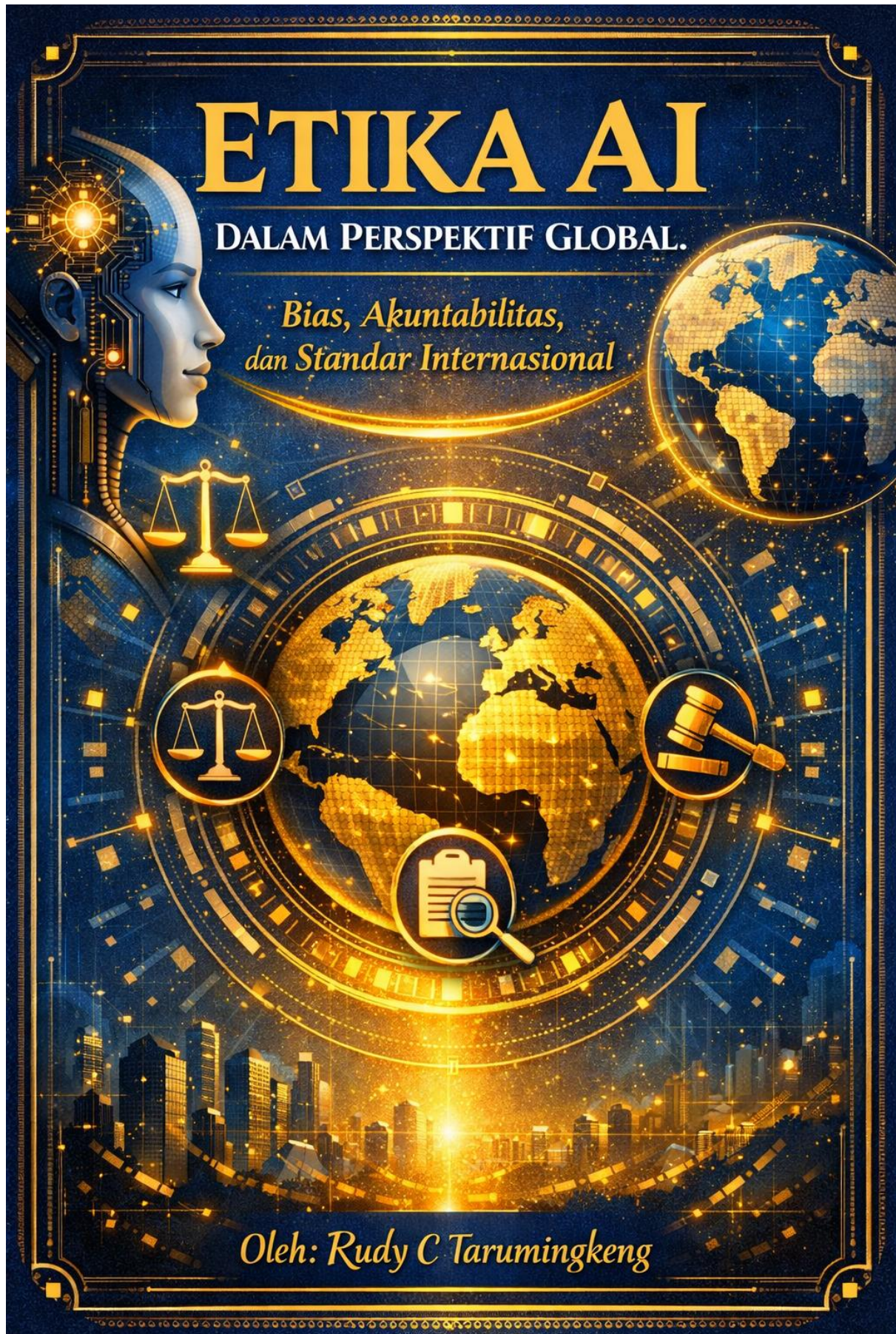




*Rudy C Tarumingkeng: Etika AI dalam Perspektif Global: Bias, Akuntabilitas, dan Standar Internasional*

Oleh:

[Prof Ir Rudy C Tarumingkeng, PhD](#)

Professor of Management NUP: 9903252922

Professor Emeritus, IPB-University

Rektor, Universitas Cenderawasih, Papua (1978-1988, dan

Rektor, Kampus AGRO Manokwari sekarang Universitas Papua Manokwari

Coordinator, CIDA/DIKTI SFU Burnaby BC Canada 1988-1991

Rektor, Universitas Kristen Krida Wacana, Jakarta (1991-2000)

Chairman. Board of Professors, IPB-University, Bogor (2005-2006)

AI - Data Analyst, dan Chairman, Academic Senate, IBM-ASMI, Jakarta 2024-

---

© RudyCT Academic Series

[rudyc75@gmail.com](mailto:rudyc75@gmail.com)

