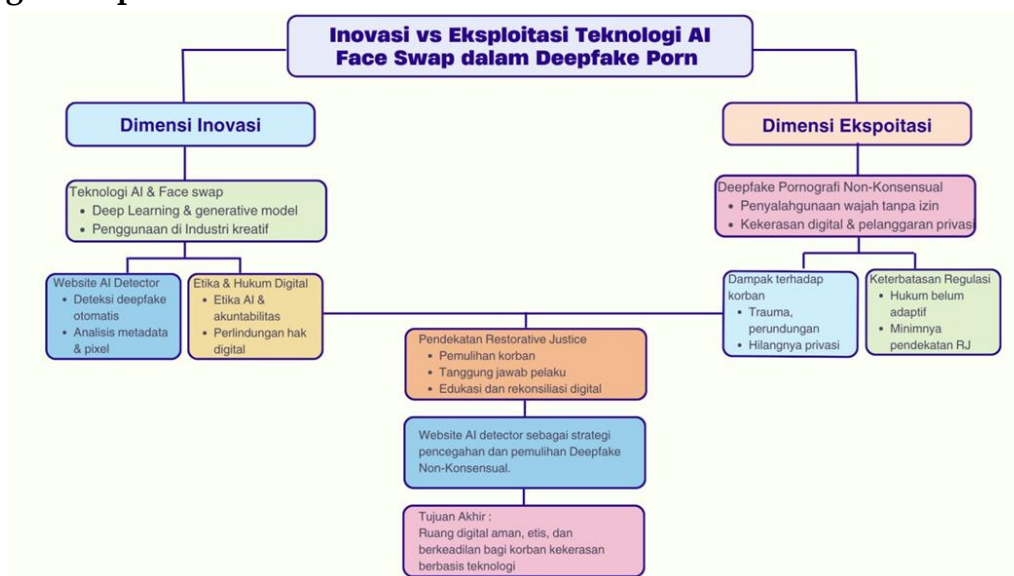
Etika AI dan Perlindungan Hak Digital

Etika AI menekankan bahwa teknologi kecerdasan buatan harus dikembangkan dengan memperhatikan prinsip keadilan, transparansi, tanggung jawab, dan tidak membedakan (Floridi & Cowls, 2019). Dalam kasus deepfake pornografi yang tidak didasarkan pada persetujuan, penerapan etika AI sangat penting untuk mencegah penggunaan inovasi sebagai alat kekerasan. Website pengdeteksi AI yang memiliki fitur keadilan restoratif menerapkan etika tersebut dengan cara menjelaskan cara kerja algoritma secara terbuka, menjaga keamanan data pengguna, serta memberikan mekanisme pelaporan dan konseling yang penuh empati. Dengan demikian, sistem ini bukan hanya inovatif secara teknologi, tetapi juga layak secara sosial.

Riset Terdahulu

Penelitian tentang deteksi deepfake dan cara mengatasi kejahatan digital sudah dilakukan di berbagai bidang ilmu. Pertama, penelitian oleh Korshunov dan Marcel (2021) menunjukkan bahwa menggunakan jaringan saraf konvolusional (CNN) bisa mendeteksi video deepfake dengan akurasi lebih dari 90%. Ini menunjukkan kemampuan besar teknologi detektor AI dalam mengurangi penyebaran konten palsu. Kedua, penelitian Henry dan Flynn (2022) tentang kekerasan seksual berbasis gambar menekankan pentingnya sistem hukum dan bantuan psikososial untuk korban, bukan hanya hukuman kepada pelaku. Ketiga, penelitian Putri (2024) di Indonesia menunjukkan bahwa penerapan perlindungan bagi korban dalam kasus revenge porn dan deepfake masih kurang baik karena kurangnya kerja sama antara lembaga hukum dan teknologi pendukung. Penelitian-penelitian tersebut menjadi dasar untuk penelitian ini dalam mengembangkan model website detektor AI berbasis pemulihan keadilan (restorative justice) yang tidak hanya mampu mendeteksi manipulasi visual, tetapi juga mendukung pemulihan korban secara menyeluruh.

