

Dekonstruksi *Algorithmic Bias* dalam Pidana Ditinjau dari Perspektif Teori Utilitarianisme

Deconstructing Algorithmic Bias in Criminal Sentencing from the Perspective of Utilitarianism Theory

Cheryl Gladisya Kafka¹, Nazwa Fatimah Az-Zahra²

^{1,2}Universitas Pembangunan Nasional “Veteran” Jakarta

Abstract: The integration of Artificial Intelligence (AI) and machine learning technologies into Indonesia's criminal justice system raises fundamental debates regarding justice, ethical legitimacy, and legal certainty. This study aims to analyze the dangers of algorithmic bias to the utilitarian principle of justice and to formulate a reconstruction of judicial discretion in upholding the principle of due process of law. Using a normative-philosophical legal research method through a utilitarian approach, the results show that reliance on judicial algorithms has the potential to transfer historical bias from input data into modern legal decisions. This violates Utilitarianism's qualitative calculus of happiness because it reproduces discriminatory suffering among marginalized groups and reduces human dignity to mere statistical probabilities. The “Black Box AI” phenomenon and automation bias also threaten judicial independence and the presumption of innocence, while triggering a dangerous shift toward a Crime Control Model that prioritizes technical efficiency over legal procedures. Therefore, this study recommends reconstructing judicial discretion based on the principle of “meaningful public control,” which affirms the role of pure AI as an analytical tool (decision-support system) based on Explainable AI (XAI). The center of normative consideration and moral responsibility for sentencing must remain entirely in the hands of human judges to ensure transparency, the dignity of the defendant, and substantive justice in the digital age.

Abstract: Integrasi teknologi Kecerdasan Buatan (*Artificial Intelligence/AI*) dan *machine learning* ke dalam sistem peradilan pidana di Indonesia menghadirkan perdebatan mendasar terkait keadilan, legitimasi etis, dan kepastian hukum. Penelitian ini bertujuan untuk menganalisis bahaya *algorithmic bias* terhadap asas keadilan utilitarian serta merumuskan rekonstruksi diskresi hakim dalam menjaga prinsip *due process of law*. Menggunakan metode penelitian hukum normatif-filosofis melalui pendekatan Utilitarianisme, hasil penelitian menunjukkan bahwa ketergantungan pada algoritma peradilan berpotensi mereproduksi bias historis dari data masukan ke dalam keputusan hukum modern. Hal ini melanggar kalkulasi kebahagiaan kualitatif Utilitarianisme karena mereproduksi penderitaan diskriminatif pada kelompok marginal serta mereduksi martabat manusia menjadi sekadar angka probabilitas statistik. Fenomena *Black Box AI* dan *automation bias* juga mengancam independensi hakim, asas praduga tidak bersalah (*presumption of innocence*), serta memicu pergeseran berbahaya menuju *Crime Control Model* yang mengagungkan efisiensi teknis di atas prosedur hukum. Oleh karena itu, penelitian ini merekomendasikan rekonstruksi diskresi hakim berbasis prinsip “kendali bermakna” (*meaningful public control*), yang menegaskan kedudukan AI murni sebagai alat bantu analitis (*decision-support system*) berbasis *Explainable AI* (XAI). Pusat pertimbangan normatif dan tanggung jawab moral pidana harus tetap berada sepenuhnya di tangan hakim manusia guna menjamin transparansi, martabat terdakwa, dan keadilan substantif di era digital.

Article History

Received: 10 October 2026

Revised: 15 October 2026

Published: 17 October 2026

Keywords: *Algorithmic Bias, Judicial Discretion, Due Process of Law, Epistemic Ethics, Utilitarianism.*

Kata kunci: *Algorithmic Bias, Diskresi Hakim, Due Process of Law, Etika Epistemik, Utilitarianisme.*



<https://doi.org/10.5281/zenodo.23450932>

This is an open-access article under the [CC-BY-SA License](#).