28 March 2026

## **ETIKA AI DALAM PERSPEKTIF GLOBAL: BIAS, AKUNTABILITAS, DAN STANDAR INTERNASIONAL**

### **Abstrak**

Perkembangan kecerdasan buatan atau artificial intelligence (AI) telah bergerak lebih cepat daripada kapasitas banyak organisasi dan negara untuk menyiapkan tata kelola yang matang. Dalam konteks global, persoalan etika AI tidak lagi dapat dipahami hanya sebagai isu teknis, melainkan sebagai persoalan sosial, hukum, politik, dan peradaban. Tiga isu paling menentukan dalam diskursus tersebut ialah bias, akuntabilitas, dan standardisasi internasional. Bias menunjukkan bahwa sistem AI dapat mereproduksi bahkan memperkuat ketimpangan historis ketika data, desain model, tujuan optimisasi, dan konteks implementasinya tidak diawasi secara kritis. Akuntabilitas menuntut kejelasan tentang siapa yang bertanggung jawab ketika sistem AI menimbulkan kerugian, diskriminasi, manipulasi, atau keputusan yang tidak dapat dijelaskan. Sementara itu, kemunculan beragam kerangka global—seperti Recommendation on the Ethics of Artificial Intelligence dari UNESCO, OECD AI Principles yang diperbarui, NIST AI Risk Management Framework, AI Act Uni Eropa, Framework Convention on Artificial Intelligence dari Council of Europe, serta standar ISO/IEC 42001, 23894, dan 42005—menunjukkan bahwa dunia sedang bergerak menuju tata kelola AI yang lebih terstruktur, meskipun belum sepenuhnya seragam. Makalah ini membahas etika AI dari perspektif global dengan menempatkan manusia, hak asasi, keadilan, dan tata kelola lintas yurisdiksi sebagai pusat analisis. Makalah ini berargumen bahwa etika AI yang memadai harus melampaui slogan “responsible AI”

dan diterjemahkan ke dalam mekanisme kelembagaan yang konkret: penilaian dampak, dokumentasi model, pengawasan manusia, audit, transparansi, pelacakan insiden, dan akses terhadap pemulihan. Pada akhirnya, masa depan AI yang etis bukan ditentukan semata oleh kecanggihan model, melainkan oleh kualitas institusi, standar, dan keberanian moral masyarakat global untuk membatasi, mengarahkan, dan menilai teknologi secara bertanggung jawab. ([UNESCO](#))

### **Kata kunci**

Etika AI; bias algoritmik; akuntabilitas; tata kelola AI; standar internasional; UNESCO; OECD; NIST; EU AI Act; ISO/IEC.

## **Etika AI dalam Perspektif Global: Bias, Akuntabilitas, dan Standar Internasional**

### **Pendahuluan**

Dalam beberapa tahun terakhir, AI telah berubah dari sekadar teknologi bantu menjadi infrastruktur keputusan. Ia hadir dalam rekrutmen, penilaian kredit, diagnosis medis, moderasi konten, pengenalan wajah, pendidikan digital, layanan publik, logistik, sampai pembuatan konten generatif. Pergeseran ini menyebabkan etika AI tidak lagi bisa dibicarakan di pinggiran inovasi, sebab AI kini masuk ke jantung proses sosial yang menentukan akses warga terhadap kesempatan, hak, dan sumber daya. Karena itu, pertanyaan etis tentang AI bukan hanya "apakah sistem ini akurat," melainkan juga "siapa yang diuntungkan," "siapa yang dirugikan," "siapa yang bertanggung jawab," dan "standar apa yang dipakai untuk menilainya." Kerangka global yang kini berkembang hampir semuanya menempatkan hak asasi manusia, fairness, transparansi, akuntabilitas, dan pengawasan manusia sebagai prinsip sentral. UNESCO, misalnya, menegaskan bahwa perlindungan hak asasi manusia dan martabat manusia merupakan fondasi rekomendasi globalnya tentang etika AI; OECD

menekankan AI yang inovatif, dapat dipercaya, dan menghormati hak asasi serta nilai demokratis; sedangkan NIST menekankan karakter AI yang dapat dipercaya, termasuk fairness, akuntabilitas, transparansi, keamanan, dan privasi. ([UNESCO](#))

Dari perspektif global, etika AI berkembang dalam situasi yang paradoksal. Di satu sisi, hampir semua pemerintah dan organisasi besar berbicara tentang "responsible AI." Di sisi lain, dunia menghadapi ledakan variasi regulasi, perbedaan kapasitas kelembagaan, kesenjangan data, dominasi beberapa aktor teknologi, dan perbedaan budaya hukum antara Global North dan Global South. OECD mencatat bahwa prinsip-prinsipnya telah diadopsi oleh negara-negara anggota dan sejumlah mitra global, serta menginspirasi lebih dari seribu inisiatif kebijakan pada lebih dari 70 yurisdiksi. Namun, banyaknya inisiatif tidak otomatis berarti adanya konsistensi implementasi. Justru di titik inilah isu standardisasi internasional menjadi penting: dunia membutuhkan interoperabilitas kebijakan tanpa menghapus konteks lokal, dan membutuhkan harmonisasi prinsip tanpa memaksakan homogenisasi budaya. ([OECD](#))

Makalah ini berangkat dari tesis bahwa problem etika AI harus dibaca sebagai problem sosio-teknis. Artinya, bias AI bukan sekadar cacat matematis, akuntabilitas bukan sekadar catatan kontraktual, dan standar internasional bukan sekadar dokumen formal. Ketiganya berkaitan dengan relasi kuasa, desain organisasi, kualitas data, struktur pasar, rezim hukum, dan kapasitas institusi publik. Karena itu, analisis etika AI memerlukan pembacaan yang melintasi disiplin ilmu: manajemen, hukum, kebijakan publik, ilmu komputer, studi organisasi, dan etika terapan.

Secara sistematis, makalah ini akan membahas empat hal utama. Pertama, kerangka konseptual etika AI dalam perspektif global. Kedua, bias sebagai masalah epistemik, teknis, dan sosial. Ketiga, akuntabilitas sebagai prasyarat legitimasi teknologi. Keempat, peta standar dan rezim internasional yang kini membentuk lanskap tata kelola AI dunia. Di bagian akhir, makalah ini akan menyoroti implikasi bagi negara berkembang dan

secara khusus relevansinya bagi organisasi, perguruan tinggi, dan pembuat kebijakan di Indonesia.