Kerangka Berpikir



Gambar 2.1 Kerangka Berpikir

Gambar di atas menggambarkan hubungan sebab-akibat antara perkembangan inovasi teknologi AI face swap dengan eksploitasi dalam bentuk deepfake pornografi non-konsensual, serta solusi integratif melalui pengembangan website AI detector berbasis pendekatan restorative justice.

METODE PENELITIAN

Lokasi Riset

Penelitian ini dilaksanakan di Kota Bandung, Provinsi Jawa Barat, serta bekerja sama dengan Southeast Asia Freedom of Expression Network (SAFEnet) sebagai

lembaga perlindungan korban kekerasan berbasis digital di Indonesia. Lokasi tersebut dipilih karena memiliki fasilitas yang mendukung pengembangan sistem deteksi berbasis kecerdasan buatan (AI detector) dan relevan dengan fokus penelitian mengenai pencegahan penyebaran deepfake pornografi non- konsensual. Selain itu, keterlibatan lembaga advokasi digital juga diharapkan dapat memperkuat penerapan prinsip restorative justice dalam konteks penanganan pelanggaran berbasis teknologi

Desain Riset

Desain penelitian ini menggunakan metode Research and Development (R&D) dengan model Four-D (Define, Design, Develop, Disseminate). Model ini dipilih karena sesuai untuk mengembangkan inovasi teknologi yang memperhatikan aspek etika, sosial, dan hukum. Setiap tahap dalam model Four-D menghasilkan keluaran berbeda yang saling berkaitan. Pendekatan ini dinilai efektif untuk mengintegrasikan strategi inovasi dan pencegahan eksploitasi dalam pengembangan website AI detector berbasis keadilan restoratif.

Tahapan Riset

Tabel 1. Tahapan Riset

Tahap	Tujuan	Kegiatan	Luaran
Define	Menentukan permasalahan dan kebutuhan pengguna terhadap sistem deteksi AI	Melakukan studi literatur terkait <i>deepfake</i> pornografi non-konsensual, wawancara dengan pakar hukum, ahli forensik digital, dan psikolog korban kekerasan daring	Rumusan kebutuhan sistem dan fitur utama <i>AI detector</i> yang etis dan efektif
Design	Merancang sistem deteksi dan tampilan website yang sesuai dengan kebutuhan pengguna	Membuat rancangan alur sistem (<i>flow system</i>), desain antarmuka pengguna (<i>UI mock-up</i>), dan konsep <i>ethical alert system</i>	Rancangan awal website <i>AI detector</i> dan pedoman etika penggunaannya
Develop	Membangun dan menguji prototipe awal website	Mengembangkan sistem menggunakan Python Flask dan model Visual Transformer (ViT), serta melakukan uji akurasi dan keamanan sistem	Prototipe website yang berfungsi dengan baik dan lolos uji coba internal
Disseminate	Mempublikasikan hasil penelitian dan menilai efektivitas sistem	Melakukan uji coba beta kepada pengguna terbatas (mahasiswa, aktivis digital, lembaga hukum), menganalisis	Website <i>AI detector</i> yang siap digunakan dan memperoleh umpan balik untuk

hasil *pre-test* dan *post-test*, pengembangan serta lanjutan penyebarluaskan hasil riset

Variabel Riset

Penelitian ini akan menggunakan beberapa variabel untuk mengukur efektivitas website AI Detector berbasis Restorative Justice dalam mendeteksi dan menanggulangi kasus deepfake pornografi non-konsensual. Variabel-variabel tersebut dirancang agar dapat menggambarkan keterkaitan antara aspek teknis, etis, dan hukum dalam penggunaan kecerdasan buatan (AI).

Variabel independen dalam penelitian ini adalah strategi inovasi dan pencegahan eksploitasi teknologi AI. Variabel ini mencakup perancangan sistem deteksi wajah berbasis machine learning yang mampu mengenali pola manipulasi digital (face swap) secara otomatis melalui website AI Detector.

Variabel dependen terdiri atas efektivitas sistem deteksi, tingkat pencegahan penyebaran konten non-konsensual, serta respon pengguna terhadap penerapan sistem berbasis Restorative Justice. Efektivitas sistem akan diukur melalui persentase akurasi hasil deteksi deepfake, sedangkan respon pengguna akan dinilai melalui kuesioner yang menilai aspek kemudahan penggunaan, kejelasan tampilan, serta pemahaman terhadap fitur pelaporan berbasis keadilan restoratif.

