

# IMPLIKASI *ARTIFICIAL INTELLIGENCE* TERHADAP HAK ASASI MANUSIA DAN INTEGRASI PRINSIP-PRINSIP HAM INTERNASIONAL DALAM SIKLUS HIDUP AI

**Ayu Riska Amalia**

Universitas Mataram, Indonesia

E-mail: aramel@unram.ac.id

**Diva Pitaloka**

Universitas Mataram, Indonesia

**Syamsul Mujtahidin**

Universitas Mataram, Indonesia

## *Abstract*

Kecerdasan Buatan (AI) telah mengubah cara manusia bekerja, berinteraksi, dan membuat keputusan. AI tidak hanya berkontribusi pada efisiensi dan inovasi di seluruh sektor seperti kesehatan, pendidikan, dan tata kelola, tetapi juga menghasilkan risiko sistemik terhadap privasi, kesetaraan, non-diskriminasi, kebebasan berekspresi, dan hak untuk bekerja. Tantangan-tantangan ini menimbulkan kesenjangan normatif antara perkembangan teknologi dan kerangka kerja hak asasi manusia internasional yang ada. Artikel ini merumuskan kerangka kerja tata kelola berbasis hak asasi manusia untuk menanamkan prinsip-prinsip hak asasi manusia di seluruh siklus hidup AI mengacu pada instrumen hak asasi manusia utama seperti Deklarasi Universal Hak Asasi Manusia (UDHR), Kovenan Internasional tentang Hak Sipil dan Politik (ICCPR), dan Kovenan Internasional tentang Hak Ekonomi, Sosial dan Budaya (ICESCR), serta kerangka kerja tata kelola AI global termasuk Rekomendasi UNESCO tentang Etika Kecerdasan Buatan (2021), Prinsip AI OECD (2019), dan Undang-Undang AI UE (2024). Penelitian ini mengidentifikasi enam prinsip yang harus digunakan dalam memandu setiap tahap siklus hidup AI, yakni martabat manusia, legalitas, kebutuhan dan proporsionalitas, kesetaraan dan non-diskriminasi, privasi dan perlindungan data, transparansi dan penjelasan, dan pengawasan manusia yang bermakna dengan akuntabilitas dan pemulihan. Penelitian ini berargumen bahwa pendekatan human rights by design sangat penting untuk memastikan bahwa sistem kecerdasan buatan tetap sah, adil, dan transparan. Selain itu, negara-negara, khususnya Indonesia, perlu untuk menyelaraskan tata kelola AI nasional dengan standar internasional untuk menjaga hak asasi manusia di tengah perkembangan teknologi.

**Kata kunci:** *artificial intelligence; HAM; hukum internasional; AI governance; human rights by design*

## *Abstract*

*Artificial is changing the way humans work, interact, and make decisions. AI not only contributes to efficiency and innovation across sectors such as health, education, and governance, but also generates systemic risks to privacy, equality, non-discrimination, freedom of expression, and the right to work. These challenges expose the normative gap between technological developments and the existing international human rights framework. Drawing from key human rights instruments such as the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the International Covenant on Economic, Social and Cultural Rights (ICESCR), as well as global AI governance frameworks including the UNESCO Recommendations on the Ethics of Artificial Intelligence (2021), the OECD AI Principles (2019), and the EU AI Act (2024), This article formulates a human rights-based governance framework for embedding human rights principles throughout the AI lifecycle. It identifies six intersecting principles that should guide each stage of the AI lifecycle: human dignity, legality, necessity*

and proportionality, equality and non-discrimination, data privacy and protection, transparency and explanation, and meaningful human oversight with accountability and restoration. The study argues that a human rights approach by design, integrating these principles from conception to implementation, is critical to ensuring that AI systems remain legitimate, fair, and transparent. Finally, the article emphasizes the urgency for countries, especially Indonesia, to align national AI governance with international standards to safeguard human rights in the midst of technological development.

**Keywords:** artificial intelligence; human rights; international law; AI governance; human rights by design

## A. PENDAHULUAN

*Artificial Intelligence* (AI) merupakan representasi dari sebuah lompatan jauh dalam bidang sains dan digital hari ini. Penggunaan AI telah merambah pada setiap sektor kehidupan seperti pendidikan, sains, budaya, komunikasi dan informasi, kesehatan, pertanian, lingkungan, dan berbagai aspek kehidupan lainnya.<sup>1</sup> Secara istilah, AI merujuk pada kemampuan mesin dan sistem untuk memperoleh dan menerapkan pengetahuan serta menjalankan perilaku cerdas,<sup>2</sup> dan tidak terbatas hanya pada pemrograman, tetapi juga mencakup pemodelan proses kognitif manusia yang dapat belajar, berpikir, dan memecahkan masalah.<sup>3</sup>

Dampak AI hari ini terlihat dalam berbagai sektor kehidupan. Dalam aspek kesehatan, pasar AI di bidang kesehatan diproyeksikan tumbuh ~43% per tahun (CAGR 2024–2032).<sup>4</sup> Dalam aspek pendidikan, AI telah digunakan 60% guru untuk mempermudah pekerjaan,<sup>5</sup> dan dipercaya sebagai alat esensial untuk menunjang kesuksesan siswa di perguruan tinggi dan dunia kerja.<sup>6</sup> Dalam penggunaan media sosial, algoritma AI berperan dalam merekomendasikan 80% konten kepada pengguna di platform. Bahkan menurut Forbes, 71% gambar di sosial media merupakan *AI-generated*.<sup>7</sup> AI bahkan juga diprediksi akan mengubah struktur ketenagakerjaan secara signifikan dalam dekade ini. Menurut *Future of Jobs Report 2023 (WEF)*, sebanyak 83 juta pekerjaan diperkirakan akan terdisrupsi akibat automasi AI, namun 69 juta pekerjaan baru akan tercipta pada tahun 2027, menghasilkan netto -14 juta pekerjaan (sekitar -2% total tenaga kerja global).<sup>8</sup> Berbagai fenomena ini menandakan bahwa AI telah menjadi bagian integral yang akan mengubah cara manusia bekerja, berinteraksi dan menjalani kehidupan sehari-hari.

Namun demikian, masuknya AI dalam berbagai sendi kehidupan tidak hanya memberikan manfaat, tapi juga membawa tantangan terhadap kehidupan sosial dan hak asasi manusia (HAM), seperti hak atas privasi, non-diskriminasi dan kesetaraan, kebebasan berpendapat dan berekspresi, dan hak-hak fundamental lainnya. Sebagai contoh, algoritma AI yang digunakan dalam pengambilan keputusan seperti dalam rekrutmen, penegakan hukum, atau distribusi sumber daya berisiko menghasilkan bias diskriminatif jika data pelatihannya tidak inklusif atau

<sup>1</sup> UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), <https://unesdoc.unesco.org/ark:/48223/pf0000380455>

<sup>2</sup> UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000380455>

<sup>3</sup> John McCarthy, "What Is Artificial Intelligence?" (Stanford University, November 12, 2007), <http://www-formal.stanford.edu/jmc/whatisai.pdf>

<sup>4</sup> World Economic Forum, *The Future of AI-Enabled Health 2025*, SlideShare, <https://www.slideshare.net/slideshow/wef-the-future-of-ai-enabled-health-2025-pdf/275311854#:~:text=reaching%20a%20total%20value%20of,GenAI>

<sup>5</sup> World Economic Forum, "How AI and Human Teachers Can Collaborate to Transform Education," January 2025, <https://www.weforum.org/stories/2025/01/how-ai-and-human-teachers-can-collaborate-to-transform-education/#:~:text=But%20the%20real%20magic%20happens,feedback%20and%20facilitating%20group%20discussions>.

<sup>6</sup> Derek Li Haoyang dan Jason Towne, "How AI and Human Teachers Can Collaborate to Transform Education," *World Economic Forum*, 9 Januari 2025, diakses 7 September 2025, <https://www.weforum.org/stories/2025/01/how-ai-and-human-teachers-can-collaborate-to-transform-education/#:~:text=,is%20available%20to%20all%20students>.

<sup>7</sup> ArtSmart AI, "AI in Social Media Statistics," accessed September 7, 2025, <https://artsmart.ai/blog/ai-in-social-media-statistics/#:~:text=%2A%20Over%2080,%28Saufter>

<sup>8</sup> AIMultiple, "Top 18 Predictions from Experts on AI Job Loss," *AIMultiple Research*, diperbarui 2 September 2025, akses 7 September 2025, <https://research.aimultiple.com/ai-job-loss/#:~:text=World%20Economic%20Forum%20,2023>

jika desainnya mengabaikan prinsip keadilan.<sup>9</sup> Pengumpulan data pribadi dalam jumlah besar yang digunakan AI juga memiliki kaitan erat dengan pelanggaran privasi yang dijamin dalam Pasal 12 Deklarasi Universal Hak asasi Manusia (DUHAM) 1948<sup>10</sup> serta pasal 17 Kovenan Internasional tentang Hak Sipil dan Politik (ICCPR) 1966.<sup>11</sup> Dalam konteks Hukum Nasional Indonesia, praktik pengumpulan data dalam jumlah yang masif juga berpotensi melanggar hak privasi yang dijamin oleh Pasal 28G ayat (1) Undang-Undang Dasar 1945, yang menyatakan bahwa setiap orang berhak atas perlindungan diri pribadi, keluarga, kehormatan, martabat, dan harta benda.

Penggunaan AI juga berkaitan dengan hak kebebasan berpendapat dan berekspresi yang dapat dilihat pada penerapan sistem moderasi konten berbasis AI oleh platform media sosial yang kerap menghapus dan membayangi (*shadow-ban*) unggahan yang sama sekali tidak melanggar hukum. Hal ini bisa ditunjukkan dengan postingan-postingan aktivis di wilayah Gaza selama eskalasi konflik Palestina-Israel pada 2021 hingga saat ini yang mengalami penghapusan otomatis,<sup>12</sup> yang menunjukkan bagaimana bias dan kekeliruan teknis dapat secara sewenang-wenang membungkam suara kelompok rentan dan aktivis masyarakat sipil, hak yang telah dijamin oleh Pasal 19 DUHAM 1948, Pasal 19 ICCPR 1966, dan Pasal 28E ayat (3) UUD 1945<sup>13</sup>.

