



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

Oleh: Rudy C Tarumingkeng

*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

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)

Ketua Dewan Guru Besar, IPB-University, Bogor (2005-2006)

AI - Data Analyst, dan Ketua Senat Akademik, IBM-ASMI, Jakarta 2024-

---

© RudyCT Academic Series

[rudyct75@gmail.com](mailto:rudyct75@gmail.com)

6 Maret 2026

Berikut esai akademik yang membahas “**Etika AI dalam Perspektif Global: Bias, Akuntabilitas, dan Standar Internasional.**”

---

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

### **1. Pendahuluan**

Kecerdasan buatan atau *artificial intelligence* (AI) telah bergerak dari wilayah laboratorium dan riset terbatas ke ruang kehidupan sehari-hari manusia. AI kini digunakan dalam pencarian informasi, sistem rekomendasi, layanan pelanggan, diagnosis medis, penilaian kredit, rekrutmen tenaga kerja, pengawasan keamanan, analisis perilaku konsumen, hingga pengambilan keputusan kebijakan publik. Perubahan ini membawa janji besar: efisiensi, kecepatan, personalisasi, dan kemampuan analitik yang jauh melampaui kapasitas manusia dalam mengolah data berskala besar. Namun, semakin luas AI digunakan, semakin jelas pula bahwa persoalan AI bukan lagi semata-mata persoalan teknis, melainkan persoalan etis, sosial, hukum, dan politik global. ([UNESCO](#))

Dalam konteks ini, diskusi mengenai etika AI menjadi sangat penting. AI bukan sekadar alat netral. Ia dibangun dari data, dirancang oleh manusia, diterapkan dalam organisasi, dan beroperasi di dalam struktur sosial yang sering kali sudah mengandung ketimpangan. Karena itu, AI dapat mempercepat manfaat, tetapi sekaligus memperbesar ketidakadilan yang sudah ada. UNESCO menegaskan bahwa risiko AI mencakup kemungkinan tertanamnya bias, ancaman terhadap hak asasi manusia, dan penguatan ketimpangan yang sudah terlebih dahulu dialami kelompok-kelompok rentan. ([UNESCO](#))

Dari sekian banyak isu etika AI, tiga tema sangat menonjol dalam perdebatan global. Pertama, **bias**, yaitu kecenderungan sistem AI menghasilkan keluaran yang tidak adil, diskriminatif, atau menyimpang karena masalah pada data, model, tujuan desain, atau konteks penggunaannya. Kedua, **akuntabilitas**, yaitu pertanyaan tentang siapa yang bertanggung jawab ketika AI membuat keputusan yang merugikan, salah, atau tidak dapat dijelaskan. Ketiga, **standar internasional**, yaitu usaha berbagai organisasi global dan regional untuk menciptakan prinsip, kerangka, dan instrumen yang dapat menjadi acuan bersama dalam merancang, menerapkan, dan mengawasi AI secara bertanggung jawab. ([UNESCO](#))

Secara global, perkembangan tata kelola AI menunjukkan pergeseran dari sekadar seruan normatif menuju instrumen yang lebih operasional. OECD memiliki prinsip AI yang pertama kali diadopsi pada 2019 dan diperbarui pada 2024. UNESCO mengeluarkan Rekomendasi Etika AI yang berlaku untuk seluruh 194 negara anggotanya. NIST di Amerika Serikat mengembangkan *AI Risk Management Framework* untuk mendukung pengelolaan risiko secara praktis. ISO/IEC menerbitkan standar seperti ISO/IEC 42001 untuk sistem manajemen AI dan ISO/IEC 23894 untuk manajemen risiko AI. Di tingkat regulasi, Uni Eropa telah menetapkan AI Act sebagai kerangka hukum berbasis risiko, sementara Council of Europe membuka untuk penandatanganan konvensi

internasional yang disebut sebagai traktat yang mengikat secara hukum pertama di bidang ini. ([OECD](#))

Dengan demikian, etika AI dalam perspektif global tidak dapat dipahami hanya sebagai daftar nilai abstrak seperti keadilan, transparansi, atau kemanusiaan. Ia harus dilihat sebagai medan perjumpaan antara teknologi, pasar, negara, hukum, organisasi, dan masyarakat sipil. Di sinilah AI menjadi isu strategis. Sebuah perusahaan dapat mengklaim bahwa sistemnya akurat secara teknis, tetapi pertanyaan etis tetap muncul: akurat bagi siapa, dengan data siapa, untuk tujuan apa, dan dengan risiko sosial seperti apa? Demikian pula, sebuah negara dapat mendorong inovasi AI atas nama daya saing ekonomi, tetapi pertanyaan moralnya tetap ada: apakah percepatan inovasi itu memperkuat martabat manusia atau justru mengubah manusia menjadi sekadar objek kalkulasi algoritmik? ([UNESCO](#))

Esai ini bertujuan menjelaskan etika AI dalam perspektif global dengan menyoroti tiga poros utama: bias, akuntabilitas, dan standar internasional. Pembahasan akan dimulai dari dasar filosofis dan konseptual etika AI, lalu beralih ke jenis-jenis bias dan manifestasinya dalam praktik, dilanjutkan dengan isu akuntabilitas dalam rantai keputusan algoritmik, dan akhirnya menelaah lanskap standar internasional yang sedang berkembang. Pendekatan ini penting agar pembaca tidak melihat etika AI hanya sebagai topik moral yang idealistis, melainkan sebagai kerangka nyata untuk tata kelola teknologi yang adil, bertanggung jawab, dan berorientasi pada kepentingan kemanusiaan global. ([OECD](#))

## **2. Landasan Konseptual Etika AI**

Etika AI berakar pada pertanyaan klasik filsafat moral: apa yang baik, apa yang adil, siapa yang bertanggung jawab, dan bagaimana teknologi seharusnya melayani manusia. Namun, AI menambahkan kompleksitas baru karena keputusan yang sebelumnya dibuat manusia kini sebagian

didelegasikan kepada sistem berbasis data dan model statistik. Dalam praktiknya, AI sering tidak "berpikir" seperti manusia, melainkan mengenali pola dari data, menghasilkan prediksi, lalu memengaruhi keputusan. Hal ini membuat etika AI menjadi wilayah lintas disiplin yang menggabungkan filsafat, ilmu komputer, hukum, kebijakan publik, sosiologi, ekonomi, dan manajemen. ([UNESCO](#))

Salah satu pergeseran penting dalam etika teknologi kontemporer adalah berpindahnya fokus dari pertanyaan "apakah teknologi ini bisa dibuat?" menjadi "apakah teknologi ini layak dibuat dan bagaimana menggunakannya secara benar?" Dalam konteks AI, pertanyaan ini menjadi mendesak karena skala dan dampak sistemnya sangat besar. Sistem rekomendasi dapat memengaruhi opini publik. Sistem pemeringkatan kandidat kerja dapat membentuk peluang hidup seseorang. Sistem pemantauan dapat mengubah relasi antara negara dan warga. Artinya, AI tidak hanya menyelesaikan masalah; ia juga membentuk struktur kekuasaan baru. ([Digital Strategy](#))