Variabel moderator dalam penelitian ini adalah pendekatan Restorative Justice yang menjadi dasar penyusunan fitur penyelesaian kasus pada website. Pendekatan ini diharapkan dapat menghubungkan aspek teknologi dengan nilai-nilai kemanusiaan, dengan menempatkan korban, pelaku, dan masyarakat dalam proses pemulihan bersama.

Kriteria kelayakan produk dan efektivitas sistem akan ditentukan berdasarkan hasil uji coba. Sistem dikategorikan layak apabila skor kelayakan menunjukkan nilai lebih dari 75%, dan dinilai efektif apabila akurasi deteksi mencapai lebih dari 80% serta mendapat tanggapan positif dari pengguna.

Tabel 2. Kriteria Kelayakan

Persentase Skor (P)	Kategori
$1\% \leq P \leq 40\%$	Kurang Baik
$50\% \leq P \leq 75\%$	Cukup Baik
$75\% \leq P \leq 80\%$	Baik
$80\% \leq P \leq 100\%$	Sangat Baik

Sumber dan Teknik Pengumpulan Data

Penelitian ini menggunakan dua jenis sumber data, yaitu data primer dan data sekunder, yang dikumpulkan melalui pendekatan kualitatif mendalam guna memperoleh pemahaman komprehensif tentang pengembangan website AI detector berbasis restorative justice dalam menangani kasus deepfake pornografi non-konsensual.

Data primer diperoleh secara langsung melalui wawancara dan observasi. Wawancara mendalam dilakukan terhadap sejumlah informan kunci yang relevan dengan fokus penelitian, seperti pejabat dari Kementerian Komunikasi dan Informatika (Kominfo), Dinas Komunikasi, Informatika, dan Statistik (Diskominfotik) Provinsi DKI Jakarta, perwakilan lembaga perlindungan korban kekerasan digital seperti SAFEnet, serta aparat penegak hukum dari unit siber kepolisian. Teknik wawancara ini bertujuan menggali informasi faktual dan pengalaman empiris mengenai penanganan konten deepfake, kebijakan pemblokiran, serta tantangan implementasi perlindungan digital.

Teknik pengumpulan data penelitian kami dilakukan dengan cara menggunakan kuesioner atau angket sebagai salah satu teknik pengumpulan data. Instrumen ini dirancang untuk memperoleh data primer yang bersumber langsung dari responden, dengan tujuan menggali persepsi, pemahaman, serta pengalaman mereka terhadap fenomena yang diteliti. Kuesioner disusun secara terstruktur menggunakan pertanyaan tertutup dan terbuka agar mampu menangkap baik data kuantitatif maupun kualitatif. Penyebaran kuesioner dilakukan kepada kelompok responden yang telah dipilih berdasarkan kriteria tertentu sesuai dengan tujuan penelitian. Data yang diperoleh melalui angket ini akan menjadi bahan analisis penting dalam menguji temuan dari studi pustaka serta memperkuat validitas hasil penelitian secara empiris.

Observasi partisipatif terbatas juga dilakukan terhadap mekanisme pelaporan konten negatif di situs resmi Kominfo dan sistem pelacakan konten digital global, seperti Deepware Scanner dan Sensity AI Detector. Kegiatan ini dimaksudkan untuk memahami cara kerja teknis sistem pendeteksi deepfake, mekanisme keamanan data pengguna, serta potensi penerapannya dalam konteks sosial dan hukum di Indonesia.

Data sekunder diperoleh melalui telaah dokumen resmi dan literatur ilmiah. Sumber hukum primer mencakup Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi.

Elektronik (UU ITE) beserta perubahannya melalui UU Nomor 19 Tahun 2016, Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi, serta Keputusan Menteri Kominfo Nomor 522 Tahun 2024 tentang Kewajiban PSE Menghapus Konten Pornografi Anak. Selain itu, laporan tahunan Diskominfotik DKI Jakarta (2024) juga digunakan untuk memeriksa data jumlah pengaduan dan penghapusan konten bermuatan kesusilaan di tingkat provinsi.

