 p-ISSN: 2548-9267 | e-ISSN : 2549-2365, Open Access at : *http://unramlawreview.unram.ac.id/index.php/ulrev*

Gender Discrimination in Artificial Intelligence: An International Human Rights Law Perspective and the Quest for Binding Regulation

Ayu Riska Amalia
University of Mataram
aramel@unram.ac.id

Diva Pitaloka
University of Mataram
divapitaloka@unram.ac.id

Adhitya Nini Rizki Apriliana
University of Mataram
adhityanini@unram.ac.id

ABSTRACT

Artificial intelligence (AI) has transformed and reshaped the way people work and interact. While AI provides convenience, it also poses significant challenges to human rights, particularly gender equality. The use of AI in recruitment processes, healthcare diagnosis, and discriminatory content moderation illustrates how it can exacerbate existing inequalities. This study employs a normative juridical method with a qualitative approach, analysing primary instruments of international human rights law such as the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), the International Covenant on Economic, Social and Cultural Rights (ICESCR), and the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW). It also examines non-binding frameworks, namely the UNESCO Recommendation on the Ethics of AI *and the OECD AI Principles, and compares them with the binding EU AI Act. The findings indicate that AI has the potential to violate fundamental rights of women, including the rights to equality and non-discrimination, work, privacy, health, participation in public and political life, as well as representation and identity. Furthermore, soft-law mechanisms remain insufficient to prevent gender bias, as their implementation relies heavily on states' political will. Nevertheless, states have a positive obligation under international law to respect, protect, and fulfil the right to equality; thus, a binding international legal framework is urgently needed to ensure accountability and gender-sensitive AI governance.*

Keyword : *Gender Discrimination, Artificial Intelligence, Human Rights.*

INTRODUCTION

Artificial Intelligence (AI) has been developing exponentially and has become an integral part of human life in ways previously unimaginable. With the conveniences it offers, AI penetrates almost every aspect of daily activities. Moreover, often without our awareness, AI influences individual decision-making through the content we consume and the advertisements we encounter. This technology has also revolutionized the way people interact with machines

DOI: https://doi.org/10.29303/ulrev.v9i2.450

and enhanced the overall quality of life.[1] The use of AI is not limited to everyday life but has also expanded into more complex domains, including the military sector. The United States,[2] Israel,[3] dan Rusia[4] are among the countries that have incorporated AI as a key component of their military strategies.

Despite the rapid advancement of AI, scholars and researchers continue to hold differing views on its definition, and to date, no universally accepted definition has been established.[5] However, in simple terms, artificial intelligence can be understood as a technology that enables computers and machines to mimic human processes such as learning, understanding, problem-solving, decision-making, creativity, and autonomy. Thus, AI allows machines to replicate a wide range of complex human skills.[6]

Artificial intelligence holds significant potential to accelerate sustainable development and reduce the digital divide. However, its rapid advancement also carries risks that may hinder progress.[7] This duality presents new challenges for society,[8] particularly in relation to human rights Responsible use of AI can bring substantial benefits, such as enhancing access to services and supporting the detection of human rights violations. Conversely, its misuse may harm critical sectors such as justice, healthcare, and education by fostering arbitrary surveillance, information censorship, and discriminatory practices.[9]

Several studies have shown that the output of generative AI has the potential to produce discrimination against women.[10] This is supported by a report of the Committee on Equality and Non-Discrimination of the Council of Europe, which affirms that women and other minority groups experience higher levels of discrimination through the use of AI.[11] Cases include Google's 2015 job ads favouring men for high-paying roles, a 2016 recidivism algorithm labeling Black defendants as higher-risk, and Amazon's 2017 recruitment tool that disadvantaged women. Similar biases led to the suspension of data-profiling by Immigration New Zealand (2018), healthcare allocation tools in U.S. hospitals (2019), and Austria's Public Employment Service recruitment algorithm (2020), all for producing discriminatory outcomes.[12] The various cases above illustrate how AI reflects and exacerbates gender discrimination.