Kerangka global yang menonjol umumnya menempatkan etika AI di atas landasan hak asasi manusia, keadilan, keamanan, transparansi, akuntabilitas, dan pengawasan manusia. UNESCO secara eksplisit menyebut perlindungan hak asasi manusia dan martabat manusia sebagai batu penjuruk rekomendasinya, sambil menekankan transparansi, fairness, auditabilitas, keterlacakan, dan pengawasan manusia. OECD juga menempatkan penghormatan terhadap hak asasi manusia dan nilai-nilai demokratis sebagai inti prinsip AI yang dapat dipercaya. ([UNESCO](#))

Di titik ini, tampak bahwa etika AI bukan sekadar etika "tentang mesin", melainkan etika tentang relasi antara manusia, organisasi, dan kekuasaan yang dimediasi oleh mesin. Sebuah sistem AI mungkin canggih secara komputasional, tetapi tetap tidak etis bila ia memperlakukan kelompok tertentu secara tidak adil. Sebaliknya, sistem yang tidak sepenuhnya sempurna secara teknis masih dapat diterima bila ia dirancang dengan

mitigasi risiko, pengawasan manusia, mekanisme banding, dan transparansi yang memadai. Dengan kata lain, etika AI menuntut keseimbangan antara inovasi dan perlindungan. ISO/IEC 42001 sendiri menekankan bahwa tata kelola AI harus menyeimbangkan risiko dan peluang, sekaligus mendukung penggunaan AI yang bertanggung jawab. ([ISO](#))

Dalam diskursus global, ada kecenderungan untuk memandang AI sebagai sistem sosio-teknis. Ini berarti kegagalan AI tidak selalu berasal dari model matematisnya saja. Kegagalan bisa timbul dari data yang tidak representatif, tujuan bisnis yang sempit, kebijakan organisasi yang lemah, kurangnya literasi pengguna, atau absennya mekanisme audit. Karena itu, pembahasan etika AI harus melampaui kode dan algoritma. Ia harus mencakup siapa yang mengembangkan sistem, siapa yang terdampak, bagaimana keputusan diuji, dan bagaimana kerugian diperbaiki. NIST menegaskan bahwa pengelolaan risiko AI harus mempertimbangkan individu, organisasi, dan masyarakat, bukan hanya kinerja teknis model. ([NIST](#))

### **3. Bias dalam AI: Pengertian dan Akar Permasalahan**

Bias dalam AI dapat dipahami sebagai kecenderungan sistem menghasilkan keluaran yang secara sistematis menguntungkan atau merugikan kelompok tertentu secara tidak adil. Dalam bahasa yang lebih sederhana, bias terjadi ketika AI tidak memperlakukan orang atau situasi secara proporsional, rasional, dan setara menurut tujuan yang sah. Bias ini sering dianggap sebagai "masalah data", padahal sesungguhnya ia jauh lebih kompleks. Bias dapat muncul sejak tahap penentuan tujuan sistem, pemilihan variabel, desain model, strategi evaluasi, hingga implementasi di lapangan. ([UNESCO](#))

Sumber pertama bias adalah **bias historis**. Data masa lalu mencerminkan sejarah sosial, termasuk diskriminasi, ketimpangan, dan pola eksklusi. Bila AI dilatih dari data tersebut tanpa koreksi, sistem akan

“mempelajari” sejarah ketidakadilan itu sebagai pola yang dianggap normal. Misalnya, bila data perekrutan selama bertahun-tahun lebih banyak memilih kandidat dari kelompok tertentu, maka model seleksi otomatis bisa menyimpulkan bahwa ciri-ciri kelompok itulah yang identik dengan “kandidat baik”, padahal itu mungkin hanyalah pantulan bias institusional masa lalu. UNESCO mengingatkan bahwa AI dapat mengompound ketimpangan yang sudah ada dan semakin merugikan kelompok yang termarginalkan. ([UNESCO](#))

Sumber kedua adalah **bias representasi**. Data yang digunakan sering kali tidak mewakili keberagaman populasi sebenarnya. Kelompok minoritas bisa kurang terwakili, data dari negara berkembang bisa lebih sedikit dibanding negara maju, bahasa-bahasa dominan lebih banyak dilatih daripada bahasa lokal, dan konteks budaya tertentu kurang tercermin dalam set data global. Akibatnya, model tampak bekerja baik secara umum, tetapi gagal pada kelompok atau konteks tertentu. Ini menjadi persoalan serius dalam AI global, sebab standar performa yang dibangun dari lingkungan sosial tertentu dapat dipaksakan ke masyarakat lain yang sangat berbeda. ([UNESCO](#))

Sumber ketiga adalah **bias pengukuran**. Dalam banyak kasus, apa yang ingin diukur sesungguhnya rumit, tetapi diubah menjadi indikator sederhana agar dapat dimasukkan ke model. Kualitas calon pegawai, misalnya, dapat direduksi menjadi skor tes tertentu; risiko kredit direduksi menjadi sejarah pembayaran; kualitas siswa direduksi menjadi performa ujian standar. Reduksi semacam ini menimbulkan persoalan etis karena realitas sosial yang kompleks dipadatkan menjadi angka yang belum tentu adil. Ketika proksi yang dipilih tidak tepat, AI dapat memperkuat keputusan yang secara statistik tampak rapi tetapi secara moral cacat. ([Digital Strategy](#))

Sumber keempat adalah **bias tujuan optimalisasi**. Sistem AI sering dioptimalkan untuk efisiensi, akurasi, klik, konversi, atau keuntungan.

Tetapi apa yang optimal secara bisnis belum tentu adil secara sosial. Mesin rekomendasi yang dioptimalkan untuk durasi keterlibatan, misalnya, bisa terdorong menyajikan konten sensasional atau polarisatif. Sistem moderasi otomatis yang dioptimalkan untuk kecepatan bisa lebih sering salah menandai bahasa kelompok tertentu. Pada titik ini, bias bukan hanya soal data buruk, tetapi juga soal tujuan organisasi yang terlalu sempit. ([NIST](#))

Sumber kelima adalah **bias implementasi dan penggunaan**. Sebuah model yang dirancang baik dapat tetap menimbulkan ketidakadilan jika digunakan di lingkungan yang salah atau tanpa pelatihan dan pengawasan yang cukup. Misalnya, sistem penilaian risiko mungkin dimaksudkan sebagai alat bantu, tetapi di lapangan dipakai sebagai penentu akhir tanpa ruang pertimbangan manusia. NIST menekankan pentingnya memahami konteks penggunaan dan penggunaan kerangka berbasis risiko yang berjalan sepanjang siklus hidup AI. ([NIST Publications](#))