PENDAHULUAN

Integrasi teknologi kecerdasan buatan (*Artificial Intelligence/AI*) dan analitik prediktif dalam ranah peradilan menandai pergeseran paradigma yang radikal dalam sistem penegakan hukum kontemporer. Pengadopsian sistem algoritmik tidak lagi terbatas pada penanganan tugas administratif sekadar efisiensi kantor peradilan, melainkan telah memasuki ke dalam proses pengambilan keputusan

substansif pidana. Penerapan ini terwujud dalam bentuk *predictive policing* untuk pemetaan lokasi kejahatan, sistem penilaian risiko residivisme (*risk assessment*) untuk pertimbangan penahanan dan jaminan, hingga pembentukan algoritma rekomendasi bobot pemidanaan (*sentencing*). Penggunaan instrumen seperti algoritma COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) di Amerika Serikat atau sistem *System Risk Indication* (SyRI) di Belanda mencerminkan ambisi institusi peradilan untuk mengoptimalkan efisiensi operasional, meningkatkan akurasi analisis, serta mengikis ketidakpastian hukum akibat variabilitas putusan hakim manusia.¹ Namun, otomatisasi yang kian masif ini secara langsung menggeser mekanisme penegakan hukum yang semula bertumpu pada diskresi nurani, pertimbangan etis kasuistik, dan kepekaan moral manusia menuju kalkulasi probabilistik berbasis statistik.²

Pergeseran mekanisme ini melahirkan problem normatif dan filosofis yang sangat krusial, yaitu timbulnya fenomena *Black Box AI*. Rekomendasi vonis yang dihasilkan oleh arsitektur pembelajaran mesin (*machine learning*) beroperasi secara *opaque* (tidak transparan), di mana alur penalaran rasional komputasional serta pembobotan variabel di dalam jaringan saraf tiruan tidak dapat dilacak secara eksplisit oleh terdakwa, penasihat hukum, maupun hakim itu sendiri. Kondisi ini berimplikasi destruktif terhadap hak-hak dasar konstitusional, khususnya prinsip *due process of law*, hak atas pembelaan diri yang efektif, serta hak atas penjelasan rasional (*right to explanation*) yang mensyaratkan bahwa setiap keputusan hukum yang mengikat individu harus dapat diuji pertimbangan rasionalnya.³ Situasi ini semakin diperparah karena fakta bahwa algoritma prediktif dilatih menggunakan data historis peradilan (*training data*) yang merekam ketimpangan sosial, diskriminasi struktural, serta bias rasial masa lalu.⁴ Ketika algoritma mengolah data tersebut, variabel proksi, seperti tingkat kejahatan di lingkungan tempat tinggal, riwayat penangkapan kepolisian, atau kondisi sosiologi ekonomi diidentifikasi sebagai prediktor risiko. Hal tersebut mengakibatkan AI secara sistematis mereplikasi dan mengotomatisasi bias historis berbasis suku, ras, gender, dan kelas sosial di balik stempel "objektivitas dan netralitas matematis".

Berdasarkan pandangan filsafat hukum, ketergantungan pada kalkulasi statistik merupakan pemicu konfrontasi epistemologis yang mendasar. Penggunaan algoritma secara inheren memicu pergeseran dari paradigma *individualized justice* menuju *actuarial justice*. *Individualized justice* memandang keadilan material sebagai pencapaian hukum yang wajib memperhitungkan konteks spesifik, kedalaman motivasi subjektif (*subjective culpability*), latar belakang personal, serta martabat unik dari setiap individu sebagai subjek hukum yang rasional.⁵ Sebaliknya, *actuarial justice* mengabaikan subjektivitas individu dan mereduksi manusia menjadi sekadar titik data yang diklasifikasikan ke dalam kategori-kategori risiko kelompok statistik.⁶ Penalaran hukum yang terotomatisasi melahirkan risiko dehumanisasi hukum, di mana pertimbangan kasuistik yang membutuhkan empati, kebijaksanaan, dan pemahaman nilai-nilai kemanusiaan digantikan oleh logika mesin yang mekanistik dan deterministik.

Pengujian atas legitimasi dan batas epistemis penggunaan AI dalam pemidanaan menemukan pisau analisis yang sangat kontekstual melalui teori Utilitarianisme. Sebagai teori etika konsekuensial yang dipelopori oleh Jeremy Bentham dan John Stuart Mill, Utilitarianisme menempatkan batas

¹ Naeem AllahRakha, Sholahuddin Al Fatih, Khaydarov Shukhratjon Djumaevich, and Ruzinazarov Shuhrat Nuraliyevich, "Legal, Ethical, and Regulatory Challenges of AI Integration in the Judicial System: Between Efficiency and Fairness," *Prophetic Law Review* 8, no. 1 (2026): 83-104

² Rusdi, Muh Rustiawan Ardiansyah, dan Zainal Arifin Hoesein, "Filsafat Hukum di Era Kecerdasan Buatan: Menelaah Kembali Konsep Realitas, Norma, dan Kode dalam Kerangka Tiga Dunia Karl Popper," *Integrative Perspectives of Social and Science Journal* vol. 2, no. 1 (2025): 4677-4689.