## **1. Kerangka Konseptual Etika AI dalam Perspektif Global**

Etika AI secara umum dapat dipahami sebagai upaya menilai dan mengarahkan perancangan, pengembangan, penerapan, penggunaan, serta pengawasan AI agar selaras dengan nilai-nilai kemanusiaan. Namun secara praktis, etika AI global dibangun di atas beberapa lapis norma. Lapis pertama adalah norma moral, seperti keadilan, non-diskriminasi, keselamatan, dan penghormatan terhadap martabat. Lapis kedua adalah norma hukum, yaitu kewajiban yang lahir dari regulasi, perjanjian, dan prosedur administratif. Lapis ketiga adalah norma teknis-organisasional, berupa standar, audit, dokumentasi, pengujian, dan mekanisme kontrol. Dengan demikian, etika AI yang serius selalu bergerak dari prinsip menuju proses, dan dari proses menuju bukti kepatuhan.

UNESCO menempuh pendekatan yang sangat normatif dan luas. Recommendation on the Ethics of Artificial Intelligence, yang diadopsi pada 2021 dan berlaku bagi 194 negara anggota UNESCO, menjadikan perlindungan hak asasi manusia dan martabat manusia sebagai batu penjur, sambil menekankan transparansi, fairness, dan pentingnya human oversight. Di sini terlihat bahwa etika AI global tidak lahir dari kekhawatiran teknologis semata, melainkan dari kesadaran bahwa AI berpotensi memengaruhi struktur peluang hidup manusia. Kerangka seperti ini penting karena mengingatkan bahwa keberhasilan AI tidak boleh dinilai hanya dari efisiensi atau skala adopsi, melainkan dari kualitas relasinya dengan manusia dan masyarakat. ([UNESCO](#))

OECD mengambil jalan yang lebih kebijakan-operasional. Prinsip-prinsip AI OECD menekankan inclusive growth, sustainable development, human rights, fairness, privacy, transparency, robustness, accountability, dan tanggung jawab aktor AI sepanjang siklus hidup sistem. Revisi OECD pada 2024 penting karena memasukkan konteks general-purpose dan generative AI, menyoroti mis/disinformation, keselamatan, transparansi,

sustainability, dan terutama interoperabilitas kebijakan antar-yurisdiksi. Ini menandakan bahwa diskursus etika AI telah bergeser: bukan lagi hanya soal algoritme yang dipakai untuk tugas sempit, tetapi juga model umum yang menjadi fondasi ekosistem digital lintas sektor. ([OECD](#))

NIST, walaupun berasal dari konteks Amerika Serikat dan bersifat sukarela, memberi kontribusi penting melalui AI Risk Management Framework. NIST mendefinisikan trustworthiness AI melalui karakteristik seperti valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-enhanced, serta fair with harmful bias managed. Kerangka ini sangat berguna karena menerjemahkan bahasa etika ke dalam fungsi organisasi: govern, map, measure, dan manage. Dengan kata lain, etika tidak berhenti pada deklarasi nilai, tetapi dioperasionalisasikan melalui tata kelola, pemetaan dampak, pengukuran risiko, dan manajemen respons. ([NIST Publications](#))

Dalam perspektif global, ketiga kerangka tersebut memperlihatkan satu pola umum. Pertama, etika AI makin dipahami sebagai isu lifecycle, bukan isu tahap pengembangan semata. Kedua, yang dinilai bukan hanya model, tetapi juga organisasi, pemasok data, pengguna, dan konteks implementasi. Ketiga, ukuran keberhasilan AI yang etis bukan hanya absence of harm, tetapi juga adanya mekanisme akuntabilitas, transparansi, dan remedy. Di titik ini, etika AI global bertemu dengan tata kelola modern: teknologi yang kuat menuntut institusi yang kuat.

## **2. Bias AI: Dari Data ke Ketidakadilan Sosial**

Bias adalah salah satu isu paling sering dibicarakan dalam etika AI, tetapi juga salah satu yang paling sering disederhanakan. Banyak organisasi mengira bias hanya muncul jika data latih "kurang seimbang." Pandangan ini terlalu sempit. Bias dapat muncul pada banyak lapis: definisi masalah, pemilihan variabel, tujuan optimisasi, desain label, kualitas data, representasi kelompok, ambang keputusan, antarmuka pengguna, hingga praktik institusional di tempat sistem itu diterapkan. Karena itu, bias AI

sebetulnya adalah hasil pertemuan antara statistik, organisasi, dan sejarah sosial.

Bayangkan sebuah sistem rekrutmen yang dilatih menggunakan data karyawan "berprestasi" dari sepuluh tahun terakhir. Secara teknis, model itu mungkin berhasil meniru pola historis. Namun jika data historis mencerminkan preferensi terhadap lulusan kampus tertentu, jenis kelamin tertentu, atau gaya komunikasi tertentu, maka model akan memproduksi masa depan yang ditawan oleh masa lalu. Di sini bias bukan terjadi karena mesin "jahat," tetapi karena organisasi gagal membedakan antara pola historis dan norma keadilan. AI lalu berubah menjadi mesin konservasi ketimpangan.