Dari sini jelas bahwa bias dalam AI harus dilihat sebagai masalah sistemik. Ia tidak dapat diatasi hanya dengan menambah data atau mengganti algoritma. Diperlukan refleksi atas nilai, struktur insentif, cara organisasi bekerja, serta siapa yang dilibatkan dalam desain dan evaluasi. Dengan kata lain, bias adalah cermin hubungan antara teknologi dan masyarakat. Sistem AI belajar dari dunia yang tidak sempurna, dan tanpa intervensi etis yang sungguh-sungguh, ia akan mengulang bahkan memperkeras ketidaksempurnaan itu. ([UNESCO](#))

#### **4. Manifestasi Bias dalam Berbagai Sektor**

Dalam sektor ketenagakerjaan, bias AI dapat terlihat pada sistem penyaringan CV, penilaian kandidat, atau analitik produktivitas. Narasinya dapat dibayangkan demikian. Sebuah perusahaan multinasional ingin mempercepat rekrutmen dengan AI. Sistem dilatih dari data karyawan yang dianggap "berprestasi tinggi" selama sepuluh

tahun. Namun, data itu ternyata didominasi oleh latar belakang pendidikan, bahasa, dan pola karier tertentu. Akibatnya, AI memberi skor lebih rendah kepada kandidat yang kompeten tetapi berasal dari jalur pendidikan alternatif, wilayah nonmetropolitan, atau kelompok yang historically kurang masuk ke perusahaan tersebut. Secara kasat mata sistem tampak objektif, padahal ia hanya mengulangi sejarah preferensi institusi. Uni Eropa bahkan secara eksplisit menjadikan penggunaan AI dalam konteks kerja dan keputusan yang memengaruhi hak-hak individu sebagai area yang perlu diatur secara ketat dalam pendekatan berbasis risiko. ([Digital Strategy](#))

Dalam sektor keuangan, bias AI dapat muncul dalam penilaian kredit, deteksi fraud, atau penetapan harga asuransi. Sering kali model memanfaatkan variabel yang tampak netral, tetapi sebenarnya berkorelasi dengan status sosial-ekonomi, lokasi geografis, atau kelompok tertentu. Seseorang dapat dinilai lebih berisiko bukan karena perilaku individualnya, melainkan karena pola statistik kelompok tempat ia diasosiasikan. Di sini masalah etika muncul karena AI berpotensi menggantikan penilaian individual dengan kategorisasi probabilistik yang mengunci orang dalam identitas risiko yang tidak sepenuhnya mereka ciptakan sendiri. UNESCO menekankan bahwa fairness dan non-discrimination harus menjadi prinsip utama, bukan sekadar pertimbangan tambahan. ([UNESCO](#))

Dalam bidang kesehatan, persoalannya lebih rumit lagi. AI dapat membantu diagnosis, triase, prediksi penyakit, dan alokasi sumber daya. Tetapi jika data pelatihan lebih banyak berasal dari populasi tertentu, model bisa tampil unggul pada kelompok mayoritas dan menurun akurasi pada kelompok lain. Dalam konteks kesehatan global, ketimpangan data antara negara kaya dan negara berkembang menjadi isu besar. Sebuah model yang dilatih dari rumah sakit besar di negara maju belum tentu adil untuk pasien di daerah dengan struktur penyakit, kondisi sosial, atau kualitas data yang berbeda. Karena itu, persoalan

bias dalam AI kesehatan bukan hanya teknis, melainkan juga persoalan keadilan global. ([UNESCO](#))

Dalam pendidikan, AI digunakan untuk pembelajaran adaptif, proctoring, penilaian, dan prediksi keberhasilan siswa. Narasi praktisnya bisa seperti ini: sebuah universitas mengadopsi AI untuk mendeteksi risiko putus studi. Model dilatih dari data kehadiran, nilai, interaksi platform, dan latar belakang ekonomi. Sistem kemudian memberi label “berisiko tinggi” pada mahasiswa yang jarang aktif di sistem digital. Namun, label itu ternyata lebih sering jatuh pada mahasiswa dari daerah dengan koneksi internet lemah atau yang harus bekerja sambil kuliah. Dalam situasi ini, AI bukan sekadar memprediksi risiko; ia berpotensi memberi stigma pada kelompok yang justru membutuhkan dukungan struktural. ([UNESCO](#))

Dalam penegakan hukum dan keamanan, bahaya bias menjadi lebih tajam karena konsekuensinya langsung menyentuh kebebasan, hak, dan martabat manusia. Sistem pengenalan wajah, prediksi risiko kriminal, atau analisis perilaku dapat menghasilkan salah identifikasi atau penandaan yang tidak proporsional terhadap kelompok tertentu. Council of Europe menempatkan risiko terhadap hak asasi manusia, demokrasi, dan rule of law sebagai perhatian utama, bahkan mendorong risk and impact assessment yang iteratif serta membuka kemungkinan pelarangan atau moratorium untuk aplikasi tertentu. Uni Eropa juga melarang beberapa praktik yang dianggap berisiko tak dapat diterima, termasuk social scoring, prediksi risiko kriminal individual semata dari data biometrik, dan pengenalan emosi di tempat kerja serta institusi pendidikan. ([Portal](#))

Dalam ekosistem media digital, bias AI beroperasi lewat rekomendasi konten, peringkat informasi, moderasi otomatis, dan personalisasi iklan. Bias di sini tidak selalu berupa diskriminasi langsung, tetapi bisa hadir dalam bentuk amplifikasi wacana tertentu, invisibilitas kelompok

tertentu, atau penciptaan ruang gema informasi. Sebuah algoritma mungkin tidak “membenci” siapa pun, tetapi jika ia terus mengutamakan konten yang paling memancing interaksi, maka ia bisa berkontribusi pada polarisasi, misinformasi, dan penurunan kualitas ruang publik. Karena itu, etika AI dalam media digital berkaitan erat dengan demokrasi, bukan hanya dengan kenyamanan pengguna. OECD dan Council of Europe sama-sama menekankan hubungan AI dengan nilai-nilai demokratis. ([OECD](#))

## **5. Mengapa Bias Sulit Dihilangkan Sepenuhnya**

Bias dalam AI sering diperlakukan seolah-olah bisa diselesaikan dengan resep teknis sederhana. Kenyataannya, bias sulit dihilangkan sepenuhnya karena konsep “adil” sendiri bersifat kontekstual. Dalam statistika dan machine learning, ada beragam metrik fairness yang tidak selalu dapat dipenuhi sekaligus. Sebuah sistem dapat setara dari segi tingkat kesalahan agregat, tetapi tidak setara dari segi probabilitas positif palsu pada subkelompok tertentu. Di sini etika dan kebijakan harus melengkapi matematika. Pertanyaannya bukan hanya “berapa akurasi model,” tetapi “definisi keadilan mana yang layak diprioritaskan dalam konteks sosial tertentu?” ([NIST AI Resource Center](#))