³ Evellyn Octavia, Jeffaldo Thio, Sabrina Audelia, dan Yuwono Prianto, "Perluasan Tanggung Jawab Pencipta dan Pengguna Kecerdasan Buatan dalam Perspektif Filsafat Hukum," *JURAGAN* vol. 3, no. 1 (2026): 118-128.

⁴ Dhani Prasetyo dan Elsa Sofiana Rahma, "Transparansi Algoritma Prediktif dalam Proses Peradilan Pidana: Analisis Kesenjangan antara Akurasi Teknis dan Hak Asasi Manusia," *Jurnal Ilmiah Multi Disiplin Ilmu (MURAMU)* no. 2, no. 3 (2026): 36-48

⁵ Meghan J. Ryan, "Framing Individualized Sentencing for Politics and the Constitution," *American Criminal Law Review* no. 58, no. 4 (2021): 1747-1773.

⁶ Actuarial Justice," *The Sociology Dictionary*, 2026, <https://thesociologydictionary.com/sociology-dictionary/a/actuarial-justice/>, diakses 27 September 2026.

keabsahan moral dan hukum pada pencapaian kemanfaatan netto (*net utility*) dan kebahagiaan terbesar bagi jumlah masyarakat terbanyak (*the greatest happiness for the greatest number*).⁷ Dalam sains kuantitatif dan perumusan kebijakan publik, paradigma utilitarian dapat dijadikan pembenaran utama (*epistemic justification*) atas pengadopsian teknologi AI demi mengejar efisiensi penanganan perkara, penekanan biaya anggaran peradilan, dan kepastian hukum. Namun, bedah kritis filsafat hukum mengindikasikan bahwa kalkulasi utilitas agregat memiliki kelemahan mendasar terkait agregasi kebaikan. Logika utilitarian cenderung mentoleransi *margin of error* dan pengorbanan hak-hak mendasar individu atau kelompok rentan demi memaksimalkan skor utilitas masyarakat secara keseluruhan.⁸ Ketika seseorang dijatuhi bobot pemidanaan yang lebih berat karena adanya bias data statistik kelompoknya, martabat individu tersebut dikorbankan demi efisiensi sistemik.

Keabsahan hukum tidak hanya bergantung pada efisiensi prosedural, tetapi juga pada legitimasi moral dan kepercayaan publik. Ketika masyarakat mengetahui bahwa keputusan peradilan tidak lagi ditentukan oleh keadilan yang rasional dan adil, melainkan oleh kalkulasi *black box* yang dapat menciptakan diskriminasi, maka akan timbul rasa cemas, penderitaan sosial, dan erosi legitimasi terhadap institusi peradilan secara luas. Oleh karena itu, dilihat dari prinsip utilitas itu sendiri, penggunaan algoritma pemidanaan yang memuat *algorithmic bias* tidak dapat dilegitimasi secara filsafat hukum.

Maka, dekonstruksi terhadap *algorithmic bias* dalam pemidanaan melalui perspektif Utilitarianisme menjadi sangat mendesak. Dengan adanya penelitian ini yang bertujuan untuk menguji batas keabsahan epistemis AI dalam diskresi pemidanaan, serta merumuskan pembatasan filosofis dan normatif. Urgensi rekonstruksi ini mengarah pada penegakan doktrin *human-in-the-Loop* (HITL) dimana AI harus ditempatkan murni sebagai alat bantu keputusan (*decision-support*), bukan penentu keputusan (*decision-maker*), serta penetapan syarat *Explainable AI* (XAI) guna menjamin transparansi, akuntabilitas, dan perlindungan hak asasi manusia dalam sistem peradilan modern.

Rumusan Masalah

1. Bagaimana manifestasi *algorithmic bias* mengancam prinsip keadilan hukum material dan *due process of law* jika dibedah melalui lensa teori utilitarianisme?
2. Bagaimana rekonstruksi atas diskresi hakim agar penggunaan AI tidak mendegradasi nilai kemanusiaan dan martabat hukum?

METODE PENELITIAN

Penelitian ini menggunakan jenis penelitian hukum normatif (*doctrinal legal research*) yang difokuskan pada pengkajian asas-asas, doktrin, serta norma hukum yang berkaitan dengan integrasi teknologi *Artificial Intelligence* (AI) dalam sistem peradilan pidana. Menggunakan pendekatan filosofis (utilitarianisme) dan pendekatan konseptual. Sumber data yang digunakan dalam penelitian ini adalah sumber data sekunder dengan bahan hukum primer serta bahan hukum sekunder. Data yang dikumpulkan berdasarkan studi kepustakaan terhadap literatur ilmiah.