Dalam sektor keuangan, skenario serupa bisa terjadi. Sistem penilaian risiko kredit yang bergantung pada riwayat formal mungkin mendiskriminasi pekerja informal, usaha mikro, atau kelompok yang akses perbankannya memang secara historis terbatas. Dalam pendidikan, sistem proctoring otomatis dapat salah menandai mahasiswa dengan kondisi pencahayaan buruk, ekspresi wajah tertentu, atau keterbatasan fisik sebagai perilaku mencurigakan. Dalam layanan publik, model prediksi risiko dapat menempatkan kelompok rentan dalam siklus pengawasan berlebihan. Setiap contoh ini menunjukkan bahwa bias bukan sekadar error; ia dapat menjelma menjadi distribusi ketidakadilan yang sistemik.

Kerangka internasional merespons persoalan ini secara cukup konsisten. UNESCO menekankan fairness dan human oversight. OECD menyebut non-discrimination, equality, diversity, fairness, social justice, privacy, dan human-centred values sebagai bagian dari prinsip dasarnya. NIST lebih operasional: fairness and bias harus dievaluasi dan hasilnya didokumentasikan; risiko terhadap individu, kelompok, komunitas, organisasi, dan ekosistem sosial perlu dipetakan; dan tim pengambil keputusan harus beragam agar risiko tidak dinilai secara sempit. Artinya, standar global kini tidak lagi menganggap bias sebagai isu opsional. Ia adalah variabel inti tata kelola AI. ([UNESCO](#))

Hal yang penting untuk dipahami ialah bahwa fairness sendiri tidak tunggal. Secara praktis, organisasi dapat menghadapi konflik antara fairness statistik, fairness prosedural, dan fairness substantif. Sebuah model mungkin memenuhi ukuran statistik tertentu tetapi tetap dirasakan tidak adil oleh kelompok yang terkena dampak. Sebaliknya, model yang dijelaskan secara transparan belum tentu menghasilkan outcome yang adil. Inilah sebabnya NIST, ISO, dan OHCHR sama-sama mendorong pendekatan penilaian dampak dan dokumentasi, bukan hanya evaluasi performa model. ISO/IEC 42005, misalnya, secara eksplisit memosisikan AI system impact assessment untuk memahami bagaimana sistem AI dapat memengaruhi individu, kelompok, dan masyarakat sepanjang lifecycle-nya. ([ISO](#))

Karena itu, upaya mengurangi bias tidak dapat dibatasi pada “membersihkan data.” Organisasi harus menanyakan: data siapa yang hilang, siapa yang mendefinisikan keberhasilan model, apa konsekuensi kesalahan bagi kelompok berbeda, apakah ada non-AI alternative, dan apakah individu yang terdampak memiliki saluran keberatan. Etika AI yang matang selalu memandang bias sebagai masalah desain institusional.

### **3. Akuntabilitas: Siapa Bertanggung Jawab Ketika AI Salah?**

Akuntabilitas adalah pilar kedua etika AI global. Dalam sistem keputusan konvensional, rantai tanggung jawab biasanya lebih jelas: ada pejabat, manajer, profesional, atau lembaga yang dapat dimintai penjelasan. AI mengganggu kejelasan ini karena keputusan sering dihasilkan melalui jaringan aktor yang lebih kompleks: pengembang model, penyedia komputasi, vendor data, integrator sistem, pengguna organisasi, dan pihak ketiga yang memodifikasi atau mengoperasikan sistem. Akibatnya, ketika terjadi kerugian, semua pihak dapat saling menunjuk. Di sinilah “akuntabilitas” tidak boleh dipahami sekadar sebagai prinsip abstrak, melainkan sebagai arsitektur tanggung jawab.

NIST memberi kontribusi penting dengan menempatkan accountable and transparent sebagai karakteristik trustworthiness, lalu menghubungkannya

dengan fungsi GOVERN, MAP, MEASURE, dan MANAGE. Dalam fungsi GOVERN, peran dan tanggung jawab terkait pemetaan, pengukuran, dan pengelolaan risiko AI harus terdokumentasi dan jelas di seluruh organisasi; kepemimpinan eksekutif harus mengambil tanggung jawab atas keputusan terkait risiko; dan organisasi perlu membangun budaya kritis serta safety-first. Ini sangat penting bagi manajemen modern: AI tidak boleh dianggap sebagai proyek teknis yang “diserahkan ke tim data,” melainkan sebagai urusan strategis yang harus naik ke level dewan, pimpinan, dan tata kelola risiko perusahaan. ([NIST Publications](#))

Dalam praktik, akuntabilitas AI sekurang-kurangnya memiliki tujuh dimensi. Pertama, akuntabilitas desain: siapa yang menetapkan tujuan model dan memilih data. Kedua, akuntabilitas dokumentasi: apakah keputusan pengembangan, keterbatasan model, asumsi, dan risiko telah dicatat. Ketiga, akuntabilitas operasional: siapa yang memantau sistem setelah diterapkan. Keempat, akuntabilitas penjelasan: apakah organisasi mampu menjelaskan keluaran model secara memadai. Kelima, akuntabilitas remedial: apakah ada mekanisme koreksi, banding, atau kompensasi. Keenam, akuntabilitas pemasok: bagaimana risiko pihak ketiga diatur. Ketujuh, akuntabilitas publik: bagaimana organisasi melaporkan insiden dan membangun kepercayaan.