---

[1]Raymond S.T. Lee (2020) Artificial intelligence in daily life. *Singapore: Springer*. https://doi.org/10.1007/978-981-15-7695-9

[2]Bloomberg. (2024, February 29). "Inside Project Maven: The US Military's AI Project". *Bloomberg*. https://www.bloomberg.com/news/newsletters/2024-02-29/inside-project-maven-the-us-military-s-ai-project

[3]Serhan, Y. (2024, December 18). *How Israel uses AI in Gaza—and what it might mean for the future of warfare*. Time. https://time.com/7202584/gaza-ukraine-ai-warfare/?utm_source=com

[4]Bendett, S., Boulègue, M., Connolly, R., Konaev, M., Podvig, P., & Zysk, K. (2021, September). *Advanced military technology in Russia: Capabilities and implications*. Chatham House. https://www.chathamhouse.org/sites/default/files/2021-09/2021-09-23-advanced-military-technology-in-russia-bendett-et-al.pdf

[5]H. Sheikh, C. Prins, & E. Schrijvers. (2023). "Artificial Intelligence: Definition and Background" dalam *Mission AI. Research for Policy*. Cham: Springer, hal. 15. https://doi.org/10.1007/978-3-031-21448-6_2

[6]*Ibid*

[7]USAID. (2024). *AI in Global Development Playbook*. Washington, DC: United States Agency for International Development, hal.4. https://www.usaid.gov/sites/default/files/2024-09/Artificial%20Intelligence%20in%20Global%20Development%20Playbook.pdf

[8]Ünver, H. A. (٢٠٢٤). *Artificial Intelligence (AI) and Human Rights: Using AI as a Weapon of Repression and Its Impact on Human Rights* (PE ٧٥٤,٤٥٠). Policy Department for External Relations, Directorate General for External Policies of the Union, European Parliament, hal. ٨.

[9]U.S. Department of State. (2020). *Guidance on Implementing the UN Guiding Principles for Transactions Linked to Foreign Government End-Users for Products or Services with Surveillance Capabilities*. Washington, DC: U.S. Department of State. https://www.state.gov/wp-content/uploads/2020/10/DRL-Industry-Guidance-Project-FINAL-1-pager-508-1.pdf

[10]Wan, Y., Pu, G., Sun, J., Garimella, A., Chang, K.-W., & Peng, N. (2023). "Kelly is a warm person, Joseph is a role model": Gender biases in LLM-generated reference letters. *arXiv preprint*, arXiv:2310.07371 [v5]. https://doi.org/10.48550/arXiv.2310.09219.

[11]Lacroix, C. (2020, September). *Preventing discrimination caused by the use of artificial intelligence*.Committee on Equality and Non-Discrimination, Socialists, Democrats and Greens Group, Council of Europe. https://assembly.coe.int/LifeRay/EGA/Pdf/TextesProvisoires/2020/20200915-PreventingDiscriminationAI-EN.pdf

[12]*Ibid*

Discrimination and gender bias in AI occur at various stages, including algorithm development, dataset training, and decision-making processes. AI systems operate by using algorithms to transform input data into computational outputs. Consequently, the type and quality of data fed into the system directly influence its subsequent decisions. If the data contains inherent biases, these biases may be replicated by the algorithm and, over time, reinforced in its decision-making processes.[13] If AI is trained on data that associates women and men with different and specific skills or interests, it will generate outputs that reflect those biases.[14] Gender bias in AI can exacerbate existing inequalities and discrimination, leading to unfair decisions, the marginalization of certain groups, and unequal opportunities. Moreover, such bias can reinforce and perpetuate gender norms that constrain societal roles.

The prohibition of gender-based discrimination has been established in various international instruments, most notably the International Bill of Human Rights, which consists of 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). More specific provisions concerning women's rights are set forth in the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW), a convention d that defines gender-based discrimination as any distinction, exclusion, or restriction made on the basis of sex, whether intentional or unintentional, that disadvantages women, hinders society as a whole from recognizing women's rights in both private and public spheres, and prevents women from exercising their human rights and fundamental freedoms to which they are entitled.