Hal tersebut di atas hanyalah sebagian dari isu HAM terkait AI, sehingga untuk memastikan bahwa manfaat AI tidak melebihi dampak negatifnya, maka diperlukan batasan berupa regulasi yang memadai. Jika menilik pada kerangka hukum Internasional, maka belum ada traktat khusus yang mengikat tentang AI dan HAM secara komprehensif, namun beberapa prinsip terkait itu dapat ditemukan dalam DUHAM 1948 dan ICCPR 1966 yang memberikan landasan perlindungan privasi dan kebebasan berekspresi. Selain itu, terdapat aturan khusus yang mengatur AI seperti *UNESCO Recommendation on the Ethics of Artificial Intelligence* (2021) yang menekankan transparansi dan keadilan dalam pengembangan AI, *EU AI Act* (2024) yang mengatur AI beresiko tinggi, dan *OECD AI Principles* yang mendorong inovasi AI yang bertanggung jawab dan inklusif.

Berangkat dari peta masalah tersebut, artikel ini menguraikan lebih jauh dampak AI terhadap hak asasi manusia, menganalisis keterbatasan Instrumen HAM yang ada, dan merumuskan prinsip HAM yang harus ditanamkan di sepanjang siklus hidup AI berdasarkan Instrumen hukum dan kerangka etik global, serta mengelaborasi pentingnya harmonisasi regulasi AI di Tingkat nasional.

## B. METODE PENELITIAN

Penelitian ini menggunakan pendekatan hukum normatif dengan berfokus kepada kajian konseptual terhadap instrument internasional tentang AI yang utama, yakni *Universal Declaration of Human Rights* (UDHR), *International Covenant on Civil and Political Rights* (ICCPR), dan *International Covenant on Economic, Social and Cultural Rights* (ICESCR) dan 3 instrumen tata Kelola AI tingkat internasional yang utama mencakup *UNESCO Recommendation on the Ethics of Artificial Intelligence* (2021), *OECD AI Principles*

<sup>9</sup> Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016).

<sup>10</sup>United Nations, "Universal Declaration of Human Rights," United Nations General Assembly Resolution 217A (III), December 10, 1948, <https://www.un.org/en/about-us/universal-declaration-of-human-rights>.

<sup>11</sup>United Nations, *International Covenant on Civil and Political Rights*, United Nations General Assembly Resolution 2200A (XXI), December 16, 1966, <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>.

<sup>12</sup>Human Rights Watch, *Meta's Broken Promises: Systemic Censorship of Palestine Content on Instagram and Facebook* (New York: Human Rights Watch, December 21, 2023), <https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content-instagram-and-facebook>.

<sup>13</sup>Indonesia, *Undang-Undang Dasar Negara Republik Indonesia Tahun 1945* (Jakarta: Sekretariat Negara Republik Indonesia, 1945).

(2019)), dan *EU Artificial Intelligence Act (2024)*. Pendekatan ini digunakan untuk menelaah secara sistematis norma dan prinsip yang mbingkai hubungan antara perkembangan AI dan perlindungan hak asasi manusia (HAM) dalam konteks hukum internasional.

Analisis dilakukan melalui studi kepustakaan dengan mengkaji bahan hukum primer, sekunder, dan tersier untuk menemukan asas dan prinsip hukum yang dapat dijadikan dasar dalam membangun tata kelola AI berbasis HAM. Penelitian ini bertujuan untuk menelaah dampak AI terhadap hak asasi manusia, mengidentifikasi prinsip-prinsip HAM yang terkandung dalam instrumen internasional tersebut, menilai relevansinya terhadap konteks hukum nasional Indonesia, serta merumuskan prinsip-prinsip HAM yang perlu ditanamkan ke dalam setiap tahapan siklus hidup AI. Dengan demikian, penelitian ini diharapkan dapat memberikan landasan normatif bagi pengembangan kebijakan hukum nasional yang sejalan dengan standar dan nilai-nilai HAM internasional dalam menghadapi perkembangan teknologi AI

## C. PEMBAHASAN

### 1. *Artificial Intelligence* dalam perspektif Hak Asasi Manusia

#### A. Perkembangan AI

Kecerdasan buatan atau AI secara konseptual dapat diartikan sebagai sistem komputer atau mesin yang memiliki kemampuan untuk meniru kecerdasan manusia, termasuk diantaranya belajar dari pengalaman (*learning*), mengenali pola (*pattern recognition*), membuat keputusan (*decision-making*), dan menyelesaikan masalah (*problem-solving*).<sup>14</sup> Definisi ini bisa saja berubah seiring waktu sesuai dengan perkembangan teknologi.<sup>15</sup> Ide awal kecerdasan buatan yang kita kenal saat ini pertama kali muncul dalam karya Alan Turing pada 1950 melalui artikelnya “*Computing Machinery and Intelligence*” yang memperkenalkan konsep Tes Turing dan kemudian dikemukakan secara resmi sebagai sebuah konsep pada tahun 1956 oleh John McCarthy dalam konferensi Dartmouth di Hanover, New Hampshire, AS.<sup>16</sup>

Pada awalnya, AI dikembangkan untuk menjalankan tugas-tugas otomatisasi sederhana seperti pengendalian mesin industri dan komputasi matematis. Namun, dalam dua dekade terakhir, muncul lonjakan signifikan dengan hadirnya *machine learning* (ML), yakni cabang AI yang memungkinkan sistem belajar dari data tanpa diprogram secara eksplisit.<sup>17</sup> Selanjutnya, ML melalui subbidang bernama *deep learning* memanfaatkan jaringan saraf tiruan berlapis-lapis (*neural networks*) untuk menganalisis data dalam jumlah besar dengan tingkat akurasi yang tinggi, terutama dalam pengenalan gambar, suara, dan bahasa.<sup>18</sup>

AI kemudian mengalami pergeseran paradigma seiring dengan hadirnya kecerdasan generatif (*AI-generative*) seperti GPT-4, Gemini, dan DALL-E yang menghasilkan kemampuan-kemampuan baru yang muncul secara tak terduga dan mengubah cara manusia berinteraksi dan bertukar informasi.<sup>19</sup> Kecerdasan generatif ini memunculkan paradigma baru, di mana mesin tidak hanya menjadi alat bantu tetapi juga *co-creator* dalam aktivitas sosial, ekonomi, hingga hukum. Lebih lanjut lagi, AI membawa pergeseran otoritas dari manusia ke mesin dengan kemampuan *autonomous decision-making* nya, seperti yang ditunjukkan oleh

<sup>14</sup>Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. (Boston: Pearson Education, 2021).

<sup>15</sup>United Nations Educational, Scientific and Cultural Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), 3, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

<sup>16</sup>Bruce G. Buchanan, “A (Very) Brief History of Artificial Intelligence,” *AI Magazine* 26, no. 4 (2006): 53–60, <https://www.aaai.org/ojs/index.php/aimagazine/article/view/1868>

<sup>17</sup>Michael I. Jordan and Tom M. Mitchell, “Machine Learning: Trends, Perspectives, and Prospects,” *Science* 349, no. 6245 (2015): 255–60, <https://doi.org/10.1126/science.aaa8415>

<sup>18</sup>Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, “Deep Learning,” *Nature* 521, no. 7553 (2015): 436–44, <https://doi.org/10.1038/nature14539>.

<sup>19</sup>Rishi Bommasani et al., *On the Opportunities and Risks of Foundation Models* (Stanford, CA: Stanford Center for Research on Foundation Models, 2021), <https://arxiv.org/abs/2108.07258>.

sistem senjata otonom dan algoritma pengambilan keputusan dalam perekrutan tenaga kerja. Hal ini menimbulkan pertanyaan tentang legitimasi, akuntabilitas, dan etika, terutama ketika keputusan tersebut berdampak langsung terhadap hak asasi manusia.

## B. Implikasi AI terhadap Hak Asasi Manusia

Kehadiran AI dalam kehidupan manusia seperti dua sisi mata uang. Di satu sisi, AI memberikan kemudahan dan memberi solusi dari permasalahan sehari-hari hingga hal-hal kompleks seperti perubahan iklim dan bidang Kesehatan.<sup>20</sup> Namun di sisi lain, penerapan AI tanpa prinsip kehati-hatian dan akuntabilitas yang kuat dapat memperlebar ketimpangan sosial, memperkuat diskriminasi sistemik, dan melemahkan perlindungan HAM.<sup>21</sup>

Hukum HAM Internasional telah menetapkan berbagai instrumen HAM yang menjadi norma dasar dan prinsip universal dalam menjalankan kewajiban negara dalam menghormati melindungi, dan memenuhi HAM, termasuk dalam pengembangan dan penerapan teknologi seperti AI. Berbagai hak fundamental yang dijamin dalam *International Bill of Rights* (UDHR 1948, ICCPR 1966, dan ICESCR 1966) sangat mungkin terdampak oleh penggunaan AI. Untuk keperluan penelitian, tulisan ini akan ditekankan pada hak-hak fundamental yang memiliki potensi besar terdampak oleh perkembangan AI, diantaranya privasi, kesetaraan dan non-diskriminasi, kebebasan berekspresi, serta hak atas pekerjaan namun tidak terbatas pada hak-hak lainnya, mengingat sifat HAM yang saling bergantung dan tak terpisahkan.