Selain itu, AI bekerja dalam dunia yang dinamis. Data berubah, perilaku pengguna berubah, konteks hukum berubah, dan dampak sistem dapat muncul setelah implementasi. Karena itu, fairness bukan keadaan final, melainkan proses pengawasan berkelanjutan. NIST menekankan bahwa manajemen risiko AI harus berjalan terus-menerus sepanjang siklus hidup sistem, dan fungsi pengukuran serta pengelolaan harus diperbarui seiring perubahan risiko. ([NIST Publications](#))

Ada pula persoalan politik dan ekonomi. Organisasi mungkin mengetahui adanya bias, tetapi memilih toleransi terhadap bias tertentu demi efisiensi biaya, kecepatan rilis, atau posisi kompetitif. Dalam konteks ini, bias menjadi gejala tata kelola yang lemah. Ia menunjukkan

bahwa organisasi belum memiliki struktur keputusan yang menempatkan kepentingan pengguna dan masyarakat setara dengan target bisnis. Di sinilah etika AI bersentuhan dengan governance perusahaan. ISO/IEC 42001 menekankan perlunya kebijakan, proses, dan perbaikan berkelanjutan pada tingkat organisasi, bukan hanya pada level model. ([ISO](#))

## 6. Akuntabilitas dalam AI: Siapa Bertanggung Jawab?

Jika bias berkaitan dengan keluaran yang tidak adil, maka akuntabilitas berkaitan dengan pertanyaan yang segera menyusul: **siapa yang bertanggung jawab?** Ini merupakan salah satu isu paling sentral dalam etika AI. Dalam sistem tradisional, tanggung jawab biasanya lebih mudah dilacak. Namun, dalam AI, keputusan lahir dari interaksi kompleks antara pembuat data, pengembang model, penyedia infrastruktur, organisasi pengguna, operator manusia, dan pengambil kebijakan. Ketika sebuah sistem salah memprediksi, mendiskriminasi, atau menyebabkan kerugian, tanggung jawab sering terpecah dan menjadi kabur. ([UNESCO](#))

UNESCO secara tegas menyatakan bahwa sistem AI tidak boleh menggantikan tanggung jawab dan akuntabilitas manusia yang bersifat final. Ini sangat penting. AI dapat membantu, merekomendasikan, atau mengotomatisasi sebagian proses, tetapi ia tidak boleh dijadikan tameng untuk menghindari tanggung jawab moral dan hukum. Di belakang setiap sistem selalu ada aktor manusia dan kelembagaan yang merancang tujuan, menentukan ambang toleransi risiko, dan memutuskan penerapan. ([UNESCO](#))

Akuntabilitas memiliki beberapa dimensi. Pertama, **akuntabilitas desain**. Pengembang dan organisasi harus dapat menjelaskan mengapa sistem dibangun, untuk tujuan apa, menggunakan data apa, dengan asumsi apa, dan dengan mitigasi risiko seperti apa. Kedua, **akuntabilitas operasional**. Organisasi pengguna harus memastikan AI dipakai sesuai tujuan, dengan pelatihan memadai, pengawasan manusia, dan prosedur

intervensi jika terjadi anomali. Ketiga, **akuntabilitas remedial**. Bila AI merugikan seseorang, harus ada mekanisme pengaduan, koreksi, banding, dan perbaikan. Tanpa tiga lapis ini, klaim "AI kami bertanggung jawab" hanya menjadi slogan. ([UNESCO](#))

Masalah besar lainnya adalah **black box accountability**. Banyak model AI, terutama yang kompleks, sulit dijelaskan secara intuitif. Ini menimbulkan ketegangan antara performa dan keterjelasan. UNESCO mengakui bahwa transparansi dan explainability penting, tetapi tingkatnya perlu disesuaikan dengan konteks karena dapat berbenturan dengan keamanan atau privasi. Artinya, penjelasan tidak harus selalu berupa pemaparan seluruh kode model, tetapi harus cukup untuk memungkinkan audit, evaluasi, dan pembelaan hak pihak yang terdampak. ([UNESCO](#))

Dalam konteks organisasi, akuntabilitas AI menuntut pembagian peran yang jelas. NIST Playbook menekankan perlunya definisi tanggung jawab manusia yang eksplisit, pelacakan informasi risiko, dan penyelarasan kebijakan AI dengan hukum, tata kelola data, serta toleransi risiko organisasi. Ini berarti organisasi yang matang tidak boleh melempar tanggung jawab ke tim data science semata. Fungsi hukum, etika, audit internal, keamanan, SDM, dan manajemen puncak harus terlibat. ([NIST AI Resource Center](#))

Mari kita lihat narasi kasus sederhana. Sebuah bank menggunakan AI untuk menilai kelayakan pinjaman UMKM. Sistem ternyata banyak menolak pemohon dari wilayah tertentu. Tim teknis berkata bahwa model hanya mengikuti pola data. Tim bisnis berkata keputusan dibuat otomatis sesuai kebijakan efisiensi. Vendor perangkat lunak berkata mereka hanya menyediakan platform. Manajemen puncak berkata mereka tidak memahami detail model. Dalam kasus semacam ini, korban diskriminasi sulit menemukan pintu pertanggungjawaban. Di sinilah pentingnya kerangka akuntabilitas end-to-end: sejak pengadaan

teknologi, kontrak vendor, dokumentasi model, validasi dampak, hingga hak pelanggan untuk meminta penjelasan dan peninjauan ulang. ([ISO](#))

## 7. Komponen-Komponen Akuntabilitas yang Baik

Akuntabilitas AI yang baik setidaknya memerlukan tujuh komponen. Pertama, **governance structure** yang jelas. Harus ada penanggung jawab kebijakan AI pada tingkat pimpinan, bukan sekadar tim teknis. Kedua, **documentation and traceability**. Sistem harus terdokumentasi: data asal, proses pelatihan, asumsi model, hasil uji, dan perubahan versi. UNESCO menekankan auditabilitas dan traceability sebagai syarat penting. ([UNESCO](#))

Ketiga, **risk and impact assessment** sebelum dan sesudah implementasi. Council of Europe menekankan penilaian risiko dan dampak yang bersifat iteratif terhadap hak asasi manusia, demokrasi, dan rule of law. Ini menunjukkan bahwa akuntabilitas tidak berhenti saat sistem diluncurkan; ia terus berjalan selama sistem beroperasi. ([Portal](#))

Keempat, **human oversight** yang bermakna. Pengawasan manusia tidak boleh sekadar formalitas. Operator harus memahami kapan perlu mengintervensi, kapan perlu menghentikan sistem, dan kapan hasil AI tidak layak dipercaya. UNESCO secara eksplisit menggarisbawahi pentingnya human oversight dan menolak penggantian tanggung jawab manusia oleh AI. ([UNESCO](#))