Contoh naratif dapat memperjelas hal ini. Misalkan sebuah rumah sakit menggunakan AI untuk membantu penentuan prioritas pasien. Sistem bekerja cukup baik secara rata-rata, tetapi beberapa pasien dari kelompok tertentu secara konsisten menerima prioritas lebih rendah karena data latihan kurang mewakili kondisi klinis mereka. Siapa yang bertanggung jawab? Vendor mungkin berkata model hanya alat bantu. Rumah sakit dapat mengatakan keputusan akhir tetap pada dokter. Tim dokter dapat menyatakan mereka terlalu mengandalkan sistem karena kebijakan internal mendorong efisiensi. Dari sudut etika AI, jawaban yang benar bukan memilih satu pihak dan membebaskan yang lain, tetapi membangun

pembagian tanggung jawab yang dapat ditelusuri sejak awal. Tanpa itu, AI menciptakan “kabut organisasi” yang menutupi tanggung jawab moral.

Karena itu, akuntabilitas AI memerlukan tiga instrumen utama. Pertama, model governance: kebijakan internal, register risiko, kontrol perubahan, dokumentasi model, dan penunjukan pemilik risiko. Kedua, impact governance: penilaian dampak terhadap hak, kelompok rentan, keselamatan, privasi, dan integritas informasi. Ketiga, incident governance: mekanisme deteksi, pelaporan, investigasi, dan pembelajaran pasca-insiden. NIST GenAI Profile menambahkan urgensi baru pada isu ini dengan menyoroti risiko synthetic content, provenance, dan pelacakan sumber untuk memperkuat integritas informasi dan kepercayaan publik. ([NIST](#))

Dalam dunia global, akuntabilitas juga berkaitan dengan lintas batas. Banyak model dikembangkan di satu negara, dilatih pada data multinasional, di-host di wilayah lain, lalu digunakan oleh organisasi di negara berkembang. Ini berarti tanggung jawab hukum dan moral dapat terfragmentasi oleh yurisdiksi. Maka, standar internasional dibutuhkan bukan hanya untuk membatasi risiko, tetapi juga untuk mencegah kekosongan tanggung jawab.

#### **4. Standar dan Rezim Internasional: Peta Global Tata Kelola AI**

##### **4.1 UNESCO: Fondasi Normatif Universal**

UNESCO menempati posisi istimewa karena Recommendation on the Ethics of Artificial Intelligence merupakan standar global pertama di bidang etika AI dan berlaku bagi 194 negara anggota. Sifatnya universal dan berakar pada hak asasi manusia, martabat manusia, fairness, transparency, dan human oversight. Di tingkat filosofis, kerangka UNESCO mengingatkan bahwa AI harus dinilai dari sudut kemanusiaan, bukan hanya dari sudut produktivitas. Ini penting terutama bagi negara-negara yang masih membangun kapasitas digital, sebab UNESCO menyediakan bahasa

normatif bersama yang dapat dijadikan titik tolak sebelum ada perangkat regulasi domestik yang rinci. ([UNESCO](#))

Kekuatan UNESCO terletak pada keluasan visi. Ia tidak memaksa pendekatan teknis tunggal, tetapi memberikan horizon etis bersama. Keterbatasannya justru pada sifat rekomendatif: keberhasilan implementasinya sangat bergantung pada kemauan politik, kapasitas birokrasi, dan penerjemahan ke kebijakan sektoral. Namun dalam tata kelola global, sebuah norma universal sering kali lebih penting sebagai kompas daripada sebagai manual teknis.

#### **4.2 OECD: Interoperabilitas Kebijakan dan Siklus Hidup AI**

OECD AI Principles merupakan referensi utama bagi banyak pemerintah karena menjembatani nilai dan kebijakan. OECD menyatakan bahwa AI harus inovatif dan trustworthy serta menghormati human rights and democratic values. Revisi 2024 menambahkan fokus pada general-purpose dan generative AI, mis/disinformation, keselamatan, transparansi, sustainability, dan interoperable governance. OECD juga menekankan lifecycle approach dan responsible business conduct sepanjang rantai pasok AI. ([OECD](#))

Nilai utama OECD ada pada perannya sebagai jembatan antar-yurisdiksi. Dalam dunia di mana negara-negara bergerak dengan kecepatan regulasi berbeda, prinsip yang cukup fleksibel tetapi cukup konkret menjadi sangat penting. OECD berhasil memainkan fungsi ini karena ia menyediakan kosakata bersama bagi pembuat kebijakan, regulator, dan organisasi bisnis.

#### **4.3 NIST: Operasionalisasi Risiko dan Tata Kelola Organisasi**

NIST AI RMF 1.0 dan GenAI Profile 2024 merupakan contoh bagaimana etika diterjemahkan menjadi tata kelola operasional. AI yang trustworthy menurut NIST harus valid, safe, secure, accountable, transparent, explainable, privacy-enhanced, dan fair with harmful bias managed. Untuk mencapai itu, organisasi didorong menjalankan fungsi govern, map, measure, dan manage. GenAI Profile kemudian memperluas perhatian

pada risiko khas AI generatif, termasuk deepfakes, synthetic content, provenance, integritas informasi, dan mekanisme transparansi digital. ([NIST Publications](#))

Kerangka NIST sangat berguna untuk organisasi bisnis, perguruan tinggi, rumah sakit, lembaga publik, dan penyedia teknologi karena tidak berhenti di level prinsip. Ia meminta bukti: peran terdokumentasi, evaluasi bias, penilaian privasi, pengelolaan third-party risk, dan rencana response-recovery. Dalam bahasa manajemen, ini menjadikan etika AI sebagai bagian dari enterprise risk management.