### 1) *Big data* dan Hak Privasi

Kemajuan teknologi *deep learning* pada AI ditopang oleh kemajuan komputasi dan ketersediaan *big data* yang digunakan dalam pelatihan.<sup>22</sup> Pengumpulan data pribadi dengan jumlah yang sangat besar menimbulkan tantangan signifikan pada perlindungan hak privasi.<sup>23</sup> Jejak digital yang tercipta dari data tersebut dapat dilacak dan dimonitor, sehingga ketika serpihan-serpihan data digabungkan, dapat terbentuk sebuah profil detail mengenai individu yang bersangkutan. Sebagai contoh, Google memiliki akses terhadap beragam data pribadi, mulai dari riwayat pencarian, pendidikan, hobi, tontonan, hingga alamat dan aspek kehidupan lainnya.<sup>24</sup>

Data yang dikumpulkan untuk suatu tujuan tertentu seringkali digunakan untuk kembali untuk tujuan lain tanpa sepengetahuan atau persetujuan individu yang bersangkutan. Sehingga, tanpa adanya regulasi yang jelas, data yang ada berpotensi disalahgunakan dan membahayakan privasi maupun keamanan individu. Penggunaan data tanpa persetujuan eksplisit mengancam hak atas privasi sebagaimana dijamin dalam Pasal 12 UDHR, Pasal 17 ICCPR, dan Pasal 28G UUD 1945. Di negara seperti Indonesia, insfastruktur perlindungan data yang lemah membuat potensi pelanggaran privasi lebih tinggi. Menurut laporan VIDA (2024), 56% bisnis di Indonesia telah mengalami penipuan identitas berbasis AI, termasuk *deepfake fraud* yang melonjak 1.550% pada 2023-2024.<sup>25</sup> Oleh sebab itu, regulasi terkait AI perlu menetapkan aturan yang lebih ketat dalam pengumpulan, penyimpanan, dan penggunaan data, disertai mekanisme perlindungan yang efektif terhadap pelanggaran privasi.

<sup>20</sup>Raúl Vinuesa et al., "The Role of Artificial Intelligence in Achieving the Sustainable Development Goals," *Nature Communications* 11, no. 233 (2020), <https://doi.org/10.1038/s41467-019-14108-y>.

<sup>21</sup>Mark Latonero, *Governing Artificial Intelligence: Upholding Human Rights & Dignity* (New York: Data & Society Research Institute, 2018), [https://datasociety.net/wp-content/uploads/2018/10/DataSociety\\_Governing\\_Artificial\\_Intelligence\\_Upholding\\_Human\\_Rights.pdf](https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf).

<sup>22</sup>Amirhosein Toosi et al., "A Brief History of AI: How to Prevent Another Winter (A Critical Review)," *PET Clinics* 16, no. 4 (October 2021): 449–69, <https://www.sciencedirect.com/science/article/pii/S1936879821004493>.

<sup>23</sup>K Emmanuel Kuye, *Digital Information Privacy in the Age of Big Data: Challenges and Solutions* (2023), [https://www.researchgate.net/publication/376448455\\_Digital\\_Information\\_Privacy\\_in\\_the\\_Age\\_of\\_Big\\_Data\\_Challenges\\_and\\_Solutions](https://www.researchgate.net/publication/376448455_Digital_Information_Privacy_in_the_Age_of_Big_Data_Challenges_and_Solutions).

<sup>24</sup>Executive Office of the President, *Big Data – Seizing Opportunities, Preserving Values* (Washington, DC: The White House, 2014).

<sup>25</sup>VIDA, *Where's The Fraud: Protecting Indonesian Businesses from AI-Generated Digital Fraud*. Jakarta: VIDA, 2024.

## 2) Bias algoritmik dan Non-diskriminasi

AI seringkali memperburuk ketidakadilan yang ada karena hasil bias yang diberikan berasal dari data latih dan algoritma yang bias pula. Data latih yang berasal dari sumber maupun proses seleksi yang tidak netral akan menghasilkan hasil yang serupa. Sehingga jika data digunakan pada sistem AI bias terhadap kelompok tertentu maka output yang sama akan dihasilkan terhadap kelompok tersebut. Model AI berbasis pembelajaran mesin membutuhkan data dalam jumlah besar untuk pelatihan, tetapi data yang tidak merata sering kali hanya mewakili kelompok dominan, seperti populasi urban atau etnis mayoritas sehingga terjadi bias yang membatasi akses bagi kelompok yang kurang terwakili. Hal ini tergambar dalam penelitian di AS yang menunjukkan algoritma kesehatan memberikan prioritas lebih rendah kepada pasien kulit hitam karena data pelatihan didominasi oleh pasien kulit putih, yang menyebabkan ketidaksetaraan akses terhadap layanan kesehatan.<sup>26</sup> Demikian pula, sistem pengenalan wajah menunjukkan akurasi lebih rendah untuk perempuan berkulit gelap karena dataset yang tidak representatif, mempersulit akses mereka ke teknologi yang adil.<sup>27</sup>

Selain ketidaksetaraan dalam hal luaran AI, ketidaksetaraan juga berpotensi muncul dalam akses terhadap AI, mengingat bahwa teknologi seperti AI seringkali hanya dapat diakses oleh segmen yang lebih mapan atau yang tinggal di daerah perkotaan, sehingga menyebabkan Masyarakat miskin dan pedesaan tertinggal. Oleh sebab itu, AI dapat meminggirkan kelompok tertentu, seperti mereka yang tidak memiliki literasi digital yang baik dan mereka yang merupakan penyandang disabilitas. Dalam pendidikan, teknologi AI sering kali hanya relevan bagi komunitas urban dengan infrastruktur digital, sehingga komunitas pedesaan menjadi tertinggal.<sup>28</sup> Ketidaksetaraan akses ini melanggar prinsip kesetaraan HAM, memperlebar kesenjangan sosial dan digital. Bias AI dan ketidaksetaraan akses terhadap AI melanggar prinsip kesetaraan dalam Pasal 2 dan 26 ICCPR.

## 3) Moderasi konten otomatis terhadap Hak Kebebasan Berekspresi

Penggunaan AI dan algoritma dapat menjadi aset bagi kebebasan berekspresi dan informasi, dengan memanfaatkan kecerdasan buatan dalam menyampaikan informasi. Namun, berbagai studi juga mengungkapkan bahwa teknologi ini juga berpotensi menimbulkan ancaman serius terhadap hak-hak tersebut.<sup>29</sup> Agar berfungsi optimal, alat algoritmik memerlukan data pribadi pengguna dalam jumlah yang besar, namun jika ditilik dalam sistem *surveillance capitalism*, hal ini dapat menjadi penghalang kebebasan. Pengumpulan informasi tersebut dapat membuat individu menahan diri dalam menggunakan hak kebebasan berekspresi mereka.<sup>30</sup> Untuk alasan yang sama anonimitas dan enkripsi memegang peran penting dalam memajukan kebebasan berbicara.

Pasal 19 ICCPR menegaskan bahwa pembatasan terhadap hak kebebasan berekspresi hanya dapat dilakukan apabila memenuhi tiga syarat kumulatif sebagaimana diatur dalam Pasal 19 ayat (3) ICCPR dan ditafsirkan dalam General Comment No. 34 (2011) yakni pembatasan tersebut harus ditetapkan oleh hukum, ditujukan untuk tujuan yang sah seperti melindungi

<sup>26</sup>Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations. *Science*, 366(6464), 447-453.

<sup>27</sup>Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (2018).

<sup>28</sup>United Nations Educational, Scientific and Cultural Organization (UNESCO), *AI and Education: Guidance for Policy-Makers* (Paris: UNESCO, 2021).

<sup>29</sup>Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: PublicAffairs, 2019), hlm. 8–12; Lihat juga: Executive Office of the President of the United States, *Big Data: Seizing Opportunities, Preserving Values* (Washington, DC: White House, 2014), hlm. 55–60 tentang efek pengawasan terhadap kebebasan berekspresi.

<sup>30</sup>UN Human Rights Council, Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, A/HRC/29/32 (22 May 2015), paras. 7–11,

hak atau reputasi orang lain maupun keamanan nasional, dan harus benar-benar diperlukan serta proporsional. Sehingga moderasi konten yang tidak tunduk pada kriteria ini merupakan pelanggaran pada hak berekspresi.

Dalam moderasi konten otomatis di platform media sosial, algoritma AI sering kali membuat keputusan untuk menghapus konten atau menyensor diam-diam melalui *filtering* dan *shadow banning* yang membentuk sensor digital yang efektif dan tidak terlihat.<sup>31</sup> Konten pengguna akan disembunyikan atau mengalami penurunan visibilitas tanpa pemberitahuan dari platform.<sup>32</sup> Karena sifat *black box*, pengguna tidak dapat memahami alasan di balik keputusan ini, yang dapat membatasi kebebasan berekspresi.

Ketidaktransparan ini juga menyulitkan regulator untuk memastikan bahwa platform tidak menyalahgunakan kekuasaan mereka. Melalui *algorithmic repression* ini, teknologi, perusahaan media sosial, dan aktor negara dapat mempertahankan kekuasaan hegemonik dan membungkam perbedaan pendapat melalui cara yang halus namun kuat.<sup>33</sup> Lebih jauh lagi, *shadow banning* dapat dimanfaatkan untuk manipulasi opini secara tersembunyi dengan menggeser opini publik secara sistematis sambil tetap terlihat “netral”<sup>34</sup>

Perkembangan AI tidak hanya mengubah cara individu dalam menjalankan kebebasan berekspresi namun membawa dampak serius pada kemampuan media dalam menyebarkan informasi dan hak publik untuk mencari dan memperoleh informasi.<sup>35</sup> Hak ini dapat terhalang dengan penggunaan algoritma untuk penyaringan berita yang secara otomatis dianggap beresiko oleh system AI tanpa proses verifikasi yang transparan dan layak. Jika hal ini terjadi, maka publik kehilangan akses terhadap informasi penting yang seharusnya tersedia bagi semua orang. Situasi ini bertentangan dengan Pasal 19 UDHR, Pasal 19 ICCPR, dan Pasal 28E ayat (3) UUD 1945 di Indonesia yang menjamin hak setiap orang untuk mencari, menerima, dan menyampaikan informasi tanpa gangguan.

#### 4) Otomatisasi dan Rekrutmen Berbasis AI terhadap Hak atas Pekerjaan

Penggunaan kecerdasan buatan dapat mempengaruhi hak atas pekerjaan dengan dua cara utama yakni dengan menciptakan pengangguran struktural karena otomatisasi, dimana AI menggantikan pekerjaan manusia dan dengan penggunaan AI dalam system perekrutan yang melibatkan bias algoritma.