Kelima, **audit independen dan evaluasi berkala**. Sistem yang menyentuh hak atau kesempatan hidup orang tidak cukup diuji sekali saat awal. Perlu evaluasi internal dan, untuk kasus berisiko tinggi, audit eksternal. NIST menekankan penggunaan metrik, pengukuran risiko, dan dokumentasi atas aspek trustworthiness yang dapat maupun tidak dapat diukur. ([NIST AI Resource Center](#))

Keenam, **mekanisme pengaduan dan banding** bagi pihak terdampak. Akuntabilitas menjadi kosong bila individu tidak memiliki jalan untuk

menantang keputusan AI. Ketujuh, **budaya organisasi yang etis**. Ini sering diabaikan. Teknologi terbaik pun dapat dipakai secara sembrono dalam budaya kerja yang hanya menghargai kecepatan dan profit. Di sinilah etika AI bersentuhan dengan kepemimpinan, integritas, dan tata kelola perusahaan. ISO/IEC 42001 dirancang justru untuk membangun sistem manajemen yang terus diperbaiki, bukan sekadar kepatuhan sesaat. ([ISO](#))

## **8. Standar dan Kerangka Internasional: Dari Prinsip ke Implementasi**

Lanskap standar internasional AI saat ini menunjukkan pola yang menarik. Di satu sisi, ada dokumen prinsipil bernuansa nilai. Di sisi lain, ada kerangka operasional untuk manajemen risiko dan standar sistem manajemen. Lalu, muncul pula aturan hukum dan traktat internasional. Kombinasi ini menandai pendewasaan governance AI global. Dulu diskusi etika AI banyak berhenti pada deklarasi normatif. Kini diskusi bergerak menuju audit, pelaporan, interoperabilitas, dan kewajiban kelembagaan. ([OECD](#))

### **a. OECD AI Principles**

OECD AI Principles merupakan standar antarpemerintah pertama di bidang ini. Prinsip ini diadopsi pada 2019 dan diperbarui pada 2024, terdiri atas lima prinsip berbasis nilai dan lima rekomendasi bagi pembuat kebijakan serta aktor AI. Relevansinya terletak pada posisinya sebagai jembatan antara inovasi dan kepercayaan. OECD menekankan AI yang inovatif dan dapat dipercaya sambil menghormati hak asasi manusia dan nilai-nilai demokratis. Pembaruan 2024 menunjukkan bahwa prinsip ini berusaha tetap relevan terhadap perkembangan baru, termasuk general-purpose dan generative AI. ([OECD](#))

Bagi dunia manajemen, OECD penting karena memberi bahasa bersama untuk kebijakan lintas negara. Prinsipnya fleksibel, tidak terlalu teknis, dan cocok sebagai acuan awal organisasi yang ingin membangun

kebijakan AI internal. Kelemahannya, seperti banyak prinsip global lain, adalah bahwa ia tidak otomatis menjamin implementasi. Namun, pada 2025 OECD meluncurkan kerangka pelaporan global untuk memonitor penerapan Hiroshima Code of Conduct, dengan tujuan meningkatkan keterbandingan, transparansi, dan interoperabilitas antar mekanisme governance AI. Ini menunjukkan pergeseran dari prinsip ke pelaporan praktik nyata. ([OECD](#))

### **b. UNESCO Recommendation on the Ethics of AI**

Rekomendasi UNESCO memiliki bobot moral-politik yang sangat besar karena berlaku untuk semua 194 negara anggota UNESCO. Dokumen ini menempatkan human rights and human dignity sebagai fondasi, disertai prinsip seperti fairness, transparency, explainability, auditability, oversight, sustainability, dan non-discrimination. Yang menarik, UNESCO tidak berhenti pada nilai abstrak, tetapi juga mengembangkan metodologi implementasi seperti *Readiness Assessment Methodology* dan *Ethical Impact Assessment*. ([UNESCO](#))

Kekuatan UNESCO terletak pada pendekatan yang lebih luas daripada sekadar risiko teknis. UNESCO memasukkan lingkungan, ekosistem, gender, pendidikan, penelitian, dan kesejahteraan sosial. Ini penting untuk negara berkembang, karena etika AI tidak bisa hanya diukur dari kepatuhan teknis perusahaan besar, tetapi juga dari dampaknya terhadap pembangunan manusia dan keadilan sosial. Bagi perspektif global Selatan, UNESCO menawarkan pendekatan yang lebih inklusif dan lebih sensitif terhadap ketimpangan struktural. ([UNESCO](#))

### **c. NIST AI Risk Management Framework**

NIST AI RMF menonjol karena sifatnya praktis. NIST menyatakan bahwa kerangka ini dimaksudkan untuk penggunaan sukarela guna membantu organisasi mengelola risiko terhadap individu, organisasi, dan masyarakat. AI RMF berfokus pada trustworthiness dan dioperasionalkan

melalui empat fungsi: Govern, Map, Measure, dan Manage. Playbook NIST menambahkan langkah-langkah praktis untuk masing-masing fungsi. ([NIST](#))

Bagi organisasi, NIST sangat berguna karena membantu menerjemahkan ideal etis menjadi proses kerja. *Govern* berkaitan dengan kebijakan dan struktur tanggung jawab. *Map* menuntut organisasi memahami konteks, tujuan, aktor, dan potensi dampak. *Measure* mengharuskan identifikasi metrik dan evaluasi risiko. *Manage* mendorong penanganan dan pemantauan risiko secara berkelanjutan. Kelebihannya adalah operasional. Kekurangannya, karena bersifat sukarela, keberhasilannya sangat bergantung pada komitmen organisasi. Namun, justru di situlah nilai etisnya: organisasi ditantang membangun kedewasaan governance, bukan sekadar mematuhi aturan minimum. ([NIST Publications](#))

#### **d. ISO/IEC 42001 dan ISO/IEC 23894**

ISO/IEC 42001:2023 adalah standar internasional pertama untuk sistem manajemen AI. Standar ini menetapkan persyaratan untuk membangun, menerapkan, memelihara, dan terus meningkatkan *Artificial Intelligence Management System* dalam organisasi. ISO menekankan bahwa standar ini membantu organisasi mengelola risiko dan peluang AI secara terstruktur, mendukung transparansi, keterlacakan, reliabilitas, dan kepatuhan. ([ISO](#))

Sementara itu, ISO/IEC 23894:2023 memberi panduan khusus tentang manajemen risiko AI dan integrasinya ke dalam aktivitas organisasi. Kedua standar ini penting karena memindahkan etika AI dari ruang wacana ke ruang sistem manajemen. Bagi perusahaan, universitas, rumah sakit, bank, atau lembaga publik, pertanyaan etika AI tidak lagi cukup dijawab dengan slogan “kami peduli fairness,” tetapi harus masuk ke kebijakan, audit, proses pengadaan, evaluasi vendor, dan perbaikan berkelanjutan. ([ISO](#))