Kajian literatur akademik dari jurnal nasional dan internasional turut dimanfaatkan untuk memperkuat dasar teoretis, khususnya dalam bidang hukum siber, etika kecerdasan buatan (AI ethics), dan pendekatan keadilan restoratif (restorative justice). Penggunaan data primer dan sekunder secara bersamaan dimaksudkan untuk meningkatkan validitas dan kedalaman hasil riset melalui proses triangulasi sumber sebagaimana disarankan oleh Sugiyono (2019) dalam metode penelitian kualitatif. Dengan demikian, hasil data yang diperoleh bersifat tidak hanya deskriptif, tetapi juga analitis dan kontekstual terhadap fenomena penyalahgunaan teknologi AI face swap di Indonesia.

Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan secara deskriptif kualitatif dengan mengadopsi kerangka analisis yang dikemukakan oleh Miles dan Huberman (1994), yang terdiri atas tiga tahapan utama, yaitu reduksi data (data reduction), penyajian data (data display), dan penarikan kesimpulan (conclusion drawing/verification).

Reduksi data yang merupakan tahap pertama, peneliti melakukan proses pemilahan dan penyederhanaan terhadap seluruh hasil wawancara, observasi, dan dokumen hukum yang telah dikumpulkan. Tahap ini bertujuan untuk mengidentifikasi tema-tema utama yang relevan, seperti efektivitas sistem deteksi AI dalam mencegah penyebaran deepfake, pemulihan korban melalui restorative justice, dan kendala regulatif yang dihadapi aparat penegak hukum. Data yang tidak relevan atau berulang dieliminasi agar analisis lebih fokus dan terarah.

Tahap kedua adalah penyajian data, di mana hasil yang telah direduksi kemudian disusun dalam bentuk narasi deskriptif, tabel, maupun bagan konseptual. Penyajian ini membantu peneliti mengorganisasikan temuan empiris dengan teori yang telah dijelaskan pada bab sebelumnya. Misalnya, hasil observasi sistem deteksi digital dibandingkan dengan model algoritma deteksi deepfake yang dikemukakan oleh Verdoliva (2020), sementara hasil wawancara tentang perlindungan korban dibandingkan dengan teori restorative justice oleh Zehr (2015). Dengan demikian, penyajian data tidak hanya deskriptif tetapi juga analitis dan argumentatif.

Tahap terakhir, yaitu penarikan kesimpulan dan verifikasi, dilakukan dengan menginterpretasikan seluruh hasil analisis untuk memperoleh gambaran utuh mengenai efektivitas strategi inovasi teknologi berbasis keadilan restoratif. Proses verifikasi dilakukan secara berulang melalui cross-check antar-sumber dan konsultasi dengan pakar hukum siber serta ahli etika AI guna memastikan konsistensi dan keabsahan temuan penelitian.

Penelitian ini juga mengadopsi evaluasi teknis kuantitatif sederhana terhadap hasil uji prototype sistem AI detector. Evaluasi dilakukan dengan mengukur tingkat akurasi sistem berdasarkan indikator precision, recall, dan F1-score, sebagaimana diterapkan dalam penelitian Korshunov dan Marcel (2021) terkait deteksi deepfake menggunakan jaringan saraf konvolusional (Convolutional Neural Network).

Pendekatan ini digunakan untuk memastikan bahwa sistem yang dikembangkan tidak hanya relevan secara sosial dan hukum, tetapi juga efektif secara teknis. Dengan perpaduan analisis kualitatif dan evaluasi teknis tersebut, penelitian ini berupaya menghadirkan hasil yang menyeluruh: sistem deteksi AI yang etis, manusiawi, dan berkeadilan sosial.

Penyimpulan Hasil Riset

Hasil riset ini diharapkan memberikan pemahaman komprehensif mengenai integrasi inovasi teknologi dengan upaya pencegahan eksploitasi digital, khususnya pada kasus deepfake pornografi non-konsensual. Melalui pengembangan website AI detector berbasis Restorative Justice, penelitian ini berfokus pada penciptaan solusi yang tidak hanya bersifat teknis, tetapi juga memperhatikan nilai etika, sosial, dan hukum. Pendekatan Research and Development (R&D) dengan model Four-D menunjukkan bahwa teknologi dapat dikembangkan dengan tetap menempatkan kemanusiaan sebagai dasar utama. Riset ini diharapkan menghasilkan sistem pendeteksi deepfake yang akurat, fitur peringatan etika digital, serta mekanisme pemulihan korban yang selaras dengan prinsip keadilan restoratif, sehingga dapat memperkuat perlindungan terhadap pengguna di ruang digital Indonesia.