Salah satu tantangan utama yang dihadapi oleh negara berkembang adalah dampak dari otomatisasi terhadap pekerja. AI dapat menggantikan pekerjaan manusia yang dapat berakibat pada peningkatan pengangguran. Penerapan AI berdampak nyata pada penurunan tingkat pekerjaan dan upah, serta menciptakan bentuk pengangguran struktural yang sulit ditanggulangi dalam jangka pendek.<sup>36</sup> Hak atas pekerjaan diakui sebagai salah satu hak fundamental manusia, sebagaimana yang tercantum dalam pasal 23 UDHR dan 6 ICESCR, dimana negara memiliki kewajiban untuk menjamin akses yang adil terhadap pekerjaan yang layak bagi seluruh warganya.<sup>37</sup> Di bawah ICESCR negara tidak hanya memiliki kewajiban untuk

<sup>31</sup>Z. Tufekci, 'Twitter and tear gas: The power and fragility of networked protest', Yale University Press (New Haven: United States of America), 2017. Shadow banning occurs when a social media platform restricts a user's content from showing up without notifying the user, in a hidden and unannounced fashion.

<sup>32</sup>Article 19, *Content Moderation Handbook* (2024), 31, [https://www.article19.org/wp-content/uploads/2024/03/BAHA-SA-Final-SM4P-Content-moderation-handbook-7-Aug-ID-translated-revised-022924\\_YHM.pdf](https://www.article19.org/wp-content/uploads/2024/03/BAHA-SA-Final-SM4P-Content-moderation-handbook-7-Aug-ID-translated-revised-022924_YHM.pdf).

<sup>33</sup>H. Akin Ünver, *Artificial Intelligence and Human Rights: Using AI as a Weapon of Repression and Its Impact on Human Rights* (Brussels: European Parliament, Policy Department for External Relations, Directorate-General for External Policies, May 2024), PE 754.450, 10.

<sup>34</sup>Y. Chen and T. Zaman, "Shaping Opinions in Social Networks with Shadow Banning," *PLOS ONE* 19 (2023), <https://doi.org/10.1371/journal.pone.0299977>

<sup>35</sup>A. Quintavalla and J. Temperman, eds., *Artificial Intelligence and Human Rights* (Oxford: Oxford University Press, 2023), 76–90.

<sup>36</sup>Daron Acemoglu and Pascual Restrepo, "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy* 128, no. 6 (2020): 2188–2244, <https://doi.org/10.1086/705716>.

<sup>37</sup>United Nations, *International Covenant on Economic, Social and Cultural Rights* (ICESCR), December 16, 1966, art. 6.

untuk menahan diri dari melakukan tindakan yang melanggar HAM tetapi mereka juga harus memastikan bahwa hak-hak dasar orang-orang di bawah yurisdiksi mereka tidak dirugikan.

Dalam konteks perekrutan, dampak AI sangat berkaitan erat dengan bias algoritma dan hak atas nondiskriminasi, dimana penerapan system AI dilakukan pada tahap pencarian kandidat yang potensial, melakukan penyaringan CV, dan mencocokkan kandidat dengan posisi yang sesuai.<sup>38</sup> Penggunaan AI dalam system perekrutan pegawai dapat menghemat biaya dan meningkatkan efisiensi dan produktivitas dalam proses pengembangan SDM. Namun, sebagaimana yang sudah dijelaskan di atas, *algorithm-based decision-making* dapat menciptakan keputusan yang bias dan ini juga berlaku pada perekrutan. EU AI Act mengklasifikasikan penggunaan AI dalam sistem rekrutmen sebagai ‘*high-risk*’,<sup>39</sup> dengan alasan potensi bahaya seperti memperkuat pola diskriminasi yang sudah ada serta mengancam hak-hak fundamental pekerja terkait perlindungan data dan privasi.

### C. Kerangka Normatif Internasional terkait AI dan Hak Asasi Manusia

Berbagai dampak yang ditimbulkan oleh penggunaan AI terhadap HAM di atas menunjukkan pentingnya hukum yang jelas untuk mengawal perkembangan teknologi ini. Sifat AI yang kompleks membutuhkan tata kelola yang memastikan penghormatan terhadap martabat dan hak-hak fundamental manusia.

Instrumen hukum HAM internasional yang secara khusus mengatur keterkaitan antara HAM dan AI masih belum ada. Permasalahan seperti perlindungan privasi, kesetaraan, kebebasan berekspresi, serta hak-hak fundamental lainnya memang dapat ditafsirkan dalam kerangka hukum HAM yang sudah ada, seperti *Universal Declaration of Human Rights (UDHR)*, *ICCPR*, dan *ICESCR*. Namun, instrumen-instrumen tersebut belum spesifik menyoal pengambilan keputusan otomatis (*automated decision-making*), profil adaptif (*adaptive profiling*), *deepfakes*, dan opasitas algoritmik (*algorithmic opacity*) yang merupakan permasalahan baru timbul dari sifat kompleks AI.

Tidak adanya norma yang secara eksplisit mengatur aspek-aspek tersebut menimbulkan kesenjangan normatif dalam hukum internasional. Akibatnya, penerapan prinsip-prinsip HAM terhadap AI sering kali bersifat analogis dan reaktif, bukan preventif dan komprehensif, sehingga efektivitas perlindungan terhadap potensi pelanggaran HAM oleh sistem AI masih sangat terbatas. Dalam hal ini, komunitas internasional terus merumuskan berbagai kerangka normatif yang berfungsi sebagai pedoman hukum dan etis untuk menjamin penerapan AI yang sejalan dengan prinsip-prinsip hak asasi manusia.

Bagian ini membahas kerangka tata kelola global yang mengatur perkembangan dan penerapan kecerdasan buatan dari perspektif etika dan hukum. Fokusnya diarahkan pada tiga instrumen utama di tingkat internasional, yakni *UNESCO Recommendation on the Ethics of Artificial Intelligence (2021)*, *OECD AI Principles (2019)*, dan *EU AI Act (2024)*, yang masing-masing memberikan pendekatan berbeda terhadap integrasi prinsip-prinsip hak asasi manusia dalam tata kelola AI.

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

##### a. Substansi dan Prinsip Utama

*UNESCO Recommendation on the Ethics of Artificial Intelligence* merupakan dokumen internasional pertama yang secara komprehensif mengatur terkait prinsip dan juga kebijakan

<sup>38</sup>Pedro R. Palos-Sánchez et al., “Artificial Intelligence and Human Resources Management: A Bibliometric Analysis,” *Applied Artificial Intelligence* 36, no. 1 (2022): 1–28. <https://doi.org/10.1080/08839514.2022.2026420>.

<sup>39</sup>European Parliament and Council, *Regulation (EU) 2024/1689 Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts* (Brussels: European Union, 2024), Annex III, 4(a).

etis dalam pengembangan dan penggunaan kecerdasan buatan, yang tidak mengikat secara umum namun tetap penting secara moral dan politik sebagai *consensus ad idem* dari negara anggota UNESCO. Secara tegas, dokumen ini menyebutkan bahwa HAM merupakan dasar utama dari seluruh prinsip rekomendasinya, yang didasarkan pada kemajuan prinsip-prinsip dasar seperti transparansi dan keadilan, *Unesco recommendation* menegaskan bahwa hak asasi manusia dan kebebasan dasar harus dihormati, dilindungi, dan dipromosikan sepanjang siklus hidup sistem AI<sup>40</sup> yang artinya bahwa HAM termasuk hak privasi, non-diskriminasi, kebebasan berekspresi, dan hak atas pekerjaan harus dihormati, dilindungi, dan dipromosikan dalam seluruh tahap desain, pengembangan, hingga implementasi teknologi AI sesuai dengan hukum internasional.

Adapun yang dimaksud dengan siklus hidup sistem AI mencakup seluruh tahapan mulai dari penelitian, perancangan dan pengembangan, hingga penerapan dan penggunaan, termasuk pemeliharaan, operasional, perdagangan, pembiayaan, pemantauan dan evaluasi, validasi, akhir penggunaan, pembongkaran, serta penghentian.<sup>41</sup> Sedangkan Aktor AI merupakan setiap orang yang terlibat dalam setidaknya satu siklus hidup AI, baik pemerintah, swasta, Masyarakat, organisasi internasional dan academia.<sup>42</sup> Pendekatan berbasis HAM dijabarkan melalui sepuluh prinsip inti (*ten core principles*) yang diantaranya adalah : *Proportionality and Do No Harm; Safety and Security; Fairness and Non-Discrimination; Sustainability; Right to Privacy and Data Protection; Human Oversight and Determination; Transparency and Explainability; Responsibility and Accountability; Awareness and Literacy; dan Multi-stakeholder and Adaptive Governance.*

#### b. Analisis Kritis

*UNESCO Recommendation on the Ethics of Artificial Intelligence* merupakan standar internasional pertama yang memuat prinsip etika AI. Namun demikian, instrument ini secara normative masih bersifat *non-binding* sehingga tidak melahirkan kewajiban hukum bagu negara anggota. Hal ini membuat kekuatan dan efektivitasnya bergantung pada itikad baik masing-masing negara anggota. Selain itu, rekomendasi ini menggambarkan HAM secara keliru sebagai nilai etik daripada kerangka hukum dengan kewajiban hukum, hal-hal yang seharusnya fundamental dan masuk ke dalam ranah *jus cogens* menjadi sekedar pedoman moral sehingga dapat melemahkan posisi instrumen HAM internasional yang bersifat mengikat seperti ICCPR dan ICESCR. Misalnya, hak atas non-diskriminasi dan martabat manusia hanya disebutkan sebagai bagian dari nilai etis yang harus dipromosikan tanpa menjabarkan konsekuensi hukum apabila dilanggar.

Dalam hal implementasi, rekomendasi ini mengalami beberapa kendala karena kurangnya infrastruktur teknologi, regulasi AI, dan tenaga ahli yang dibutuhkan. UNESCO sendiri mengakui tantangan ini dengan mengembangkan *Readiness Assessment Methodology* untuk memetakan kapasitas nasional dalam menerapkan Rekomendasi.<sup>43</sup> Tanpa langkah perlindungan hukum dan kebijakan proaktif oleh pemerintah, potensi bahaya AI tidak dapat dimitigasi di negara-negara berkembang.<sup>44</sup> Kesenjangan kapasitas antara negara yang besar membuat negara maju jauh lebih siap mengadopsi standar UNESCO dibandingkan dengan negara berkembang.

<sup>40</sup>UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), art. 16.

<sup>41</sup>UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), art. 2, 10.