### **e. EU AI Act**

EU AI Act adalah kerangka hukum komprehensif pertama di dunia untuk AI. Komisi Eropa menyebutnya sebagai *the first-ever legal framework on AI* dan menegaskan bahwa AI Act menggunakan pendekatan berbasis risiko. Regulasi ini membedakan tingkat risiko dan menetapkan larangan untuk beberapa praktik yang dianggap sebagai ancaman jelas bagi keselamatan, mata pencaharian, dan hak-hak manusia, termasuk social scoring, eksploitasi kerentanan tertentu, beberapa bentuk biometric categorisation, emotion recognition di tempat kerja dan pendidikan, serta beberapa bentuk prediksi risiko kriminal individual. ([Digital Strategy](#))

Secara global, arti penting AI Act terletak pada efek ekstrateritorial dan daya teladannya. Banyak organisasi di luar Eropa akan menyesuaikan diri karena beroperasi di pasar Eropa atau karena regulasi ini menjadi acuan global. Walaupun AI Act adalah hukum regional, pengaruhnya bersifat internasional. Ia menunjukkan bahwa governance AI kini telah bergerak dari norma sukarela menuju kewajiban hukum yang konkret. ([Digital Strategy](#))

### **f. Council of Europe Framework Convention**

Council of Europe membuka untuk penandatanganan *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law* pada 5 September 2024, dan menyebutnya sebagai traktat internasional pertama yang mengikat secara hukum di bidang ini. Konvensi ini menuntut agar aktivitas sepanjang siklus hidup AI selaras dengan hak asasi manusia, demokrasi, dan rule of law, serta mendorong penilaian risiko dan dampak yang iteratif. ([Portal](#))

Keistimewaan konvensi ini adalah orientasinya yang eksplisit pada tata hukum dan demokrasi. Ini memperkuat pandangan bahwa AI bukan sekadar persoalan inovasi industri, tetapi juga persoalan konstitusional

dan peradaban. Dalam dunia yang makin terdigitalisasi, kualitas demokrasi dapat dipengaruhi oleh algoritma. Maka, traktat ini penting sebagai pengingat bahwa governance AI harus dilandaskan pada perlindungan institusi demokratis, bukan hanya keamanan produk. ([Portal](#))

### **g. Hiroshima Process dan interoperabilitas global**

Kode etik G7 Hiroshima untuk organisasi yang mengembangkan AI canggih bersifat sukarela, namun penting karena fokus pada model canggih dan generatif. Pada 2025 OECD meluncurkan kerangka pelaporan global untuk memonitor penerapannya, dengan tujuan mendorong keterbandingan, transparansi, dan interoperabilitas antar mekanisme governance AI. Ini menunjukkan arah masa depan governance global: bukan satu standar tunggal untuk semua, melainkan ekosistem standar yang saling terhubung. ([Digital Strategy](#))

## **9. Tantangan Global dalam Membangun Standar Bersama**

Walaupun banyak kemajuan telah dicapai, membangun standar internasional AI tetap menghadapi sejumlah tantangan. Pertama adalah **perbedaan nilai politik**. Negara-negara demokratis liberal cenderung menekankan hak individu, transparansi, dan pembatasan kekuasaan. Negara lain mungkin lebih menekankan stabilitas, kontrol, atau kedaulatan data. Karena itu, standar global selalu merupakan hasil kompromi. ([Portal](#))

Kedua adalah **ketimpangan kapasitas**. Negara maju memiliki sumber daya untuk audit, penilaian dampak, infrastruktur data, dan kepatuhan. Banyak negara berkembang belum memiliki kapasitas regulasi, ahli teknis, atau anggaran yang cukup. Akibatnya, ada risiko terbentuknya "etika AI dua kecepatan": standar tinggi di negara kaya, tetapi implementasi lemah di negara miskin. UNESCO mencoba menjawab ini melalui metodologi kesiapan dan kapasitas implementasi. ([UNESCO](#))

Ketiga adalah **kecepatan inovasi**. Teknologi AI berkembang jauh lebih cepat daripada hukum dan standar formal. OECD sendiri memperbarui prinsipnya pada 2024 agar tetap relevan dengan perkembangan general-purpose dan generative AI. Ini menunjukkan bahwa governance AI harus adaptif, bukan statis. ([OECD](#))

Keempat adalah **fragmentasi rezim**. Saat ini tidak ada satu badan dunia yang memonopoli governance AI. Ada OECD, UNESCO, NIST, ISO/IEC, Uni Eropa, Council of Europe, G7, dan berbagai aturan nasional. Fragmentasi ini kadang membingungkan organisasi, tetapi juga memberi peluang untuk interoperabilitas. OECD pada 2025 secara eksplisit mendorong interoperabilitas dan consistency across international AI governance mechanisms. ([OECD](#))

## 10. Implikasi bagi Organisasi, Pemerintah, dan Pendidikan

Bagi organisasi bisnis, etika AI tidak boleh dipandang sebagai beban kepatuhan semata. Ia adalah prasyarat reputasi, keberlanjutan, dan kepercayaan. Di era ketika publik semakin peka terhadap diskriminasi algoritmik, perusahaan yang gagal mengelola bias dan akuntabilitas akan menghadapi risiko hukum, reputasi, dan sosial. Pendekatan seperti NIST AI RMF dan ISO/IEC 42001 memberikan jalan praktis untuk membangun governance yang matang. ([NIST](#))

Bagi pemerintah, tantangannya adalah menciptakan kebijakan yang melindungi warga tanpa membunuh inovasi. Pendekatan berbasis risiko seperti dalam AI Act memberi pelajaran penting: tidak semua AI perlu diperlakukan sama, tetapi aplikasi yang menyentuh hak dasar harus diawasi lebih ketat. Pada saat yang sama, negara perlu membangun kapasitas audit, literasi digital, dan infrastruktur kebijakan agar tidak hanya menjadi pasar bagi teknologi asing. ([Digital Strategy](#))

Bagi dunia pendidikan, etika AI harus menjadi bagian integral dari kurikulum teknologi, manajemen, hukum, dan ilmu sosial. Mahasiswa

tidak cukup diajarkan membuat model yang akurat; mereka perlu diajarkan memikirkan fairness, oversight, auditability, governance, dan hak pengguna. UNESCO secara eksplisit menempatkan awareness, literacy, dan AI ethics training sebagai area penting. ([UNESCO](#))

Dalam konteks Indonesia dan negara berkembang lain, pelajaran pentingnya adalah bahwa adopsi AI harus dibarengi dengan kedaulatan pengetahuan, data, dan tata kelola. Sistem yang dibangun dari logika sosial negara lain belum tentu cocok bagi konteks lokal. Karena itu, etika AI global perlu diterjemahkan ke dalam konteks nasional: bahasa lokal, struktur sosial, ketimpangan digital, budaya organisasi, dan kebutuhan pembangunan. Standar global memberi kerangka, tetapi implementasi yang adil menuntut sensitivitas lokal. Ini adalah pekerjaan besar bagi akademisi, pembuat kebijakan, regulator, dan praktisi manajemen. ([UNESCO](#))