#### **4.4 Uni Eropa: Dari Prinsip ke Kekuatan Hukum**

EU AI Act menandai fase baru karena etika AI diubah menjadi kewajiban hukum yang komprehensif. Komisi Eropa menyebut AI Act sebagai kerangka hukum AI pertama dan paling komprehensif di dunia, berbasis risk-based approach. Act ini mulai berlaku 1 Agustus 2024, dengan penerapan bertahap: larangan praktik AI tertentu dan kewajiban AI literacy mulai berlaku pada 2 Februari 2025; kewajiban untuk general-purpose AI models berlaku sejak 2 Agustus 2025; dan penerapan umum penuh terjadi pada 2 Agustus 2026, dengan beberapa sistem high-risk tertentu memiliki masa transisi lebih panjang. ([Digital Strategy](#))

Signifikansi EU AI Act terletak pada beberapa hal. Pertama, ia mengklasifikasikan penggunaan AI menurut tingkat risiko. Kedua, ia menautkan tata kelola AI dengan fundamental rights. Ketiga, ia memberi perhatian khusus pada GPAI dan model dengan systemic risk. Keempat, ia mulai membangun mekanisme implementasi melalui AI Office, guidelines, dan code of practice. Bagi organisasi global, AI Act memiliki efek ekstrateritorial praktis: sekalipun perusahaan berbasis di luar Eropa, bila produknya masuk pasar UE, standar tersebut menjadi rujukan penting. Ini memperkuat tren bahwa etika AI tidak lagi bisa dipisahkan dari kepatuhan regulasi.

#### **4.5 Council of Europe: Perjanjian Mengikat Berbasis Hak Asasi, Demokrasi, dan Rule of Law**

Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law merupakan tonggak penting karena merupakan perjanjian internasional yang mengikat secara hukum pertama di bidang ini. Konvensi ini dibuka untuk penandatanganan pada 5 September 2024 dan bertujuan memastikan bahwa aktivitas sepanjang lifecycle AI konsisten dengan hak asasi manusia, demokrasi, dan rule of law, sambil tetap kondusif bagi inovasi. Konvensi ini juga bersifat technology-neutral, sehingga dimaksudkan untuk bertahan menghadapi perubahan teknologi yang cepat. ([Portal](#))

Nilai tambah konvensi ini adalah penekanannya pada kerangka konstitusional dan hak publik, bukan hanya keselamatan produk atau kepatuhan korporasi. Dengan demikian, AI dipahami sebagai isu tata kelola demokratis, bukan sekadar tata kelola industri.

#### **4.6 ISO/IEC: Standar Manajemen, Risiko, dan Penilaian Dampak**

Standar ISO/IEC memberi lapisan yang berbeda: bukan terutama norma politik, melainkan arsitektur manajemen. ISO/IEC 42001:2023 adalah standar manajemen sistem AI yang meminta organisasi menempatkan kebijakan, prosedur, dan governance AI secara sistematis dengan pendekatan Plan-Do-Check-Act. ISO secara eksplisit menyebut bahwa standar ini membantu governance yang sound dan pengelolaan risiko serta peluang AI di seluruh organisasi. ISO/IEC 23894:2023 memberi panduan risk management spesifik AI, sedangkan ISO/IEC 42005:2025 memberi panduan untuk AI system impact assessment yang berfokus pada dampak terhadap individu, kelompok, dan masyarakat di sepanjang lifecycle sistem AI. ([ISO](#))

Bagi organisasi, kombinasi 42001, 23894, dan 42005 sangat strategis. 42001 membangun sistem manajemen; 23894 membingkai proses pengelolaan risiko; dan 42005 memperdalam analisis dampak sosial dan

kemanusiaan. Dengan kata lain, ISO menyediakan jembatan dari governance ke practice, dari prinsip ke auditability.

#### **4.7 PBB dan Hak Asasi: Due Diligence dan Tata Kelola Inklusif**

Di tingkat PBB, arah umum tata kelola AI bergerak ke pendekatan yang lebih inklusif dan berbasis hak. Global Digital Compact diposisikan sebagai kerangka global komprehensif untuk kerja sama digital dan governance of artificial intelligence, serta diadopsi bersama Pact for the Future pada 2024. Sementara itu, panduan OHCHR tentang human rights due diligence for digital technology use menekankan pendekatan berbasis risiko yang mencakup seluruh lifecycle dan value chain teknologi digital, dengan prioritas pada human rights risk assessments. Ini penting karena memperluas diskusi etika AI dari isu “teknologi yang baik” menjadi “kewajiban kelembagaan untuk mengidentifikasi, mencegah, memitigasi, melacak, dan mengomunikasikan dampak hak asasi.” ([United Nations](#))

### **5. Tantangan Global: Fragmentasi, Kesenjangan Kapasitas, dan Suara Global South**

Meskipun standar internasional makin banyak, dunia belum memiliki satu rezim AI global yang tunggal. Tantangan pertama adalah fragmentasi. Uni Eropa bergerak dengan instrumen hukum yang keras; OECD dengan prinsip kebijakan; NIST dengan kerangka sukarela; ISO dengan standar manajemen; UNESCO dengan norma universal; dan Council of Europe dengan treaty berbasis hak. Semua ini saling melengkapi, tetapi juga dapat membebani organisasi, terutama yang beroperasi lintas negara. Di satu sisi, keberagaman ini sehat karena memungkinkan eksperimentasi kelembagaan. Di sisi lain, tanpa interoperabilitas yang baik, organisasi bisa terjebak dalam compliance overload tanpa benar-benar meningkatkan keadilan dan keselamatan.