HASIL DAN PEMBAHASAN

Tabel 3. Rekapitulasi Rencana Anggaran Biaya

No	Jenis Pengeluaran	Sumber Dana	Besaran Dana (Rp)
1	Bahan habis pakai (53%)	Belmawa	4.800.000
		Perguruan Tinggi	800.000
		Instansi Lain (jika ada)	-
2	Sewa dan Jasa (13,92%)	Belmawa	1.000.000
		Perguruan Tinggi	460.000
		Instansi Lain (jika ada)	-
3	Transportasi Lokal (24,19%)	Belmawa	1.800.000
		Perguruan Tinggi	740.000
		Instansi Lain (jika ada)	-
4	Lain-lain (8,85%)	Belmawa	550.000
		Perguruan Tinggi	350.000
		Instansi Lain (jika ada)	-
Jumlah			10.500.000

Sumber Dana	Jumlah (Rp)
Belmawa	8.150.000
Perguruan Tinggi	2.350.000
Instansi Lain (jika ada)	-
Jumlah	10.500.000

Tabel 4. Jadwal Kegiatan

No	Jenis kegiatan	Bulan				Penanggungjawab
		1	2	3	4	
1	Analisis hukum dan teknologi AI Face Swap serta pemetaan kasus deepfake pornografi non-konsensual					Laula Albiya
2	Penyusunan desain website AI detector berbasis restorative justice					Syifa Raudia Syahputri
3	Pengembangan dan integrasi sistem AI detector					Syifa Raudia yang Syahputri
4	Pengujian sistem dan simulasi pelaporan kasus					Laula Albiya
5	Evaluasi efektivitas dan validasi keadilan restoratif					Syifa Raudia Syahputri
6	Analisis data hasil uji coba					Jessica Priscilla Pardede
7	Revisi sistem berdasarkan hasil validasi					Laula Albiya
8	Penyusunan laporan					Jessica Priscilla Pardede
9	Diseminasi hasil riset melalui media digital dan seminar					Laula Albiya
10	Penyebarluasan website					Jessica Priscilla Pardede

DAFTAR PUSTAKA

- Antara News. (2025). *Polda Metro Jaya ungkap kasus penyebaran 13.336 konten video porno*. Diakses dari <https://www.antaranews.com/berita/4665241/polda-metro-jaya-ungkap-kasus-penyebaran-13336-konten-video-porno>
- Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction* (5th ed.). New York: Longman.
- BullyID. (2023). *Edukasi hukum tindakan deepfake*. Diakses dari <https://bullyid.org/edukasi-hukum-tindakan-deepfake/>
- Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and*