To date, no international treaty specifically regulates artificial intelligence (AI) and its impact on human rights. Consequently, existing legal instruments serve as the basis for governing the application of AI to ensure that the technology upholds gender equality and non-discrimination. International initiatives such as the UNESCO Recommendation on the Ethics of Artificial Intelligence (2021) and the OECD AI Principles (2019) have attempted to fill this regulatory gap, but they remain in the form of soft law, whose effectiveness depends heavily on the political will of member states. The adoption of the EU AI Act represents a significant shift from ethical guidelines toward binding normative regulation; however, its application remains limited to the European region

Building on the above background, this study aims to provide an understanding of how AI broadly affects gender equality, as well as how existing international legal frameworks are able to address the challenges of discrimination that arise. It also opens space for discussion on the urgency of establishing binding international regulation.

METHOD

This research employs a normative juridical method with a qualitative approach. The focus is directed at international instruments such as the UDHR, ICCPR, ICESCR, and CEDAW, as well as non-binding instruments such as the UNESCO Recommendation on the Ethics of AI and the OECD AI Principles, which are then compared with the binding EU AI Act. The study relies on a literature review, consisting of primary legal materials (treaties and international conventions), secondary legal materials (books, journal articles, and reports of international institutions), and tertiary legal materials (legal dictionaries, encyclopaedias, and credible

[13]Manasi, A., Panchanadeswaran, S., & Sours, E. (2023, March 17). "Addressing gender bias to achieve ethical AI." *The Global Observatory*. Retrieved from https://theglobalobservatory.org/2023/03/gender-bias-ethical-artificial-intelligence/.

[14]UN Women. (2024, May 22). *Artificial intelligence and gender equality*. UN Women. https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality.

online sources). The analysis is carried out through descriptive-analytical and comparative methods to identify regulatory gaps in international law concerning gender discrimination in AI and provide normative arguments regarding the urgency of establishing binding international regulation.

DISCUSSION AND ANALYSIS

The Impact of Artificial Intelligence on Gender Equality

The Transformation of AI's Role in Social and Professional Life

AI represents the face of the Fourth Industrial Revolution and is considered to have an impact comparable to that of the eighteenth-century Industrial Revolution. Unlike earlier revolutions, however, AI not only transforms modes of production but also fundamentally reshapes how humans interact and relate to one another.[15] AI algorithms embedded in social media, conversational chatbots, and personalization systems are reshaping the ways in which people communicate and access information. Today, social media platforms are accessed by approximately 5.24 billion active users, relying heavily on AI to filter and personalize the content displayed in order to enhance user convenience. However, this personalization also risks creating *filter bubbles* or *echo chambers*, which may narrow users' perspectives as they are continuously exposed to similar content or like-minded opinions**.** [16]

Conversational AI emerges as a new social entity. This AI model serves virtual queries, provides advice to users, and even offers instant emotional support, blurring human-to-human interaction with interaction with computers. Various large companies use this AI for 24/7 services to reduce operational costs, provide fast and accurate responses, and thus increasing customer satisfaction. In addition, this system can personalize services based on user history.[17] The launch of ChatGPT marked the emergence of the most advanced AI chatbot. ChatGPT can understand complex questions and commands, generate coherent natural language responses that feel human-like, and continuously learn from user feedback to improve the quality of conversations and user experience..[18] However, this can also result in biased outputs caused by systematic deviations, attribution errors, or factual distortions that lead to favouritism toward certain groups or ideas, perpetuate stereotypes, or create false assumptions based on learned patterns.[19]

AI is not merely a tool for automation but also drives the transformation of human roles and the emergence of new professions. The *Future of Jobs 2025* report by the World Economic Forum estimates that the adoption of technology (led by AI) will create around 69 million new jobs while disrupting or replacing 83 million jobs by 2027, meaning that one-quarter of current jobs will change within the next five years..[20] Even longer-term projections suggest that 170 million new jobs could be created globally as a result of AI and automation by

[15]Tai, M. C.-T. (2020). "The impact of artificial intelligence on human society and bioethics." *Tzu Chi Medical Journal, 32*(4), 339–343. https://doi.org/10.4103/tcmj.tcmj_71_20.

[16]Chang, J.-P.-C., Cheng, S.-W., Chang, S. M.-J., & Su, K.-P. (2025). *Navigating the Digital Maze: A Review of AI Bias, Social Media, and Mental Health in Generation Z. AI*, *6*(6), 118. https://doi.org/10.3390/ai6060118, hlm. 5.