Tantangan kedua adalah kesenjangan kapasitas. Negara-negara maju relatif lebih siap dalam hal infrastruktur data, lembaga pengawas, sumber daya audit, keahlian hukum-teknis, dan posisi tawar terhadap perusahaan

teknologi besar. Sebaliknya, banyak negara berkembang justru menghadapi situasi di mana AI diadopsi lebih cepat daripada kemampuan untuk menilai dampaknya. Dalam konteks seperti itu, etika AI berisiko berubah menjadi impor normatif: prinsip datang dari luar, tetapi kapasitas implementasi tidak tumbuh secara memadai di dalam.

Tantangan ketiga ialah representasi Global South. Banyak dataset, benchmark, dan asumsi desain AI berasal dari konteks sosial-bahasa tertentu. Jika standar global tidak sungguh-sungguh menyertakan pengalaman negara berkembang, maka AI yang “etis” menurut dokumen formal bisa tetap tidak adil di lapangan. Contoh sederhana terlihat pada bahasa, akses digital, identitas informal, kualitas data administrasi, dan konteks budaya keputusan. Sistem yang tampak netral di negara maju bisa menjadi eksklusif ketika diterapkan pada masyarakat dengan infrastruktur lemah, dokumen identitas terbatas, atau keberagaman bahasa tinggi.

Karena itu, interoperabilitas global tidak boleh diartikan sebagai uniformitas tanpa konteks. Yang dibutuhkan adalah arsitektur bertingkat: prinsip universal di puncak, standar manajemen di tingkat organisasi, dan adaptasi lokal di tingkat implementasi. UNESCO, OECD, ISO, NIST, dan PBB justru paling berguna bila dibaca secara komplementer, bukan kompetitif.

## **6. Relevansi bagi Indonesia: Dari Prinsip ke Tata Kelola Institusional**

Bagi Indonesia, diskusi etika AI sebaiknya tidak berhenti pada kekaguman terhadap inovasi atau kekhawatiran normatif yang abstrak. Yang lebih penting ialah membangun tata kelola institusional yang mampu menerjemahkan prinsip global ke praktik lokal. Indonesia berada dalam posisi strategis: adopsi AI tumbuh di sektor keuangan, pendidikan, kesehatan, pemasaran, pemerintahan digital, dan UMKM; tetapi tantangan data, literasi, kepatuhan, dan koordinasi antarlembaga juga masih nyata. Karena itu, pendekatan yang paling realistis adalah membangun “governance by design.”

Untuk organisasi bisnis, langkah awalnya ialah membentuk AI governance structure yang jelas: dewan atau komite etik, pemilik risiko, kebijakan penggunaan AI, register model, persyaratan vendor, prosedur validasi, dan mekanisme pelaporan insiden. NIST dan ISO sangat berguna di sini karena keduanya menawarkan kerangka operasional yang dapat diintegrasikan ke manajemen risiko perusahaan. Untuk instansi pemerintah, fokus perlu diletakkan pada penilaian dampak terhadap hak warga, mekanisme pengaduan, dokumentasi keputusan otomatis, dan audit berkala untuk sistem yang memengaruhi akses terhadap layanan publik. Untuk perguruan tinggi, agenda utamanya ialah kurikulum literasi AI, kebijakan penggunaan AI generatif, integritas akademik, dan riset lintas disiplin tentang keadilan algoritmik. ([NIST Publications](#))

Dalam konteks pendidikan manajemen, etika AI juga perlu diajarkan bukan hanya sebagai mata kuliah filsafat teknologi, tetapi sebagai bagian dari governance, leadership, risk management, strategic decision-making, dan organizational accountability. Seorang manajer masa depan tidak cukup paham tentang prompt atau dashboard analitik. Ia harus mampu bertanya: apakah sistem ini menciptakan ketidakadilan tersembunyi, apakah ada human override, bagaimana risiko vendor dikelola, apakah keputusan dapat diaudit, dan bagaimana organisasi merespons bila terjadi kesalahan. Dengan kata lain, etika AI harus menjadi kompetensi manajerial.

Indonesia juga perlu berhati-hati agar tidak mengadopsi AI tanpa kerangka evaluasi dampak. Misalnya, bila AI dipakai untuk seleksi penerima bantuan, penilaian kredit UMKM, atau analitik pendidikan, maka penilaian fairness dan bias harus dilakukan sebelum dan sesudah deployment. Di sinilah ISO/IEC 42005 dan pendekatan human rights due diligence menjadi sangat relevan, karena keduanya mengingatkan bahwa dampak AI tidak berhenti pada performa model, tetapi menjalar ke kehidupan sosial yang konkret. ([ISO](#))

## **Penutup**

Etika AI dalam perspektif global pada akhirnya adalah persoalan tentang bagaimana umat manusia memilih hidup bersama teknologi yang semakin otonom, prediktif, dan persuasif. Bias memperlihatkan bahwa AI dapat menyerap sejarah ketimpangan dan memproyeksikannya ke masa depan. Akuntabilitas memperlihatkan bahwa tanpa struktur tanggung jawab yang jelas, teknologi canggih dapat menghasilkan kekosongan moral. Standar internasional memperlihatkan bahwa dunia sedang berusaha menjawab tantangan ini melalui kombinasi prinsip universal, regulasi berbasis risiko, kerangka manajemen, dan due diligence hak asasi.