<sup>42</sup>UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), art. 16, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

<sup>43</sup>United Nations Educational, Scientific and Cultural Organization (UNESCO), *A Tool of the Recommendation on the Ethics of Artificial Intelligence: Readiness Assessment Methodology* (Paris: UNESCO, 2023), Foreword, 6–7, <https://doi.org/10.54678/YHAA4429>

<sup>44</sup>Global Index on Responsible AI, “Main Findings,” *Global Index on Responsible AI 2024* (2024), menyimpulkan terdapat kesenjangan besar dalam praktik AI bertanggung jawab dan bahwa mekanisme perlindungan HAM dalam konteks AI masih terbatas, sehingga membutuhkan kebijakan dan pedoman komprehensif berbasis HAM; diakses 16 Oktober 2025, <https://www.global-index.ai/Main-Findings>.

## 2) OECD AI Principles

### a. Substansi dan Prinsip Utama

*OECD AI Principles* merupakan standar prinsip antar pemerintah pertama untuk mempromosikan penggunaan AI yang *innovative and trustworthy*, juga menghormati HAM dan nilai-nilai demokratis. *OECD AI principles* terdiri dari lima prinsip berbasis nilai dan lima rekomendasi kebijakan yang terdiri dari: Pertumbuhan inklusif, pembangunan berkelanjutan, dan kesejahteraan; HAM manusia dan nilai demokrasi (termasuk keadilan dan privasi); Transparansi dan keterjelasan; Ketahanan, keamanan, dan keselamatan; dan Akuntabilitas. Selain itu juga terdapat lima rekomendasi untuk kebijakan utama bagi pemerintah dalam rangka mendukung implementasi AI yang terpercaya antaranya: Investasi dalam riset dan pengembangan AI; Membangun ekosistem AI yang inklusif; Membentuk tata kelola dan kebijakan interoperable; Pengembangan kapasitas manusia dan persiapan pasar kerja; Kerja sama internasional untuk AI terpercaya.

Negara-negara anggota dan negara di luar anggota OECD seperti Argentina, Brazil, Kolombia, dan Singapura telah berkomitmen untuk memasukkan prinsip-prinsip OECD ke dalam kerangka dan strategi AI mereka. Transparansi dan keterjelasan yang menjadi prinsip kunci dalam OECD menandakan bahwa harus ada tingkat keterlibatan manusia dalam pengambilan keputusan otomatis, sebanding dengan dampak potensial dari *output* yang dihasilkan.<sup>45</sup> Prinsip 2 dalam OECD AI Principles mengharuskan aktor AI untuk menghormati hukum, HAM, nilai demokrasi, dan berlandaskan pada *human-centred values* dalam setiap siklus hidup AI. Namun tentunya komitmen ini tidak cukup jika tidak disertai dengan akuntabilitas untuk mencapainya.<sup>46</sup>

### b. Analisis Kritis

Sama seperti *Unesco recommendation*, OECD bersifat *soft law* dan tidak memiliki kekuatan hukum yang mengikat. Sebagai akibatnya, penerapan prinsip dan nilai bersifat sukarela bagi negara dan korporasi, sehingga mereka dimungkinkan untuk menghindari tanggung jawab hukum atas dampak AI terhadap hak asasi manusia. Selain itu, Prinsip nilai dan rekomendasi kebijakan instrument ini masih sangat umum sehingga sulit untuk mengukur kepatuhan aktor AI. Misalnya, dalam penjabaran prinsip “*transparent and explainable*” tidak dijelaskan standar minimum keterbukaan yang dimaksud, apakah mencakup keterbukaan algoritma, data pelatihan, atau proses pengambilan keputusan. Selanjutnya, prinsip yang ada lebih relevan dengan *machine learning* tradisional sebelum *generative AI* berkembang dengan sangat pesat sejak 2022. Sehingga isu-isu baru seperti *deepfake*, disinformasi, dan AI kreatif kurang terakomodasi dalam prinsip-prinsip OECD AI.

Hal lain yang menjadi keterbatasan instrument ini, *OECD AI Principles* dibentuk dari perspektif negara industri dan kurang relevan untuk negara berkembang. Hal ini menyebabkan terjadinya kesenjangan kapasitas dalam pemenuhan prinsip yang memerlukan sumber daya manusia dan infrastruktur yang tinggi. Sebagai contoh, Prinsip-prinsip seperti *transparency*, *accountability*, dan *robustness* memerlukan infrastruktur digital, kapasitas teknis, serta mekanisme pengawasan yang kuat, sesuatu yang belum banyak dimiliki oleh negara-negara di Global South. Dengan hal tersebut di atas, maka OECD lebih tepat menjadi *entry point* menuju regulasi, bukan kerangka yang memadai untuk menjawab tantangan HAM dalam AI.

<sup>45</sup> Organisation for Economic Co-operation and Development (OECD), *AI and the Future of Social Protection in OECD Countries* (OECD Artificial Intelligence Papers, No. 42; Paris: OECD Publishing, June 2025), [https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/ai-and-the-future-of-social-protection-in-oecd-countries\\_038f49ed/7b245f7e-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/ai-and-the-future-of-social-protection-in-oecd-countries_038f49ed/7b245f7e-en.pdf)

<sup>46</sup> Organisation for Economic Co-operation and Development (OECD), *OECD Principles on Artificial Intelligence* (Paris: OECD, 2019), Principle 1.5, <https://oecd.ai/en/ai-principles>.

### 3) EU AI Act (2024)

#### a. Substansi dan Prinsip Utama

Berbeda dengan dua instrumen yang disebutkan di atas, EU AI Act merupakan instrumen yang bersifat mengikat. Regulasi ini merupakan pelopor di dalam mengatur tata kelola AI tingkat internasional yang menetapkan kerangka peraturan dan hukum untuk AI di Uni Eropa namun bersifat ekstrateritorialitas de-facto<sup>47</sup> dengan visi utama untuk pemenuhan HAM dalam rangka menciptakan lingkungan yang kondusif dalam mendorong inovasi berbasis AI yang mencakup semua industri (kecuali militer). EU AI Act mengklasifikasikan AI berdasarkan risiko yang dapat ditimbulkannya, yakni *unacceptable*, *high-level*, *limited* and *minimal*. Semua sistem AI yang dianggap sebagai ancaman terhadap keselamatan, mata pencaharian, dan HAM dilarang. Untuk AI yang berisiko tinggi tunduk pada kewajiban ketat, dan sistem AI yang risiko terbatas mewajibkan transparansi dimana manusia wajib untuk mendapatkan informasi yang diperlukan. Sedangkan untuk sistem AI berisiko minimal atau tidak ada, tidak diatur di dalam UU ini.

Perusahaan yang menerapkan AI di wilayah sensitif harus melakukan pengujian bias, memberikan laporan transparansi, dan memastikan pengawasan manusia dalam pengambilan keputusan berbasis AI. Dimana secara teori, laporan untuk mencegah model AI menghasilkan luaran yang bias seperti melakukan *profiling* yang rasis dalam kepolisian atau bias gender dalam algoritma perekrutan kerja.

Dalam hal kebebasan berekspresi, untuk menjaminkannya maka Undang-Undang AI mengharuskan penyedia sistem AI untuk memberi tahu dan memberikan alasan yang jelas pada pengguna apabila *automated decision* AI memutuskan penghapusan konten. Sedangkan dalam hal perlindungan hak atas informasi, bukan hanya persoalan akses tetapi juga pencegahan terhadap manipulasi<sup>48</sup> Selain itu, hak atas privasi dan non-diskriminasi juga seperti dalam pelarangan *biometric categorisation* terhadap karakteristik dilindungi, dan identifikasi biometrik *real-time* di ruang publik.

#### b. Analisis Kritis

Meskipun instrumen ini menjadi tonggak regulasi mengikat pertama di dunia yang secara khusus mengatur penggunaan AI, terdapat beberapa hal yang masih perlu dikritisi terkait dengan aspek normative dan operasionalnya.

Larangan terhadap *biometric surveillance* dan *emotion recognition* masih memiliki celah hukum. Misalnya, dengan memperbolehkan penggunaannya dalam penegakan hukum tertentu dan migrasi justru dapat melegitimasi praktik pengawasan massal yang infasif.<sup>49</sup> Selain itu, larangan dalam penggunaan *Emotional Recognition Technology (ERT)* hanya bersifat parsial, yakni karena keterbatasan teknis, bukan didasarkan pada alasan etis atau hak asasi manusia.<sup>50</sup> Artinya, apabila teknologi ERT menjadi akurat di masa depan, aturan ini tidak berlaku lagi, meskipun hal ini merupakan invasi terhadap privasi dan otonomi psikologis.

EU AI Act terkait kewajiban transparansi tidak mengharuskan penyedia AI mempublikasikan dataset pelatihan yang memadai bagi publik dan peneliti independen sehingga dapat menjadi *black box*. Hal ini dapat melemahkan pengawasan publik dan hak

<sup>47</sup>European Commission, "Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence: Shaping Europe's Digital Future," *Digital Strategy*, April 21, 2021, archived January 4, 2023, accessed October 6, 2024, <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>.

<sup>48</sup>European Union, *Artificial Intelligence Act*, Regulation (EU) 2024/1689, art. 5.

<sup>49</sup>"EU Reaches Agreement on the Artificial Intelligence Act," ECNL, December 13, 2023, <https://ecn1.org/news/eu-reaches-agreement-artificial-intelligence-act>

<sup>50</sup>Alexandra Prigent, "Is There Not an Obvious Loophole in the AI Act's Ban on Emotion Recognition?," *AI & Society* (2025), <https://doi.org/10.1007/s00146-025-02289-8>

untuk menggugat keputusan otomatis (*contestability*) berbasis bukti menjadi terbatas.<sup>51</sup> Selain itu, jangkauan eksternal dari Uni Eropa justru dapat menjadi tantangan tersendiri. Ketika UE menerapkan aturan yang ketat, pelaku di Global South yang cenderung masih kekurangan sumber daya dapat terbebani dan menimbulkan resiko kolonialitas regulasi, dimana standar hukum di Uni Eropa mendominasi negara-negara berkembang tanpa dukungan sumber daya dan infrastruktur memadai.