[17]Mitha Saputri. (2025, April 4). "10 Perusahaan Besar yang Menggunakan Chatbot dan Kegunaannya." *Aptikma Blog*. Diakses dari https://aptikma.co.id/perusahaan-besar-yang-menggunakan-chatbot/.

[18]Jiaxi Liu. (2024, June 18). "ChatGPT: Perspectives from Human–Computer Interaction and Psychology." *Frontiers in Artificial Intelligence, 7*. https://doi.org/10.3389/frai.2024.1418869 , hlm. 1.

[19]Ferrara, E. (2023, April). "Should ChatGPT be biased? Challenges and risks of bias in large language models." *SSRN Preprint*. https://doi.org/10.2139/ssrn.4627814, hlm. 2.

[20]World Economic Forum. (2023). *Future of Jobs Report 2023 – New jobs to emerge, upskilling is key* [video]. Retrieved from https://www.weforum.org/videos/foj-job-market/.

2030, alongside the loss of approximately 92 million existing jobs.[21] This indicates that AI is currently transforming and will continue to reshape the world of work. This transformation not only affects social and economic structures but also carries implications for the fulfillment of human rights, including gender equality.

Mechanisms of the Emergence of Gender Bias in AI

The widespread use of AI often exacerbates existing biases. To understand how AI impacts gender discrimination, it is necessary to identify several key (structural/internal) factors that contribute to the emergence of gender bias in AI.

a. Training Data

AI models learn from the *dataset they're* given. If the training data used contains bias, whether from the source material or the selection process, such bias will be absorbed and reflected in its behavior. If the dataset contains inequality, the model will also produce biased results. As a result, the system tends to reinforce existing injustices.[22] Research has shown that automated translation systems often associate "engineer" or "CEO" with men, while "nurse" or "baker" are associated with women, because the training data used reflects the gender bias present in society.[23] This illustrates the potential dangers when AI is applied, both in generating text and as a component within classification systems.[24]

b. *Measurement bias*

If the training variables do not reflect existing realities, they may reinforce structural discrimination. For example, when measuring crime indicators, using arrest history.[25] Thus, if AI relies on variables and historical data already influenced by gender discrimination, its predictions or decisions will automatically be biased. For instance, if an automated recruitment tool is trained on ten years of job applications dominated by men, the system will teach itself that male candidates are more desirable and downgrade CVs containing the word "women" (such as "president of the women's club") or graduates of women-only colleges, as was the case with Amazon.com Inc.[26] This example shows that when training data reflects past biases (such as more men being hired), AI will evaluate new candidates using the same biased criteria, thereby disadvantaging female applicants.

c. Algorithm

Algorithmic bias arises from the way algorithms process and learn from data. This means that even if the dataset used is relatively neutral, the way the algorithm processes and prioritizes certain features over others can result in discriminatory outcomes.[27] For example, an investigation conducted by *The Guardian* found that in the use of AI for visual content moderation, algorithms tend to be stricter when assessing photos of women. When flagging women's photos, the algorithm applies a lower body threshold, causing content to be labeled

[21]Jessen, J. (2025, January 17). "WEF: AI Will Create and Displace Millions of Jobs." *Sustainability Magazine*. Retrieved from https://sustainabilitymag.com/articles/wef-report-the-impact-of-ai-driving-170m-new-jobs-by-2030

[22]Women in Tech Network. (2025, July 31). *How does gender bias manifest in AI data collection and labeling?*WomenTech. Retrieved from https://www.womentech.net/how-to/how-does-gender-bias-manifest-in-ai-data-collection-and-labeling#:~:text=Stereotypical%20Data%20Representation

[23]Prates, M., Avelar, P., & Lamb, L. C. (n.d.). *Assessing gender bias in machine translation: A case study with Google Translate*. *arXiv preprint*.

[24]Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). "On the dangers of stochastic parrots: Can language models be too big?" In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM. https://s10251.pcdn.co/pdf/2021-bender-parrots.pdf, hlm. 614.

[25]Ferrara, E. (n.d.). *The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness*. *SSRN preprint*.