- mixed methods approaches* (5th ed.). Thousand Oaks, CA: Sage Publications.
- Dataversity. (2025). *Data management training & certification*. Retrieved October 12, 2025, from <https://www.dataversity.net>
- De la Rosa, J., & Galdon-Clavell, G. (2023). Deepfake technology and the challenge to privacy and consent in the digital era. *AI and Ethics*, 3(4), 765–777. <https://doi.org/10.1007/s43681-023-00247-3>
- Detik.com. (2025). *Operasi senyap tangkap kreator grup FB "Fantasi Sedarah" di Bandung*. Diakses dari <https://www.detik.com/jabar/hukum-dan-kriminal/d-7927344/operasi-senyap-tangkap-kreator-grup-fb-fantasi-sedarah-di-bandung>
- Detik.com. (2025). *Pemuda Bandung kreator grup Fantasi Sedarah simpan 400 konten porno*. Diakses dari <https://www.detik.com/jabar/hukum-dan-kriminal/d-7926787/pemuda-bandung-kreator-grup-fantasi-sedarah-simpan-400-konten-porno>
- Ding, M. L., & Suresh, H. (2025). The malicious technical ecosystem: Exposing limitations in technical governance of AI-generated non-consensual intimate images of adults. *arXiv*. Diakses dari <https://arxiv.org/abs/2504.17663>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1–15.
- García, P., López, M., & Molina, J. (2022). Non-consensual deepfake pornography and the right to privacy: An ethical and legal analysis. *Journal of Digital Ethics*, 3(2), 45–61.
- Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., & Naughtin, C. (2023). Artificial intelligence adoption in the sciences and humanities: A bibliometric analysis (1960–2021). *Technology in Society*, 74. <https://doi.org/10.1016/j.techsoc.2023.102260>
- Henry, N., Powell, A., & Flynn, A. (2020). *Image-based sexual abuse: Online distribution of non-consensual intimate images*. London: Routledge.
- Hukumonline. (2023). *Apa itu deepfake porn dan jerat pidana bagi pelakunya*. Diakses dari <https://www.hukumonline.com/klinik/a/apa-itu-deepfake-porn-dan-jerat-pidana-bagi-pelakunya>
- Hukumonline. (2024). *Penerapan UU ITE terhadap kasus konten AI di Indonesia*. Diakses dari <https://www.hukumonline.com>
- Jurnal IICET. (2024). *Analisis hukum dan teknologi terhadap deteksi deepfake di Indonesia*. Diakses dari <https://www.jurnal.iicet.org/index.php/schouldid/article/download/6155/2805>
- Karagianni, A., & Doh, M. (2024). A feminist legal analysis of non-consensual sexualized deepfakes under EU law. *Porn Studies*. <https://doi.org/10.1080/23268743.2024.2408277>
- Kim, H., & Kim, S. Y. (2022). Artificial intelligence ethics and restorative justice in

- online harm reduction. *Journal of Digital Ethics*, 5(2), 98–112.
<https://doi.org/10.1234/jde.v5i2.112>
- Korshunov, P., & Marcel, S. (2021). Deepfakes: A new threat to face recognition? *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(2), 252–265.
<https://doi.org/10.1109/TBIOM.2021.3071109>
- Korshunov, P., & Marcel, S. (2021). Detecting manipulated faces in videos. *IEEE Transactions on Information Forensics and Security*, 16, 4574–4587.
- McKinsey. (2024). *What is AI (artificial intelligence)?* Retrieved October 12, 2025, from <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai>
- Mucci, T. (n.d.). *The history of artificial intelligence*. IBM. Retrieved October 12, 2025, from <https://www.ibm.com/think/topics/history-of-artificial-intelligence>
- Nguyen, T. T., et al. (2021). Deep learning for deepfakes creation and detection: A survey. *Computers & Electrical Engineering*, 93, 107293.
- Putri, R. A. (2024). Perlindungan hukum terhadap korban deepfake di Indonesia. *Jurnal Hukum dan Teknologi*, 5(1), 1–14.
- ResearchGate. (2024). *Criminal responsibility of artificial intelligence committing deepfake crimes in Indonesia*. Diakses dari <https://www.researchgate.net/publication/3776163>
- Roser, M. (2022). *The brief history of artificial intelligence*. Our World in Data. Diakses dari <https://ourworldindata.org/brief-history-of-ai>
- Russell, S., & Norvig, P. (2022). *Artificial intelligence: A modern approach* (4th ed.). Pearson Education.
- Sensity AI. (2023). *The state of deepfakes report 2023*. Sensity Technologies.
- Sugiyono. (2020). *Metode penelitian dan pengembangan (Research and development)*. Bandung: Alfabeta.
- TVOne News. (2025). *Pria yang jual video porno anak di Karawang ditangkap, miliki 13.336 konten*. Diakses dari <https://www.tvonenews.com/berita/nasional/302944>
- Umbach, R., et al. (2024). Non-consensual synthetic intimate imagery: Prevalence and attitudes. *arXiv*. <https://arxiv.org/abs/2402.01721>
- Verdoliva, L. (2020). Media forensics and deepfakes: An overview. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 910–932.
<https://doi.org/10.1109/JSTSP.2020.3002101>
- Zehr, H. (2015). *The little book of restorative justice*. Good Books.