## 2. Penanaman Prinsip Hak Asasi Manusia ke dalam Siklus Hidup AI

### A. *AI Lifecycle* dan Penanaman Prinsip HAM

Martabat yang melekat pada setiap manusia adalah dasar bagi sistem hak asasi manusia dan kebebasan fundamental yang universal, tidak dapat dipisahkan, tidak dapat dicabut, saling bergantung, dan saling terkait. Oleh karena itu, penghormatan, perlindungan, dan promosi martabat dan hak asasi manusia sebagaimana ditetapkan oleh hukum internasional, termasuk hukum hak asasi manusia internasional, sangat penting sepanjang siklus hidup sistem AI.

Meskipun *UNESCO Recommendation on the Ethics of AI* menanamkan prinsip dan kewajiban tata kelola di seluruh tahapan daur hidup AI, dokumen ini tidak memberikan definisi daur hidup secara eksplisit, melainkan menggambarkan setiap tahapannya secara luas mengenai rentang kegiatan sepanjang daur hidup sistem AI.<sup>52</sup> Daur hidup AI mencakup tahapan riset, desain, dan pengembangan hingga penerapan dan penggunaan, termasuk pemeliharaan, operasi, perdagangan, pembiayaan, pemantauan & evaluasi, validasi, akhir-pakai, pembongkaran, dan penghentian.<sup>53</sup> Sehingga, seluruh kewajiban etika berlaku *end-to-end*, bukan hanya saat *deployment* sehingga setiap tahap daur hidup AI harus menghormati, melindungi hak asasi manusia.<sup>54</sup> Hal ini menunjukkan bahwa sejak sebuah sistem AI dirancang, dikembangkan, hingga dioperasikan, prinsip-prinsip seperti privasi, non-diskriminasi, transparansi, dan akuntabilitas harus sudah diintegrasikan. Pendekatan inilah yang dikenal sebagai pendekatan berbasis HAM sejak desain (*human rights by design*), analog dengan konsep *privacy by design*, di mana pertimbangan HAM ditanamkan sejak awal dalam arsitektur dan proses pengembangan AI dan dapat diaudit melalui dokumentasi yang memadai serta mekanisme pengawasan independen.<sup>55</sup>

Bertolak dari UDHR, ICCPR, ICESCR, serta kerangka UNESCO 2021, OECD 2019, dan EU AI Act 2024, tulisan ini merumuskan setidaknya enam irisan prinsip yang paling konsisten untuk diintegrasikan sepanjang daur hidup AI.

#### 1) Martabat manusia

Perlindungan terhadap hak asasi dan martabat manusia merupakan fondasi dari seluruh sistem hak asasi manusia yang berlaku sepanjang siklus hidup AI.<sup>56</sup> Dengan prinsip ini, maka para aktor AI berkewajiban untuk menghormati HAM di sepanjang daur hidup AI dengan menempatkan *human-centred values/respect* sebagai prinsip utama<sup>57</sup> Agar implementatif, maka

<sup>51</sup>Marleen van Bekkum, "Using Sensitive Data to De-bias AI Systems: Article 10(5) of the EU AI Act," *Computer Law & Security Review* 56 (2025): 105990, <https://doi.org/10.1016/j.clsr.2025.105990>

<sup>52</sup>United Nations Educational, Scientific and Cultural Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), PDF, accessed October 9, 2025, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

<sup>53</sup>Ibid

<sup>54</sup>UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (adopted 23 Nov 2021), Section II – Aims and objectives, para. (c), p. 15.

<sup>55</sup>Penney, J., McKune, S., Gill, L., & Deibert, R. (2018). Advancing human-rights-by-design in the dual-use technology industry. *Columbia Journal of International Affairs*, 71(2), 103-110. P

<sup>56</sup>Unesco

<sup>57</sup>Organisation for Economic Co-operation and Development (OECD), *Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449)* (Paris: OECD, May 22, 2019; revised 2024), pt. I.2(a) "Human-centred values and fair-

prinsip ini harus diterjemahkan ke dalam persyaratan teknik, proses, dan tata kelola sepanjang siklus hidup AI dan diwujudkan melalui *human rights by design* sehingga praktik-praktik AI tertentu yang secara inheren merendahkan manusia, seperti penilaian risiko kriminal hanya berbasis profil, atau inferensi emosi di tempat kerja dan pendidikan, tidak diperbolehkan<sup>58</sup> dan membutuhkan pengawasan manusia dan transparansi yang lebih kuat sesuai dengan tingkat risiko.<sup>59</sup>

## 2) Legalitas, kebutuhan, dan proporsionalitas.

Setiap penggunaan AI harus punya tujuan yang sah (*lawful*), benar-benar diperlukan (*necessary*) untuk mencapai tujuan itu, dan proporsional (*proportionate*) terhadap potensi dampak pada hak.<sup>60</sup> Prinsip ini berfungsi untuk mencegah perluasan tujuan tanpa dasar (*function creep*) dan agar implementatif, prinsip ini harus diterjemahkan ke dalam kebijakan dan mekanisme pengendalian risiko yang jelas di setiap tahap siklus hidup AI, mulai dari perancangan hingga penggunaan.<sup>61</sup> Oleh sebab itu, “*Proportionality*” dan “*do not harm*” merupakan prinsip inti yang harus dihormati sepanjang siklus hidup AI<sup>62</sup> menyesuaikan tingkat pengamanan dan pengawasan manusia dengan besaran risiko yang mungkin timbul, sehingga sistem AI beroperasi secara sah, perlu, dan seimbang antara manfaat teknologi dan perlindungan hak asasi manusia.<sup>63</sup>

## 3) Kesetaraan & non-diskriminasi.

Prinsip kesetaraan dan non-diskriminasi menegaskan bahwa setiap individu berhak atas perlakuan yang adil tanpa pembedaan berdasarkan identitas mereka termasuk dalam konteks pengguna AI. Dalam praktiknya, penggunaan *big data* pada sistem AI dapat memunculkan diskriminasi,<sup>64</sup> oleh karena itu, untuk menghindari praktik yang merugikan kelompok terdindungi, maka uji bias pra dan pasca “*deployment*”, menggunakan dataset yang representatif, dan penerapan “*fairness metrics*” merupakan kontrol minimum yang harus diterapkan.<sup>65</sup> Karenanya, prinsip *fairness & non-discrimination* merupakan prinsip penting yang harus ditegaskan sebagaimana secara eksplisit dimasukkan dalam *UNESCO recommendation* dan ditegaskan dalam OECD Principles. Hal ini juga diperkuat dalam EU AI act dengan melarang praktik tertentu melalui penggolongan berbagai skenario diskriminatif sebagai *high-risk*. Dengan demikian, pengawasan manusia yang bermakna, transparansi algoritmik, serta akuntabilitas institusional menjadi syarat mutlak untuk memastikan bahwa sistem AI beroperasi secara adil dan tidak memperdalam ketimpangan struktural yang ada.

---

ness,” <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>.

<sup>58</sup>European Union, *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 L 2024/1689 (12 July 2024), art. 5(1)(d), (e), (f), (g) (melarang penilaian risiko kriminal hanya berbasis profil; *untargeted scraping* citra wajah; inferensi emosi di tempat kerja dan pendidikan; serta kategorisasi biometrik atas atribut sensitif), ELI: <http://data.europa.eu/eli/reg/2024/1689/oj>

<sup>59</sup>European Union, *Regulation (EU) 2024/1689 ... (Artificial Intelligence Act)*, art. 14 (“Human oversight”—kewajiban rancangan dan langkah pengawasan manusia untuk sistem berisiko tinggi), ELI: <http://data.europa.eu/eli/reg/2024/1689/oj>

<sup>60</sup>Juri Milaj, “Privacy, Surveillance, and the Proportionality Principle,” *International Review of Law, Computers & Technology* 30, no. 1–2 (2016): 35–50, <https://www.tandfonline.com/doi/abs/10.1080/13600869.2015.1076993>

<sup>61</sup>Bert-Jaap Koops, “The Concept of Function Creep,” *Law, Innovation and Technology* 13, no. 1 (2021): 1–23, <https://www.tandfonline.com/doi/full/10.1080/17579961.2021.1898299>

<sup>62</sup>UNESCO, “Recommendation on the Ethics of Artificial Intelligence,” 2021, bagian “Proportionality and Do No Harm,” <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>

<sup>63</sup>European Union, “Artificial Intelligence Act,” OJ L 2024/1689 (12 July 2024) (kerangka berbasis risiko), ELI: <http://data.europa.eu/eli/reg/2024/1689/oj>

<sup>64</sup>Solon Barocas and Andrew D. Selbst, “Big Data’s Disparate Impact,” *California Law Review* 104, no. 3 (2016): 671–732. <https://doi.org/10.15779/Z38BG31>

<sup>65</sup>Inioluwa Deborah Raji et al., “Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT ’20)* (New York: ACM, 2020), 33–44. <https://doi.org/10.1145/3351095.3372873>

#### 4) Privasi & perlindungan data.

Prinsip privasi dan perlindungan data merupakan pilar utama dalam menjamin penghormatan terhadap otonomi individu di era kecerdasan buatan. Untuk melindungi privasi secara operasional, sistem AI harus mematuhi asas *purpose limitation* dan *data minimization*, serta menerapkan *data/privacy by design and by default*, kontrol desain yang bisa diaudit sebagaimana yang ditegakkan dalam GDPR (Pasal 5 & 25) dan diposisikan UNESCO sebagai kewajiban sepanjang siklus hidup AI (*privacy & data protection*). Secara normatif, *purpose limitation* mewajibkan ditetapkannya tujuan yang spesifik sejak pengumpulan dan melarang pemrosesan lanjutan yang tidak kompatibel dan uji kompatibilitas ketika tujuan berubah di sepanjang lifecycle system.<sup>66</sup> Sementara itu, penerapan *Data minimization* dan (*data protection by design/default*) artinya kewajiban privasi bukan hanya terkait dokumen, tapi harus ditanamkan ke dalam teknologinya.<sup>67</sup>