[26]Jeffrey Dastin. (2018, October 11). "Amazon scraps secret AI recruiting tool that showed bias against women." *Reuters*. Retrieved from https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/

[27]SAP. (n.d.). *What is AI bias? Causes, effects, and mitigation strategies*. Retrieved from https://www.sap.com/resources/what-is-ai-bias

as sexual even when it merely shows a woman wearing a bra, being pregnant, or exercising. As a result, such content may be hidden, and everyday images of women can be censored or automatically have their reach reduced.[28]

d. Labeling and Annotation

Labelling bias emerges because humans label training data, and their subjective decisions influence the model's outcomes.[29] In AI based on Natural Language Processing (NLP),[30] an annotator's identity, such as ethnicity and gender, has the potential to introduce bias into the AI final system, even when the labeling team is demographically diverse.[31] As a result, an AI system trained on such data internalizes gender bias in its predictions and classifications.

e. Product design decisions

If an AI-based health application is designed using a dataset dominated by men, women's symptoms are often misdiagnosed. **This frequently occurs in the field of cardiology**. Heart attacks in women are often underestimated because diagnostic models are more sensitive to male symptom patterns, due to training data that is less representative of women's bodies.[32] Latest study conduct by London School of Economics shows that AI algorithmic used in social service A recent study conducted by the London School of Economics (LSE) showed that AI algorithms used in social services in the UK systematically downplay women's healthcare needs compared to men when summarizing socio-medical records, which risks affecting the amount of healthcare provided.[33] Besides, if the interface design (UI/UX) places greater emphasis on men's health conditions, women's health may be overlooked.[34]

f. Policy decisions

Policies established by AI developer also play a role in shaping gender bias. At the stage of language model development, companies such as OpenAI and Microsoft typically establish policies in the form of guardrails (output restrictions) to prevent harmful, toxic, or potentially abusive content to prevent harmful , toxic, or potentially abusive content. The purpose is indeed to maintain system safety, but these rules can indirectly introduce bias.[35][36] Policies have the potential to generate bias if the rules are made stricter on certain issues. AI may refuse to discuss certain perspectives, ultimately limiting the diversity of viewpoints. In the context of gender, AI content filtering policies can sometimes disproportionately affect women and minorities, for instance, when discussions about experiences of discrimination or gender-based violence are blocked or restricted because they are deemed "sensitive."[37]

---

[28]Gianluca Mauro & Hilke Schellmann. (2023, February 8). "There is no standard: Investigation finds AI algorithms objectify women's bodies." *The Guardian*. Retrieved from https://www.theguardian.com/technology/2023/feb/08/biased-ai-algorithms-racy-women-bodies.

[29]Munro, R., Bethard, S., Kuperman, V., Lai, V. T., Melnick, R., Potts, C., Schnoebelen, T., & Tily, H. (2010). "Crowdsourcing and language studies: The new generation of linguistic data." In *NAACL Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 122–130). Association for Computational Linguistics.

[30]Geva, M., Goldberg, Y., & Berant, J. (2019). *Are we modeling the task or the annotator? An investigation of annotator bias in natural language understanding datasets*. *arXiv preprint* arXiv:1908.07898.

[31]Uncovering labeler bias in machine learning annotation tasks . https://link.springer.com/article/10.1007/s43681-024-00572-w https://link.springer.com/article/10.1007/s43681-024-00572-w

[32]**Mitigation measures for addressing gender bias in artificial intelligence within healthcare settings: a critical area of sociological inquiry**Isaksson A*AI and Society (2025) 40(4) 3009-3018*

[33]Rickman, S. (2025). "Evaluating gender bias in large language models in long-term care." *BMC Medical Informatics and Decision Making, 25*(1), 274.

[34]Benjamin, R. (2020). *Race after technology: Abolitionist tools for the new Jim Code*. Cambridge: Polity Press.

[35]Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning*. *arXiv preprint*arXiv:1702.08608.

[36][Binns, R. (2018). "Fairness in machine learning: Lessons from political philosophy." In *Conference on Fairness, Accountability and Transparency* (pp. 149–159). Proceedings of Machine Learning Research (PMLR).