Peta global yang ada saat ini cukup jelas. UNESCO menyediakan fondasi normatif universal berbasis martabat dan hak asasi. OECD menyediakan prinsip kebijakan yang interoperabel. NIST menerjemahkan etika ke dalam manajemen risiko yang operasional. EU AI Act mengubah prinsip menjadi kewajiban hukum. Council of Europe menghadirkan treaty yang mengikat dengan basis demokrasi dan rule of law. ISO/IEC menyediakan alat manajemen, risiko, dan impact assessment. PBB dan OHCHR menegaskan bahwa tata kelola AI harus inklusif dan berbasis hak. Secara bersama, semua ini menunjukkan bahwa masa depan AI tidak lagi hanya ditentukan oleh laboratorium dan pasar, tetapi oleh kualitas institusi global dan nasional yang mengawasinya. ([UNESCO](#))

Bagi negara berkembang, termasuk Indonesia, pelajaran terpentingnya ialah bahwa etika AI tidak boleh ditunda sampai "teknologi matang." Justru semakin cepat AI menyusup ke keputusan organisasi, semakin mendesak kebutuhan akan kebijakan, dokumentasi, audit, human oversight, dan akses terhadap remedy. Organisasi yang mengabaikan etika mungkin tampak bergerak lebih cepat, tetapi dalam jangka panjang akan menghadapi erosi kepercayaan, risiko hukum, dan kerusakan reputasi.

Dengan demikian, pertanyaan tentang etika AI bukan apakah kita pro-atau anti-teknologi. Pertanyaan yang lebih tepat ialah: dapatkah kita membangun tata kelola yang membuat AI tetap berada dalam horizon keadilan, martabat, dan tanggung jawab manusia? Jawaban atas

pertanyaan itu akan menentukan apakah AI menjadi alat emansipasi atau mesin reproduksi ketimpangan dalam abad ke-21.

## **Glosarium**

**Akuntabilitas AI:** Kejelasan mengenai siapa yang bertanggung jawab atas desain, penggunaan, pengawasan, dampak, dan perbaikan sistem AI.

**AI governance:** Struktur kebijakan, proses, peran, dan kontrol untuk memastikan AI digunakan secara aman, etis, sah, dan bertanggung jawab.

**AI literacy:** Kemampuan memahami peluang, keterbatasan, risiko, dan penggunaan AI secara memadai dalam konteks kerja atau kehidupan publik.

**Bias algoritmik:** Kecenderungan sistem AI menghasilkan keputusan atau keluaran yang merugikan kelompok tertentu secara tidak proporsional.

**Due diligence:** Proses kelembagaan untuk mengidentifikasi, mencegah, memitigasi, dan melaporkan risiko serta dampak negatif.

**Fairness:** Prinsip keadilan dalam desain dan penggunaan AI, baik dari sisi prosedur, distribusi hasil, maupun perlakuan terhadap kelompok berbeda.

**General-purpose AI (GPAI):** Model AI serbaguna yang dapat digunakan untuk banyak tugas dan menjadi fondasi bagi banyak aplikasi turunan.

**Human oversight:** Keterlibatan manusia dalam memantau, menilai, mengoreksi, atau menghentikan sistem AI.

**Impact assessment:** Penilaian sistematis terhadap dampak teknologi terhadap individu, kelompok, organisasi, dan masyarakat.

**Interoperabilitas kebijakan:** Kemampuan berbagai rezim regulasi dan standar untuk saling selaras atau bekerja bersama tanpa harus identik.

**Lifecycle AI:** Seluruh tahapan sistem AI, dari desain, pengembangan, pengujian, deployment, penggunaan, pemantauan, hingga penghentian.

**Provenance:** Informasi tentang asal-usul dan riwayat data atau konten, penting untuk integritas informasi dan penelusuran sumber.

**Risk-based approach:** Pendekatan yang menyesuaikan intensitas kewajiban tata kelola dengan tingkat risiko sistem AI.

**Synthetic content:** Konten yang dihasilkan atau dimodifikasi oleh AI, termasuk teks, gambar, audio, atau video.

**Trustworthy AI:** AI yang memenuhi karakteristik seperti valid, aman, adil, transparan, dapat dijelaskan, menghormati privasi, dan akuntabel.

## Daftar Pustaka (APA 7)

Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and machine learning: Limitations and opportunities*. MIT Press.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 149–159.

Council of Europe. (2024). *Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law*. <https://www.coe.int>

European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*. <https://eur-lex.europa.eu>

Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society:

Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707.

International Organization for Standardization. (2023). *ISO/IEC 23894:2023 Artificial intelligence—Guidance on risk management*. ISO.

International Organization for Standardization. (2023). *ISO/IEC 42001:2023 Artificial intelligence—Management system*. ISO.

International Organization for Standardization. (2025). *ISO/IEC 42005:2025 Information technology—Artificial intelligence (AI)—AI system impact assessment*. ISO.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21.

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce.

National Institute of Standards and Technology. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1). U.S. Department of Commerce.

OECD. (2019/2024). *Recommendation of the Council on Artificial Intelligence*. OECD Legal Instruments. <https://legalinstruments.oecd.org>

OECD. (2024). *OECD updates AI Principles to stay abreast of rapid technological developments*. <https://www.oecd.org>

Office of the United Nations High Commissioner for Human Rights. (2024). *Human rights due diligence for digital technology use*. OHCHR.

O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68.

UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. UNESCO.

United Nations. (2024). *Global Digital Compact*. United Nations.

Copilot for this article: ChatGPT 5.4. Thinking (2026). Access date: 28 March 2026. Author's account ([Rudy C Tarumingkeng](#))

<https://chatgpt.com/c/69c614f3-0004-83a0-8ba5-1ecf32dccd07>