#### 5) Transparansi & keterjelasan (*explainability*).

Transparansi dan keterjelasan berarti bahwa dokumentasi dan penjelasan yang dapat diaudit, mulai dari tujuan, data, sampai logika model, agar keputusan dapat ditelusuri dan dipertanggungjawabkan.<sup>68</sup> Hal ini merupakan pilar prinsip dengan tujuan Pengguna dan otoritas perlu memahami kemampuan, keterbatasan, dan risiko system. Semakin tinggi level risiko maka semakin tinggi pula transparansi yang harus diaplikasikan dalam sebuah system AI di seluruh tahap lifecycle-nya. Dalam praktik, pendekatan berlapis direkomendasikan dan tingkat transparansi sebaiknya proporsional dengan risiko penggunaan.<sup>69</sup> keterbatasan penjelasan pasca-hoc (berpotensi menyesatkan), sehingga untuk konteks berisiko tinggi lebih aman memilih model yang interpretable-by-design dibandingkan dengan menjelaskan *black box*.<sup>70</sup>

#### 6) Pengawasan Manusia, Akuntabilitas & Remedy

Keputusan AI berdampak signifikan harus dapat ditinjau manusia melalui jalur eskalasi yang jelas (*meaningful human oversight*), dengan rekam jejak audit dan mekanisme keberatan/perbaikan yang efektif.<sup>71</sup> Pengawasan yang bermakna mewajibkan manusia yang berwenang dan memiliki informasi cukup untuk meninjau atau membatalkan keputusan otomatis, sehingga standar HAM tetap menjadi acuan utama dalam setiap tahap pengambilan Keputusan. Akuntabilitas menekankan pentingnya jejak audit menyeluruh, pelacakan asal-usul data, serta audit algoritmik internal maupun independen agar setiap keputusan dapat ditelusuri, dievaluasi, dan diperbaiki secara sistematis sepanjang AI *lifecycle*. Sementara itu, mekanisme remedy harus menyediakan jalur keberatan yang mudah diakses, alasan yang bermakna atas keputusan, dan kesempatan untuk perbaikan atau peninjauan ulang oleh manusia, sebab “hak atas penjelasan” semata belum cukup menjamin adanya kontestabilitas dan pemulihan yang

<sup>66</sup>Bert-Jaap Koops, “The Concept of Function Creep,” *Law, Innovation and Technology* 13, no. 1 (2021): 1–28, <https://doi.org/10.1080/17579961.2021.1898299>

<sup>67</sup>Lee A. Bygrave, “Data Protection by Design and by Default: Deciphering the EU’s Legislative Requirements,” *Oslo Law Review* 1, no. 2 (2017): 105–120, <https://doi.org/10.18261/issn.2387-3299-2017-02-03>.

<sup>68</sup>Riccardo Guidotti et al., “A Survey of Methods for Explaining Black Box Models,” *ACM Computing Surveys* 51, no. 5 (2018): 93:1–93:42, <https://doi.org/10.1145/3236009>

<sup>69</sup>Inioluwa Deborah Raji et al., “Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT '20)* (New York: ACM, 2020), 33–44, <https://doi.org/10.1145/3351095.3372873>

<sup>70</sup>Cynthia Rudin, “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,” *Nature Machine Intelligence* 1 (2019): 206–215, <https://doi.org/10.1038/s42256-019-0048-x>

<sup>71</sup>Inioluwa Deborah Raji et al., “Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT '20)* (New York: ACM, 2020), 33–44, <https://doi.org/10.1145/3351095.3372873>.

efektif bagi individu terdampak.<sup>72</sup> Hak atas penjelasan hanya memberikan informasi terkait logika sebuah keputusan otomatis tanpa menjamin perbaikan terhadap keputusan tersebut, sehingga hak atas penjelasan harus disertai dengan kontesabilitas yang aksesibel yakni berupa hak untuk menggugat dan memperoleh pemulihan.

Prinsip-prinsip tersebut di atas dapat dilihat sebagai prinsip utama yang harus dijadikan sebagai dasar dalam menjalankan setiap siklus hidup AI. Oleh karena itu, tidak cukup hanya sebagai konsep namun juga harus implementatif dalam kebijakan dan regulasi nasional agar perlindungan hak asasi manusia tidak berhenti pada tataran deklaratif saja. Hal ini membuat kebutuhan untuk mengharmonisasikan regulasi nasional dengan standar internasional menjadi sangat mendesak agar tata kelola AI di Indonesia dapat memastikan keseimbangan antara inovasi dan perlindungan hak asasi manusia.

## B. Pentingnya harmonisasi regulasi nasional dengan standar internasional

Berbeda dengan Uni Eropa yang merespon perkembangan AI dengan membuat regulasi yang komprehensif seperti EU AI Act, Indonesia sebagai negara dengan populasi digital keempat terbesar belum memiliki kerangka hukum khusus terkait kecerdasan buatan. Indonesia tetap dalam ruang hampa, hanya mengandalkan panduan etika yang tidak mengikat yang ditetapkan dalam Surat Edaran No. 9/2023 tentang Etika AI.<sup>73</sup> Sedangkan kerangka hukum yang mengikat seperti UU ITE dan UU PDP belum cukup untuk menjawab kompleksitas permasalahan AI yang ada. Dalam situasi seperti ini, Indonesia perlu untuk memperkuat regulasi demi memenuhi kewajibannya dalam memberikan perlindungan dan pemenuhan HAM dalam konteks dampak kemajuan AI. Salah satunya adalah dengan melakukan harmonisasi regulasi nasional dengan standar internasional.

*Unesco Recommendation* menjadi pelopor dalam memberikan arah normatif untuk kebijakan AI pada Tingkat internasional dan menunjukkan pentingnya tata kelola AI yang berakar pada perlindungan martabat dan hak-hak fundamental manusia.<sup>74</sup> Rekomendasi ini tidak hanya menetapkan nilai etis, tetapi juga membentuk kerangka konseptual yang kemudian diadopsi dan diperkuat dalam instrumen lain seperti *OECD AI Principles* dan *EU AI Act*. Dari ketiga instrumen tersebut, dapat dirumuskan sejumlah prinsip hak asasi manusia yang perlu diterapkan secara konsisten, meliputi martabat manusia, legalitas, kebutuhan, dan proporsionalitas, kesetaraan dan non-diskriminasi, privasi dan perlindungan data, transparansi dan keterjelasan, serta pengawasan manusia, akuntabilitas, dan pemulihan (*remedy*). Prinsip-prinsip tersebut harus menjadi fondasi etis yang perlu untuk diintegrasikan ke dalam setiap siklus hidup AI dan diatur di dalam regulasi nasional. Dengan kata lain, Indonesia harus menjadikan HAM sebagai dasar dan pendekatan utama dalam pengembangan kebijakan hukum terkait AI. Jika hukum atau kebijakan yang ada belum cukup mengakomodasi prinsip-prinsip tersebut, maka negara harus mengubah peraturan yang sudah ada, dan/atau menyusun undang-undang baru yang sesuai dengan prinsip etika dan HAM internasional.<sup>75</sup>

Adopsi prinsip *risk-based governance* sebagaimana diatur dalam EU AI Act juga penting untuk diterapkan di Indonesia, mengingat sistem AI berisiko tinggi, seperti yang digunakan dalam sektor kesehatan, keuangan, pendidikan, dan pemilu, memiliki potensi besar untuk melanggar privasi, memperkuat diskriminasi, atau melemahkan akuntabilitas publik jika tidak

<sup>72</sup>Lilian Edwards and Michael Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For," *Duke Law & Technology Review* 16, no. 1 (2017): 18–84, <https://scholarship.law.duke.edu/dltr/vol16/iss1/2/>

<sup>73</sup>Indonesia, *Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial*.

<sup>74</sup>Srabonty Das Gupta, "The EU AI Act and Its Adherence to the European Convention on Human Rights" (EMILDAI, 2025), <https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/>.

<sup>75</sup>UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), para. 9, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>

diatur secara tepat.<sup>76</sup> Dengan begitu, negara juga dimungkinkan untuk dapat mengklasifikasikan sistem AI berdasarkan dampaknya terhadap HAM sehingga bisa memfokuskan regulasi secara proporsional terhadap AI dengan resiko tinggi tanpa menghambat perkembangan dan penggunaan AI yang bermanfaat.

#### D. Kesimpulan

*Artificial Intelligence* tidak hanya menghadirkan manfaat dan efisiensi, tetapi juga tantangan mendasar terhadap tatanan hukum dan hak asasi manusia. Kerangka yang dirumuskan melalui *UNESCO Recommendation on the Ethics of AI (2021)*, *OECD AI Principles (2019)*, dan *EU AI Act (2024)* menunjukkan arah normatif menuju tata kelola berbasis HAM. Namun, ketiganya masih menyisakan kesenjangan normatif dan implementatif. Dalam konteks hukum internasional, integrasi prinsip HAM sepanjang siklus hidup AI merupakan keharusan normatif untuk menjamin agar kemajuan teknologi tidak merusak martabat manusia. Dalam konteks siklus hidup AI, penanaman prinsip HAM sejak tahap perancangan hingga penggunaan akhir (*human rights by design*) sangat penting. Enam prinsip utama yang perlu diintegrasikan adalah: martabat manusia, legalitas/kebutuhan/proporsionalitas, kesetaraan dan non-diskriminasi, privasi dan perlindungan data, transparansi dan keterjelasan, serta pengawasan manusia yang bermakna beserta akuntabilitas. Integrasi prinsip-prinsip ini memastikan bahwa pengembangan dan penerapan AI tetap sah, adil, transparan, dan menghormati hak-hak fundamental manusia. Indonesia, sebagai negara dengan populasi digital besar namun kerangka hukum yang masih terbatas, perlu segera mengharmonisasi regulasi nasional dengan standar internasional. Adopsi pendekatan *risk-based governance*, penerapan prinsip *human rights by design*, serta pembentukan mekanisme *independent AI oversight* merupakan langkah penting menuju tata kelola AI yang adil, transparan, dan menghormati martabat manusia. Dengan demikian, pembangunan AI nasional dapat berlangsung dalam koridor hukum internasional yang menempatkan manusia, bukan algoritma, sebagai pusat dari setiap inovasi.