HASIL DAN PEMBAHASAN

Manifestasi *Algorithmic Bias* dalam Sistem Peradilan

Integrasi kecerdasan buatan dalam sistem peradilan pidana sering dijustifikasi melalui perhitungan kalkulus utilitarian yang menjanjikan efisiensi dan objektivitas pemidanaan. Doktrin

⁷ Mitra Madanchian dan Hamed Taherdoost, "Ethical Theories, Governance Models, and Strategic Frameworks for Responsible AI Adoption and Organizational Success," *Frontiers in Artificial Intelligence* 8 (2025): 1-14

⁸ Carlos Mougán dan Joshua Brand, "Kantian Deontology Meets AI Alignment: Towards Morally Grounded Fairness Metrics," *arXiv* (2024), <https://arxiv.org/abs/2311.05227>.

utilitarianisme menurut Bentham menyatakan bahwa hukum seharusnya bertujuan untuk mengoptimalkan kebahagiaan serta mengurangi penderitaan masyarakat secara keseluruhan. Tren pemanfaatan teknologi kecerdasan buatan dalam era digitalisasi peradilan pidana Indonesia sejatinya diproyeksikan untuk mengakselerasi pembaharuan hukum.⁹ Akan tetapi, penerapan algoritma penilaian risiko berpotensi mengganggu perhitungan kebahagiaan universal disebabkan oleh adanya bias pada data yang dimasukkan.¹⁰ Algoritma diproyeksikan sebagai penentu keputusan yang objektif, padahal data historis yang digunakannya menyimpan rekam jejak diskriminasi struktural masa lalu. Ketika data yang cacat epistemik ini diproses, mesin akan meratifikasi bias tersebut menjadi sebuah norma hukum baku yang seolah-olah valid secara ilmiah. Akibatnya, alih-alih menghadirkan kemanfaatan umum, algoritma justru melegitimasi ulang penderitaan bagi kelompok-kelompok terpinggirkan dalam peradilan.¹¹

Replikasi diskriminasi sistemik oleh algoritma peradilan berakibat langsung pada berkurangnya asas kemanfaatan (utility) dalam proporsi hukuman. Utilitarianisme mengharuskan agar setiap sanksi pidana yang diterapkan harus menghasilkan total kebahagiaan yang lebih tinggi dibandingkan dengan penderitaan yang ditimbulkannya. Masalah utama muncul ketika bias algoritmik tidak hanya mengganggu penetapan hukuman resmi, tetapi juga memengaruhi mekanisme keadilan restoratif, sehingga menimbulkan ancaman serius bagi komunitas marginal.¹² Apabila sistem kecerdasan buatan mengeluarkan skor risiko kejahatan yang terlalu tinggi pada kelompok yang rentan, maka dampak dari pemidanaan akan terasa tidak adil. Kesakitan orang yang dihukum berdasarkan asumsi statistik merusak kepercayaan masyarakat terhadap objektivitas lembaga penegakan hukum secara keseluruhan. Kerusakan terhadap legitimasi moral hukum ini secara kumulatif mengurangi nilai kebahagiaan bersih yang seharusnya dilindungi oleh pemerintah. Dengan demikian, anggapan mekanis dalam algoritma telah menghambat fungsi utama filsafat pemidanaan utilitarian itu sendiri.

Pergeseran paradigma dari keadilan yang dipersonalisasi menuju keadilan akuntabilitas akibat penggunaan algoritma bertentangan dengan prinsip utilitarianisme kualitatif yang diusulkan oleh John Stuart Mill. Ia menekankan signifikansi kualitas kebahagiaan dan perlindungan martabat individu agar tidak dikorbankan demi perhitungan agregat kasar. Rekonstruksi paradigma hukum di zaman revolusi digital dan perubahan etika publik mengharuskan agar aspek moral tetap diperhatikan dalam setiap otomatisasi kebijakan hukum.¹³ Namun, dalam peradilan prediktif berbasis pemrosesan data masif, keunikan konteks moral dan latar belakang subjektif seorang terdakwa sering kali diabaikan. Orang tidak lagi dihakimi berdasarkan kejahatan spesifik yang telah dilakukannya, tetapi berdasarkan prediktor statistik dari kelompok demografis mereka. Pendekatan aktuarial ini mengubah tingkat kemanusiaan individu menjadi sekedar rangkaian angka probabilitas yang kaku. Pembatasan terhadap hak kebebasan pribadi tanpa pertimbangan moral yang mendalam memicu penurunan kualitas keadilan yang sangat mendasar.