## **11. Refleksi Normatif: Menuju AI yang Benar-Benar Human-Centered**

Pada akhirnya, pertanyaan etika AI bukan sekadar bagaimana membuat AI lebih pintar, tetapi bagaimana memastikan AI tetap berada dalam horizon kemanusiaan. Teknologi yang mampu memprediksi belum tentu mampu memahami. Sistem yang efisien belum tentu adil. Model yang akurat belum tentu bermartabat. Karena itu, paradigma *human-centered AI* tidak boleh direduksi menjadi fitur antarmuka yang ramah pengguna. Ia harus berarti bahwa manusia tetap menjadi pusat nilai, pusat tanggung jawab, dan pusat tujuan akhir teknologi. ([UNESCO](#))

Etika AI yang matang menolak dua ekstrem. Ekstrem pertama adalah optimisme naif yang menganggap teknologi akan menyelesaikan semua masalah sosial. Ekstrem kedua adalah pesimisme total yang melihat AI hanya sebagai ancaman. Perspektif global yang lebih dewasa justru mencari tata kelola yang membuat AI bermanfaat tanpa mengorbankan keadilan, kebebasan, dan martabat manusia. Kerangka OECD, UNESCO,

NIST, ISO/IEC, Uni Eropa, dan Council of Europe menunjukkan bahwa dunia sedang bergerak ke arah ini, meskipun masih belum sempurna. ([OECD](#))

## 12. Kesimpulan

Etika AI dalam perspektif global merupakan wilayah yang semakin strategis karena AI kini memengaruhi kehidupan sosial, ekonomi, politik, dan budaya pada skala yang luas. Tiga isu utama—bias, akuntabilitas, dan standar internasional—menjadi inti dari perdebatan global tentang bagaimana teknologi ini seharusnya dikembangkan dan digunakan. Bias menunjukkan bahwa AI dapat mewarisi, mereproduksi, bahkan memperbesar ketidakadilan sosial yang telah ada. Akuntabilitas menegaskan bahwa tanggung jawab atas AI tetap harus berada pada manusia dan institusi, bukan dilarutkan ke dalam kompleksitas algoritma. Sementara itu, standar internasional memperlihatkan upaya dunia untuk bergerak dari prinsip normatif menuju tata kelola yang lebih operasional, terukur, dan dalam beberapa kasus mengikat secara hukum. ([UNESCO](#))

Dari OECD yang memperbarui prinsipnya pada 2024, UNESCO yang menegaskan fondasi hak asasi manusia bagi 194 negara anggotanya, NIST yang menawarkan kerangka manajemen risiko praktis, ISO/IEC yang menyediakan standar sistem manajemen dan risiko, sampai Uni Eropa dan Council of Europe yang membawa AI ke ranah hukum formal, terlihat bahwa governance AI global sedang membentuk arsitektur baru. Arsitektur ini belum selesai, tetapi arah dasarnya sudah jelas: AI harus trustworthy, human-centered, auditable, accountable, dan selaras dengan hak-hak dasar manusia. ([OECD](#))

Bagi dunia akademik, bisnis, dan pemerintahan, pesan utamanya sederhana tetapi mendalam: AI tidak boleh hanya dinilai dari kecanggihan teknisnya, melainkan dari kualitas moral dan sosial penggunaannya. Pertanyaan sentral abad ini bukan hanya apakah mesin dapat mengambil keputusan, melainkan apakah manusia masih cukup

bijaksana untuk mengatur keputusan yang diambil dengan bantuan mesin. Masa depan etika AI pada akhirnya bukan ditentukan oleh algoritma semata, tetapi oleh keberanian manusia untuk menempatkan keadilan, tanggung jawab, dan martabat di atas sekadar efisiensi dan kecepatan. ([UNESCO](#))

---

Berikut **glosarium** dan **daftar pustaka** untuk topik **“Etika AI dalam Perspektif Global: Bias, Akuntabilitas, dan Standar Internasional.”**

## **Glosarium**

### **Artificial Intelligence (AI)**

Sistem berbasis mesin yang dirancang untuk menghasilkan keluaran seperti prediksi, rekomendasi, keputusan, atau konten yang memengaruhi lingkungan fisik maupun virtual. OECD memperbarui definisinya pada 2024 agar tetap relevan dengan perkembangan AI mutakhir. ([OECD](#))

### **Etika AI**

Bidang kajian normatif dan praktis yang membahas bagaimana AI harus dirancang, dikembangkan, digunakan, dan diawasi agar selaras dengan martabat manusia, hak asasi manusia, keadilan, transparansi, dan tanggung jawab. UNESCO menempatkan perlindungan hak asasi manusia dan martabat manusia sebagai fondasi utama etika AI. ([UNESCO](#))

### **Bias Algoritmik**

Kecenderungan sistem AI menghasilkan keluaran yang tidak adil,

diskriminatif, atau tidak proporsional terhadap individu atau kelompok tertentu karena masalah data, desain model, tujuan optimasi, atau konteks implementasi. UNESCO mengakui bias sebagai salah satu risiko sentral AI. ([UNESCO](#))

### **Bias Historis**

Bias yang berasal dari data masa lalu yang sudah mengandung ketimpangan sosial, diskriminasi, atau pola eksklusi, lalu diwariskan ke dalam model AI.

### **Bias Representasi**

Bias yang muncul ketika data pelatihan tidak cukup mewakili keragaman populasi, wilayah, bahasa, gender, atau kelompok sosial tertentu.

### **Bias Pengukuran**

Bias yang timbul ketika konsep sosial yang kompleks direduksi menjadi indikator atau proksi yang terlalu sempit untuk diproses model.

### **Fairness (Keadilan)**

Prinsip bahwa sistem AI harus memperlakukan individu dan kelompok secara adil serta tidak menghasilkan diskriminasi yang tidak sah. Fairness menjadi salah satu prinsip penting dalam kerangka UNESCO dan OECD. ([UNESCO](#))

### **Transparency (Transparansi)**

Keterbukaan mengenai cara sistem AI dirancang, digunakan, dan memengaruhi keputusan. Transparansi membantu evaluasi, audit, dan pengawasan publik. UNESCO menempatkan transparansi sebagai prinsip penting dalam AI ethics. ([UNESCO](#))

### **Explainability (Keterjelasan/Penjelasan)**

Kemampuan untuk memberikan penjelasan yang dapat dipahami mengenai bagaimana dan mengapa sistem AI menghasilkan keluaran tertentu. Dalam praktik, tingkat explainability sering menyesuaikan konteks risiko dan penggunaan. ([UNESCO](#))

### **Accountability (Akuntabilitas)**

Kewajiban individu, organisasi, atau institusi untuk bertanggung jawab atas desain, penerapan, dampak, dan konsekuensi sistem AI. UNESCO menekankan bahwa AI tidak boleh menggantikan tanggung jawab manusia yang bersifat final. ([UNESCO](#))

### **Human Oversight (Pengawasan Manusia)**

Keterlibatan manusia yang bermakna dalam mengawasi, meninjau, mengintervensi, atau menghentikan sistem AI jika diperlukan. Prinsip ini menjadi unsur inti dalam kerangka UNESCO dan pendekatan risk governance internasional. ([UNESCO](#))

### **Auditability (Kemampuan Diaudit)**

Kemampuan suatu sistem AI untuk ditelusuri, diperiksa, dan dievaluasi melalui dokumentasi, log, proses validasi, atau mekanisme audit independen.