#### DAFTAR PUSTAKA

- Acemoglu, Daron, and Pascual Restrepo. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy* 128, no. 6 (2020): 2188–2244. <https://doi.org/10.1086/705716>.
- AIMultiple. "Top 18 Predictions from Experts on AI Job Loss." *AIMultiple Research*, updated September 2, 2025. Accessed September 7, 2025. <https://research.aimultiple.com/ai-job-loss/#:~:text=World%20Economic%20Forum%20,2023>.
- Article 19. *Content Moderation Handbook*. March 2024. [https://www.article19.org/wp-content/uploads/2024/03/BAHASA-Final-SM4P-Content-moderation-handbook-7-Aug-ID-translated-revised-022924\\_YHM.pdf](https://www.article19.org/wp-content/uploads/2024/03/BAHASA-Final-SM4P-Content-moderation-handbook-7-Aug-ID-translated-revised-022924_YHM.pdf).
- ArtSmart AI. "AI in Social Media Statistics." Accessed September 7, 2025. <https://artsmart.ai/blog/ai-in-social-media-statistics/#:~:text=%2A%20Over%2080,%28Saufter>.
- Barocas, Solon, and Andrew D. Selbst. "Big Data's Disparate Impact." *California Law Review* 104, no. 3 (2016): 671–732. <https://doi.org/10.15779/Z38BG31>.
- Bommasani, Rishi, et al. *On the Opportunities and Risks of Foundation Models*. Stanford, CA: Stanford Center for Research on Foundation Models, 2021. <https://arxiv.org/abs/2108.07258>.

<sup>76</sup>European Commission, *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (EU AI Act)*, COM(2021) 206 final, Recital 5.

- Buchanan, Bruce G. "A (Very) Brief History of Artificial Intelligence." *AI Magazine* 26, no. 4 (2006): 53–60. <https://www.aaai.org/ojs/index.php/aimagazine/article/view/1868>.
- Buolamwini, Joy, and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 2018.
- Bygrave, Lee A. "Data Protection by Design and by Default: Deciphering the EU's Legislative Requirements." *Oslo Law Review* 1, no. 2 (2017): 105–120. <https://doi.org/10.18261/issn.2387-3299-2017-02-03>.
- Chen, Y., and T. Zaman. "Shaping Opinions in Social Networks with Shadow Banning." *PLOS ONE* 19 (2023). <https://doi.org/10.1371/journal.pone.0299977>.
- Das Gupta, Srabonty. "The EU AI Act and Its Adherence to the European Convention on Human Rights." *EMILDAI*(2025). <https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/>.
- Edwards, Lilian, and Michael Veale. "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For." *Duke Law & Technology Review* 16, no. 1 (2017): 18–84. <https://scholarship.law.duke.edu/dltr/vol16/iss1/2/>.
- European Commission. "Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence: Shaping Europe's Digital Future." *Digital Strategy*, April 21, 2021. Accessed October 6, 2024. <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>.
- . *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (EU AI Act)*, COM(2021) 206 final, Recital 5.
- European Parliament and Council. *Regulation (EU) 2024/1689 Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. Brussels: European Union, 2024. <http://data.europa.eu/eli/reg/2024/1689/oj>.
- European Union. *Artificial Intelligence Act, Regulation (EU) 2024/1689. Official Journal of the European Union L 2024/1689 (12 July 2024)*.
- Executive Office of the President. *Big Data – Seizing Opportunities, Preserving Values*. Washington, DC: The White House, 2014.
- Guidotti, Riccardo, et al. "A Survey of Methods for Explaining Black Box Models." *ACM Computing Surveys* 51, no. 5 (2018): 93:1–93:42. <https://doi.org/10.1145/3236009>.
- Haoyang, Derek Li, and Jason Towne. "How AI and Human Teachers Can Collaborate to Transform Education." *World Economic Forum*, January 9, 2025. Accessed September 7, 2025. <https://www.weforum.org/stories/2025/01/how-ai-and-human-teachers-can-collaborate-to-transform-education>.
- Human Rights Watch. *Meta's Broken Promises: Systemic Censorship of Palestine Content on Instagram and Facebook*. New York: Human Rights Watch, December 21, 2023. <https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content-instagram-and>.
- Indonesia. *Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial*.

- . *Undang-Undang Dasar Negara Republik Indonesia Tahun 1945*. Jakarta: Sekretariat Negara Republik Indonesia, 1945.
- Jordan, Michael I., and Tom M. Mitchell. “Machine Learning: Trends, Perspectives, and Prospects.” *Science* 349, no. 6245 (2015): 255–60. <https://doi.org/10.1126/science.aaa8415>.
- Koops, Bert-Jaap. “The Concept of Function Creep.” *Law, Innovation and Technology* 13, no. 1 (2021): 1–23. <https://doi.org/10.1080/17579961.2021.1898299>.
- Kuye, K. Emmanuel. *Digital Information Privacy in the Age of Big Data: Challenges and Solutions*. 2023. [https://www.researchgate.net/publication/376448455\\_Digital\\_Information\\_Privacy\\_in\\_the\\_Age\\_of\\_Big\\_Data\\_Challenges\\_and\\_Solutions](https://www.researchgate.net/publication/376448455_Digital_Information_Privacy_in_the_Age_of_Big_Data_Challenges_and_Solutions).
- Latonero, Mark. *Governing Artificial Intelligence: Upholding Human Rights & Dignity*. New York: Data & Society Research Institute, 2018. [https://datasociety.net/wp-content/uploads/2018/10/DataSociety\\_Governing\\_Artificial\\_Intelligence\\_Upholding\\_Human\\_Rights.pdf](https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf).
- LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning.” *Nature* 521, no. 7553 (2015): 436–44. <https://doi.org/10.1038/nature14539>.
- McCarthy, John. “What Is Artificial Intelligence?” Stanford University, November 12, 2007. <http://www-formal.stanford.edu/jmc/whatisai.pdf>.
- Milaj, Juri. “Privacy, Surveillance, and the Proportionality Principle.” *International Review of Law, Computers & Technology* 30, nos. 1–2 (2016): 35–50. <https://doi.org/10.1080/13600869.2015.1076993>.
- O’Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown, 2016.
- Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. “Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations.” *Science* 366, no. 6464 (2019): 447–453.
- Organisation for Economic Co-operation and Development (OECD). *AI and the Future of Social Protection in OECD Countries*. OECD Artificial Intelligence Papers, No. 42. Paris: OECD Publishing, June 2025. [https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/ai-and-the-future-of-social-protection-in-oecd-countries\\_038f49ed/7b245f7e-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/ai-and-the-future-of-social-protection-in-oecd-countries_038f49ed/7b245f7e-en.pdf).
- . *OECD Principles on Artificial Intelligence*. Paris: OECD, 2019. <https://oecd.ai/en/ai-principles>.
- . *Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449)*. Paris: OECD, May 22, 2019; revised 2024. <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>.
- Palos-Sánchez, Pedro R., et al. “Artificial Intelligence and Human Resources Management: A Bibliometric Analysis.” *Applied Artificial Intelligence* 36, no. 1 (2022): 1–28. <https://doi.org/10.1080/08839514.2022.2026420>.
- Penney, Jon, Samantha McKune, Lisa Gill, and Ronald Deibert. “Advancing Human-Rights-by-Design in the Dual-Use Technology Industry.” *Columbia Journal of International Affairs* 71, no. 2 (2018): 103–110.

- Quintavalla, Alessandra, and Jeroen Temperman, eds. *Artificial Intelligence and Human Rights*. Oxford: Oxford University Press, 2023.
- Raji, Inioluwa Deborah, et al. "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT '20)*, 33–44. New York: ACM, 2020. <https://doi.org/10.1145/3351095.3372873>.
- Raji, Inioluwa Deborah, et al. "Closing the AI Accountability Gap." In *Proceedings of FAccT '20*, 33–44.
- Rudin, Cynthia. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." *Nature Machine Intelligence* 1 (2019): 206–215. <https://doi.org/10.1038/s42256-019-0048-x>.
- Russell, Stuart, and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 4th ed. Boston: Pearson Education, 2021.
- Toosi, Amirhosein, et al. "A Brief History of AI: How to Prevent Another Winter (A Critical Review)." *PET Clinics* 16, no. 4 (October 2021): 449–469. <https://www.sciencedirect.com/science/article/pii/S1936879821004493>.
- Tufekci, Zeynep. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven: Yale University Press, 2017.
- United Nations. *Universal Declaration of Human Rights*. United Nations General Assembly Resolution 217A (III), December 10, 1948. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>.
- . *International Covenant on Civil and Political Rights*. United Nations General Assembly Resolution 2200A (XXI), December 16, 1966. <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>.
- . *International Covenant on Economic, Social and Cultural Rights (ICESCR)*. December 16, 1966.
- United Nations Educational, Scientific and Cultural Organization (UNESCO). *AI and Education: Guidance for Policy-Makers*. Paris: UNESCO, 2021.
- . *Recommendation on the Ethics of Artificial Intelligence*. Paris: UNESCO, 2021. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.
- . *Recommendation on the Ethics of Artificial Intelligence (Adopted 23 November 2021)*, Section II: Aims and Objectives, para. (c), p. 15.
- . *Recommendation on the Ethics of Artificial Intelligence*, Article 16, "Proportionality and Do No Harm." <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>.
- Ünver, H. Akin. *Artificial Intelligence and Human Rights: Using AI as a Weapon of Repression and Its Impact on Human Rights*. Brussels: European Parliament, Policy Department for External Relations, Directorate-General for External Policies, May 2024. PE 754.450.
- VIDA. *Where's the Fraud: Protecting Indonesian Businesses from AI-Generated Digital Fraud*. Jakarta: VIDA, 2024.
- Vinuesa, Raúl, et al. "The Role of Artificial Intelligence in Achieving the Sustainable

Development Goals.” *Nature Communications* 11, no. 233 (2020). <https://doi.org/10.1038/s41467-019-14108-y>.

World Economic Forum. *The Future of AI-Enabled Health 2025*. SlideShare. [https://www.slideshare.net/slideshow/wef\\_the\\_future\\_of\\_ai\\_enabled\\_health\\_2025-pdf/275311854](https://www.slideshare.net/slideshow/wef_the_future_of_ai_enabled_health_2025-pdf/275311854).

———. “How AI and Human Teachers Can Collaborate to Transform Education.” January 2025. <https://www.weforum.org/stories/2025/01/how-ai-and-human-teachers-can-collaborate-to-transform-education>.