Efisiensi proses peradilan yang ditawarkan oleh teknologi kecerdasan buatan hanya merupakan ilusi kebahagiaan kuantitatif bagi masyarakat. Penganut utilitarianisme umum mungkin mengagumi kecepatan algoritma dalam mengurangi tumpukan kasus dan mengoptimalkan anggaran operasional peradilan.¹⁴ Akan tetapi, perhitungan utilitarian yang cermat harus mempertimbangkan pengaruh jangka panjang dari ketidakadilan yang dihasilkan secara sistematis oleh sistem. Penilaian

⁹ Mecca, A. S. P., Hidayat, W. A., & Tuasikal, H. (2025). Pemanfaatan teknologi kecerdasan buatan (artificial intelligence) dalam sistem peradilan pidana di Indonesia. *Jurnal Sosial Teknologi*, 5(6), 1730-1746.

¹⁰ Wibowo, I. A., & Kom, M. (2026). Peran AI dalam Penegakan Hukum. yayasan penerbit.

¹¹ O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group (NY).

¹² Ariyadi, F. (2025). Algorithmic Bias Dalam Restorative Justice: Ancaman Bagi Marginal. *National Multidisciplinary Sciences*, 4(3), 148-158.

¹³ Priyatno, H. D., & SH, M. (2026). Filsafat Revolusi Digital: Rekonstruksi Paradigma Hukum Di Era Teknologi. Dunia Penerbitan buku.

¹⁴ Jonathan, S. (2017). The new penology: Notes of the emerging strategy of corrections and its implications. In *Crime, Law and Society* (pp. 327-352). Routledge.

etika algoritma menunjukkan bahwa pilihan teknis AI selalu memiliki konsekuensi moral yang berpengaruh besar pada nilai keadilan publik.¹⁵ Ketika ketakutan akan kriminalisasi berbasis algoritma menyebar di masyarakat, rasa aman warga negara akan merosot secara drastis. Biaya sosial yang timbul akibat keresahan umum tersebut jauh melampaui keuntungan teknis berupa kecepatan penanganan perkara di pengadilan. Oleh karena itu, efisiensi mekanis algoritma tidak serta-merta dapat disetarakan dengan pencapaian kemanfaatan tertinggi bagi hukum.

Penetapan bobot pidana yang bersumber dari *algorithmic bias* secara nyata merusak keseimbangan antara tujuan pencegahan kejahatan dan perlindungan masyarakat. Utilitarianisme memperbolehkan hukuman hanya jika tindakan itu berfungsi dengan baik sebagai alat pencegahan atau deterrence terhadap kejahatan yang akan datang. Jika algoritma menjatuhkan sanksi kepada individu berdasarkan asumsi prediktif yang tidak sesuai dengan kenyataan objektif, fungsi edukasi dan pencegahan kejahatan akan lenyap. Terdakwa yang dijatuhi hukuman secara tidak adil akan mengalami ketidakpastian hukum yang menimbulkan keputusan dan ketidakpercayaan terhadap peradilan. Perasaan ketidakadilan yang meluas ini justru memicu kemungkinan adanya pembangkangan sosial baru di tengah-tengah tatanan masyarakat. Pada akhirnya, alat yang awalnya dibuat untuk menciptakan ketertiban umum justru menjadi penyebab kekacauan sosial yang baru.

Rekonstruksi Pembatasan Peran AI dalam Peradilan

Diskresi hakim pada dasarnya bukan merupakan kekuasaan sewenang-wenang, tetapi kebebasan yang memiliki struktur guna mengambil keputusan dalam batas norma hukum, terutama ketika undang-undang bersifat kabur atau bertentangan dengan norma lain. Diskresi ini menjadi hal yang penting bagi hakim untuk mempertimbangkan faktor-faktor individual terdakwa, misalnya latar belakang, kondisi psikologis, dan konteks kejahatan yang tidak selalu dapat ditangkap oleh algoritma prediktif. Ketika sistem AI mulai memasuki ranah pidana, diskresi hakim menghadapi tekanan ganda, di lain sisi, algoritma menawarkan konsistensi dan efisiensi, tetapi hal tersebut berisiko membekukan penilaian hakim ke dalam output teknis yang sulit ditantang.¹⁶