### **Traceability (Keterlacakan)**

Kemampuan untuk menelusuri asal-usul data, perubahan model, keputusan desain, dan keluaran sistem AI sepanjang siklus hidupnya. ISO/IEC 42001 menekankan pentingnya pengelolaan AI yang terstruktur dan dapat ditelusuri. ([ISO](#))

### **Trustworthy AI**

AI yang dapat dipercaya karena memenuhi unsur-unsur seperti keamanan, reliabilitas, fairness, transparansi, akuntabilitas, dan keselarasan dengan hak asasi manusia. OECD dan NIST sama-sama menggunakan gagasan ini sebagai orientasi governance AI. ([OECD](#))

### **AI Risk Management**

Pendekatan sistematis untuk mengidentifikasi, memetakan, mengukur, memitigasi, dan memantau risiko AI terhadap individu, organisasi, dan masyarakat. NIST AI RMF dikembangkan khusus untuk tujuan ini. ([NIST](#))

## **NIST AI RMF**

Kerangka kerja dari National Institute of Standards and Technology untuk membantu organisasi mengelola risiko AI secara sukarela melalui empat fungsi utama: Govern, Map, Measure, dan Manage. ([NIST](#))

### **Govern–Map–Measure–Manage**

Empat fungsi utama dalam NIST AI RMF. *Govern* berfokus pada tata kelola, *Map* pada konteks dan pemetaan risiko, *Measure* pada evaluasi risiko, dan *Manage* pada penanganan serta pemantauan risiko. ([NIST Publications](#))

## **ISO/IEC 42001**

Standar internasional pertama untuk sistem manajemen AI. Standar ini membantu organisasi membangun, menerapkan, memelihara, dan meningkatkan sistem manajemen AI secara berkelanjutan. ([ISO](#))

## **ISO/IEC 23894**

Standar panduan manajemen risiko AI yang melengkapi tata kelola organisasi dalam mengelola risiko spesifik sistem AI. ISO menempatkannya dalam keluarga standar AI yang relevan untuk governance. ([ISO](#))

## **Standar Internasional**

Dokumen acuan lintas negara yang menetapkan prinsip, persyaratan, atau panduan bersama bagi pengembangan dan penggunaan AI. Contohnya meliputi OECD AI Principles, UNESCO Recommendation, NIST AI RMF, dan ISO/IEC 42001. ([OECD](#))

## **Regulasi Berbasis Risiko**

Pendekatan hukum atau kebijakan yang membedakan tingkat pengawasan dan kewajiban berdasarkan tingkat risiko aplikasi AI. Pendekatan ini menjadi ciri utama rezim AI modern seperti AI Act di Eropa. Ini saya simpulkan dari kerangka resmi yang menekankan diferensiasi kewajiban menurut tingkat risiko. ([OECD](#))

### **Human Rights-Based Approach**

Pendekatan tata kelola AI yang menempatkan hak asasi manusia sebagai kerangka dasar dalam seluruh siklus hidup AI. UNESCO menjadikannya prinsip utama dalam recommendation-nya. ([UNESCO](#))

### **Sistem Sosio-Teknis**

Pandangan bahwa AI bukan hanya artefak teknis, melainkan bagian dari jaringan manusia, institusi, data, aturan, dan budaya organisasi yang bersama-sama menentukan dampaknya.

### **Continuous Improvement (Perbaikan Berkelanjutan)**

Prinsip manajemen yang menuntut organisasi terus meninjau, memperbaiki, dan memutakhirkan sistem AI seiring perubahan risiko, teknologi, dan konteks penggunaan. ISO/IEC 42001 secara eksplisit menekankan continual improvement. ([ISO](#))

---

### **Daftar Pustaka**

Berikut daftar pustaka yang sudah dirapikan dalam gaya **APA 7**.

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. U.S. Department of Commerce. ([NIST Publications](#))

National Institute of Standards and Technology. (n.d.). *AI Risk Management Framework*. ([NIST](#))

National Institute of Standards and Technology. (n.d.). *NIST AI RMF Playbook*. ([NIST](#))

Organisation for Economic Co-operation and Development. (2019/2024). *AI Principles*. OECD. ([OECD](#))

Organisation for Economic Co-operation and Development. (2024, May 3). *OECD updates AI Principles to stay abreast of rapid technological developments.* ([OECD](#))

Organisation for Economic Co-operation and Development. (2024). *Explanatory memorandum on the updated OECD definition of an AI system.* OECD AI Papers. ([OECD](#))

Organisation for Economic Co-operation and Development. (2025, February 20). *OECD launches global framework to monitor application of G7 Hiroshima AI Code of Conduct.* ([OECD](#))

International Organization for Standardization. (2023). *ISO/IEC 42001:2023 Information technology—Artificial intelligence—Management system.* ISO. ([ISO](#))

International Organization for Standardization. (n.d.). *AI management systems: What businesses need to know.* ISO. ([ISO](#))

International Organization for Standardization. (n.d.). *ISO 42001 explained: What it is.* ISO. ([ISO](#))

International Organization for Standardization. (2023). *ISO/IEC 23894 Information technology—Artificial intelligence—Guidance on risk management.* ISO. ([ISO](#))

United Nations Educational, Scientific and Cultural Organization. (2021). *Recommendation on the Ethics of Artificial Intelligence.* UNESCO. ([UNESCO Digital Library](#))

United Nations Educational, Scientific and Cultural Organization. (n.d.). *Recommendation on the Ethics of Artificial Intelligence.* ([UNESCO](#))

United Nations Educational, Scientific and Cultural Organization. (2023, July 20). *UNESCO's Recommendation on the Ethics of Artificial Intelligence: Key facts.* ([UNESCO](#))

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

United Nations Educational, Scientific and Cultural Organization. (n.d).  
*Global AI Ethics and Governance Observatory.* ([UNESCO](#))

---

Copilot for this article - Chatgpt 5.2 Thinking. Access date: 6 Maret 2026  
Prompting on Writer's account ([Rudy C Tarumingkeng](#))

<https://chatgpt.com/c/69aa827e-b5bc-8398-9c31-72a3ba4d566b>