Penggunaan AI dalam sistem hukum pidana di Indonesia berisiko melemahkan asas keadilan prosedural (*due process of law*).¹⁷ Bias algoritmik dapat merusak keadilan prosedural karena proses analisis AI tidak selalu transparan dan sulit diuji dalam proses peradilan. Fenomena *black box* dan *automation bias* dapat menimbulkan ancaman bagi independensi hakim, karena hakim cenderung menerima *output* algoritma sebagai kebenaran objektif tanpa kritik memadai. Dalam konteks peradilan pidana Indonesia, penggunaan AI yang melebihi fungsinya bahkan berpotensi melanggar asas legalitas dan asas praduga tidak bersalah (*presumption of innocence*), yang dijamin dalam Pasal 8 UU No. 39 Tahun 1999 tentang Hak Asasi Manusia.¹⁸ Pergeseran paradigma peradilan dari *Due Process Model* yang mengutamakan prosedur dan hak asasi menuju *Crime Control Model* yang mengutamakan efisiensi penghukuman demi kepuasan publik menjadi ancaman nyata ketika algoritma mulai mendominasi proses pengambilan keputusan.¹⁹ Ketika negara menggantikan elemen penalaran normatif hakim dengan output algoritma, ini berisiko gagal menghormati status moral terdakwa sebagai agen yang memiliki martabat manusia.

Untuk menjaga martabat hukum dan nilai kemanusiaan, diperlukan rekonstruksi diskresi hakim berdasarkan prinsip "kendali bermakna" (*meaningful public control*). Prinsip ini memberikan

¹⁵ Hendrawan, J., & Nasution, M. K. (2025). Etika algoritma: Pertimbangan Moral dalam Kecerdasan Buatan dan Otomasi. Research Gate, (February).

¹⁶ Syamsul Fatoni, Erma Rusdiana, "Penggunaan Kecerdasan Buatan (*Artificial Intelligence*) sebagai Bahan Pertimbangan Putusan Hakim dalam Sistem Peradilan Pidana di Indonesia," *Jurnal Penelitian Hukum De Jure* vol. 24, no. 3 (2024): 247-264.

¹⁷ Salman Al Farisi, Sri Puji Ningsih, Arda Fairuzaki, Novita Mayasari, dan Salman Nurfari, "Kecerdasan Buatan dan Hak Asasi Manusia: Analisis Etika dan Kerangka Hukum di Era Digital," *Jurnal Hukum, Administrasi Publik dan Negara* vol. 3, no. 1 (2026): 107-12

¹⁸ Pasal 8 Undang-Undang Nomor 39 Tahun 1999 tentang Hak Asasi Manusia

¹⁹ Attarizza Satria Putra, Dudik Djaja Sidarta, Moh. Taufik, dan Dedi Wardana Nasution, "Narasi dan Stigma dalam Peradilan Pidana: Bagaimana Cerita Publik Mempengaruhi Penjatihan Hukuman," *Syiah Kuala Law Journal*, vol. 9 no. 3 (2025): 440-441

syarat bahwa hakim tetap harus mempertahankan tanggung jawab moral atas seluruh keputusan pidana, dan AI hanya berperan sebagai *decision support system*. Teknologi AI tidak dapat menggantikan peran hakim sebagai penafsir hukum, penjaga nilai keadilan, dan pemberi legitimasi moral terhadap putusan.²⁰ Masa depan *digital justice* di Indonesia tidak terletak pada dominasi algoritma, melainkan pada penguatan peran hakim sebagai pusat pertimbangan hukum yang tetap berlandaskan nilai kemanusiaan dan keadilan substantif.

Dalam kerangka ini, AI seharusnya diposisikan sebagai alat bantu, bukan pengganti hakim. Algoritma dapat digunakan untuk membantu mengidentifikasi ketidakadilan dalam putusan pidana, seperti disparitas yang tidak dapat dijustifikasi, diskriminasi, atau pelanggaran prinsip ekuitas, tetapi tidak dapat menggantikan penilaian normatif hakim. Hal ini yang menekankan bahwa AI tidak dapat menggantikan peran hakim dalam proses penemuan hukum yang secara inheren melibatkan pertimbangan etis, pengalaman manusia, dan penalaran moral.²¹ Dengan demikian, rekonstruksi diskresi hakim harus menegaskan bahwa:

1. AI hanya bekerja sebagai alat bantu analitis, bukan yang membuat keputusan final menggantikan kewenangan hakim.
2. Hakim wajib melakukan penilaian independen terhadap setiap *output* algoritma, termasuk kemampuan untuk menolaknya.
3. Transparansi dan akuntabilitas menjadi prasyarat mutlak sebelum algoritma digunakan dalam peradilan.
4. Martabat terdakwa sebagai agen moral harus diakui melalui adanya keterlibatan hakim yang berperan dalam setiap kasus yang ditangani, bukan hanya penerapan skor risiko.

Rekonstruksi ini penting untuk memastikan bahwa modernisasi sistem peradilan melalui AI tidak mengorbankan nilai-nilai fundamental hukum, keadilan, martabat manusia, dan *due process* hanya demi efisiensi teknologi.

SIMPULAN

Berdasarkan analisis filosofis berbasis teori utilitarianisme terhadap penerapan AI dalam sistem peradilan, dapat ditarik kesimpulan bahwa penerapan algoritma penilaian risiko (*risk-assessment algorithms*) dalam peradilan pidana terbukti merelokasi dan melegitimasi *historical bias* yang terekam dalam data masukan (*training data*). Dari perspektif utilitarianisme, kegagalan mekanis ini mencederai kalkulasi kebahagiaan universal (*greatest happiness*) karena menghasilkan distribusi penderitaan pidana yang tidak proporsional dan diskriminatif terhadap kelompok minoritas. Pergeseran ke arah *actuarial justice* melanggar prinsip utilitarian kualitatif John Stuart Mill, sebab mereduksi martabat serta konteks moral unik individu menjadi sekadar kalkulasi probabilitas statistik yang kaku.

Diskresi hakim adalah kebebasan yang terstruktur dan penting untuk memperhitungkan konteks moral serta nilai kemanusiaan terdakwa, yang tidak bisa disubstitusi oleh perhitungan teknis dari algoritma kecerdasan buatan (AI). Penerapan AI dalam sistem peradilan pidana Indonesia dapat membahayakan *prinsip due process of law*, otonomi hakim, serta asas praduga tidak bersalah akibat tantangan black box, bias otomatisasi, dan pergeseran paradigma ke arah *Crime Control Model*. Oleh karena itu, perlu ada rekonstruksi diskresi hakim yang berlandaskan prinsip kendali bermakna (*meaningful public control*), yang menjadikan AI semata-mata sebagai alat bantu analitis (*decision-support system*) dan bukan sebagai pengambil keputusan akhir. Akhirnya, modernisasi peradilan pidana

²⁰ Catur Yudianto, "Digital Transformation Of The Judiciary: Ethical And Legal Implications of The Use of AI in The Indonesian Criminal Justice System," *International Journal of Humanities, Social Sciences and Business (INJOSS)* vol. 5 no. 2 (2026): 316-324

²¹ Masinton Pasaribu, "Penerapan Kecerdasan Buatan (*Artificial Intelligence*) Dalam Proses Legislasi dan Sistem Peradilan di Indonesia," *SELISIK* vol. 10 no. 2 (2024)

perlu tetap menjadikan hakim sebagai pusat pemikiran normatif untuk menjaga martabat manusia, transparansi, dan keadilan substantif di atas efisiensi teknologi semata.

SARAN

Berdasarkan dari hasil pembahasan, penulis menyarankan agar penggunaan AI dalam sistem peradilan Indonesia tetap berlaku sebagai alat bantu (*decision support system*), dan bukan yang menggantikan posisi hakim, sehingga independensi dan otonomi penilaian hakim tetap terjaga. Hakim perlu diperkuat literasi dan etika mengenai AI agar dapat menggunakan output algoritma dengan kritis, dan tidak hanya menerimanya sebagai kebenaran objektif. Selain itu, diperlukan regulasi khusus yang isinya mengatur batas-batas penggunaan AI, kewajiban transparansi algoritma (*explainable AI*), dan adanya mekanisme audit berkala untuk mendeteksi serta mengevaluasi bias algoritmik. Kemudian, pengembangan dan penerapan AI dalam peradilan harus tetap menghormati hak asasi terdakwa, khususnya asas legalitas dan praduga tidak bersalah sebagaimana dijamin dalam UU No. 39 Tahun 1999 tentang Hak Asasi Manusia.

REFERENSI

- Actuarial justice. (2026). In *The Sociology Dictionary*. <https://thesociologydictionary.com/sociology-dictionary/a/actuarial-justice/>
- Al Farisi, S., Ningsih, S. P., Fairuzaki, A., Mayasari, N., & Nurfarizi, S. (2026). Kecerdasan buatan dan hak asasi manusia: Analisis etika dan kerangka hukum di era digital. *Jurnal Hukum, Administrasi Publik dan Negara*, 3(1), 107–112.
- AllahRakha, N., Al Fatih, S., Djumaevich, K. S., & Nuraliyevich, R. S. (2026). Legal, ethical, and regulatory challenges of AI integration in the judicial system: Between efficiency and fairness. *Prophetic Law Review*, 8(1), 83–104.
- Ariyadi, F. (2025). Algorithmic bias dalam restorative justice: Ancaman bagi marginal. *National Multidisciplinary Sciences*, 4(3), 148–158.
- Fatoni, S., & Rusdiana, E. (2024). Penggunaan kecerdasan buatan (*artificial intelligence*) sebagai bahan pertimbangan putusan hakim dalam sistem peradilan pidana di Indonesia. *Jurnal Penelitian Hukum De Jure*, 24(3), 247–264.
- Hendrawan, J., & Nasution, M. K. (2025). *Etika algoritma: Pertimbangan moral dalam kecerdasan buatan dan otomasi*. ResearchGate.
- Kamila, Z. (2025). Pengaturan hukum dan prospek penggunaan *artificial intelligence* dalam era digitalisasi sistem peradilan di Indonesia. *Jurnal Riset Multidisiplin Edukasi*, 2(3), 16–36.
- Madanchian, M., & Taherdoost, H. (2025). Ethical theories, governance models, and strategic frameworks for responsible AI adoption and organizational success. *Frontiers in Artificial Intelligence*, 8, 1–14.
- Mecca, A. S. P., Hidayat, W. A., & Tuasikal, H. (2025). Pemanfaatan teknologi kecerdasan buatan (*artificial intelligence*) dalam sistem peradilan pidana di Indonesia. *Jurnal Sosial Teknologi*, 5(6), 1730–1746.
- Mougan, C., & Brand, J. (2024). *Kantian deontology meets AI alignment: Towards morally grounded fairness metrics*. arXiv. <https://arxiv.org/abs/2311.05227>
- Octavia, E., Thio, J., Audelia, S., & Prianto, Y. (2026). Perluasan tanggung jawab pencipta dan pengguna kecerdasan buatan dalam perspektif filsafat hukum. *JURAGAN*, 3(1), 118–128.
- O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

- Prasetyo, D., & Rahma, E. S. (2026). Transparansi algoritma prediktif dalam proses peradilan pidana: Analisis kesenjangan antara akurasi teknis dan hak asasi manusia. *Jurnal Ilmiah Multi Disiplin Ilmu (MURAMU)*, 2(3), 36–48.
- Priyatno, H. D., & SH, M. (2026). *Filsafat revolusi digital: Rekonstruksi paradigma hukum di era teknologi*. Dunia Penerbitan Buku.
- Putra, A. S., Sidarta, D. D., Taufik, M., & Nasution, D. W. (2025). Narasi dan stigma dalam peradilan pidana: Bagaimana cerita publik mempengaruhi penjatuhan hukuman. *Syiah Kuala Law Journal*, 9(3), 440–441.
- Putra, C. Y. (2026). Digital transformation of the judiciary: Ethical and legal implications of the use of AI in the Indonesian criminal justice system. *International Journal of Humanities, Social Sciences and Business (INJOSS)*, 5(2), 316–324.
- Republik Indonesia. (1999). *Undang-Undang Nomor 39 Tahun 1999 tentang Hak Asasi Manusia*. Lembaran Negara Republik Indonesia Tahun 1999 Nomor 165, Tambahan Lembaran Negara Republik Indonesia Nomor 3886.
- Rusdi, R., Ardiansyah, M. R., & Hoesein, Z. A. (2025). Filsafat hukum di era kecerdasan buatan: Menelaah kembali konsep realitas, norma, dan kode dalam kerangka tiga dunia Karl Popper. *Integrative Perspectives of Social and Science Journal*, 2(1), 4677–4689.
- Ryan, M. J. (2021). Framing individualized sentencing for politics and the Constitution. *American Criminal Law Review*, 58(4), 1747–1773.
- Yopik Gani, S. I. P., & Sibarani, B. S. I. K. (2026). *Etika publik dan kebijakan publik*. Penerbit KBM Indonesia.