[37]West, S. M., Whittaker, M., & Crawford, K. (2019). *Discriminating systems: Gender, race and power in AI*. New York: AI Now Institute. Retrieved from https://ainowinstitute.org/publications/discriminating-systems-gender-race-and-power-in-ai-2

Besides those mentioned above, mechanisms of gender bias can also be observed in the interaction between AI and users or the social environment, namely through user feedback and cultural context.

a. **Feedback Loops**

Bias can arise because the output of AI influences the subsequent input into the system, thereby reinforcing existing biases**.** Recommendation systems can exacerbate the marginalization of minority groups by prioritizing engagement-driven algorithms**.**[38] As a result, minority groups and women may become increasingly marginalized because their perspectives are either rarely represented or overshadowed by the dominance of majority content, which can include patriarchal bias or masculine viewpoints. Such data feedback loops often reinforce gender stereotypes. For instance, studies have shown that translation software trained on online texts tends to convert gender-neutral terms (such as "the doctor" or "the nurse") into gendered forms (like *el doctor* for men and *la enfermera* for women), thereby perpetuating the stereotype of male doctors and female nurses.[39]

b. Contextual Bias.

If an AI system is developed without taking cultural context into account, it will lead to cultural bias and difficulties in adapting to local values.[40] This has an impact on gender representation in AI outputs, for example, the biased universal assumption that men are leaders and women are caregivers may be embedded across languages and cultures through AI. For instance, the word *"doctor"* (gender-neutral in English) is often rendered as *"el doctor"* (masculine) and *"nurse"* as *"la enfermera"* (feminine) in Spanish, reflecting cultural assumptions that doctors are male and nurses are female, thereby reinforcing gender stereotypes. Contextual bias may also occur when the data used is not representative, being centered mainly on male faces or majority racial groups while neglecting other demographic contexts. As shown in the **Gender Shades Project**, the study found that commercial facial recognition systems more frequently misclassified women than men, and that these systems performed especially poorly in recognizing dark-skinned women, with error rates as high as **35%**, a stark contrast to the error rate of only **0.8%** for light-skinned men.[41]

Bias in AI can emerge from datasets, algorithms, labeling processes, and even product design policies. When datasets or algorithms reinforce existing inequalities, this can be categorized as **structural discrimination**. The **CEDAW General Recommendation No. 25** obliges states to identify and eliminate indirect discrimination in both public and private spheres, including in the design of technology. Therefore, leaving biased algorithms unchecked by state oversight has the potential to conflict with these international obligations.

The Impact of AI on Women

a. Equal Employment Opportunities

The use of AI in recruitment processes can discriminate against women and widen inequality in employment opportunities. When the data, algorithms, and metrics applied are biased, the decisions produced by the system will also be biased. A well-known case is **Amazon's**

[38]Pagan, N., et al. (n.d.). *A classification of feedback loops and their relation to biases in automated decision-making systems*. *arXiv preprint*.

[39]Smith, G., & Rustagi, I. (2021, March 31). "When Good Algorithms Go Sexist: Why and How to Advance AI Gender Equity." *Stanford Social Innovation Review*. https://ssir.org/articles/entry/when_good_algorithms_go_sexist_why_and_how_to_advance_ai_gender_equity

[40]Tao, Y., et al. (2024, September). "Cultural bias and cultural alignment of large language models." *PNAS Nexus, 3*(9). https://doi.org/10.1093/PNASNEXUS/PGAE346

[41]Smith, G., & Rustagi, I. (2021, March 31). "When Good Algorithms Go Sexist: Why and How to Advance AI Gender Equity." *Stanford Social Innovation Review*. https://ssir.org/articles/entry/when_good_algorithms_go_sexist_why_and_how_to_advance_ai_gender_equity, DOI: 10.48558/a179-b138.

2018 recruitment AI, which was trained on a decade of application histories dominated by male candidates. As a result, the system disadvantaged female applicants and favored male candidates. The algorithm downgraded résumés containing the word *"women's"* and those from women's colleges, thereby reinforcing existing gender inequalities.[42] These findings eventually forced Amazon to terminate the project after realizing that its algorithm was not gender-neutral.[43]

Such bias is becoming increasingly widespread as AI-based recruitment becomes more common. This can be seen from the fact that **99% of Fortune 500 companies** have already adopted automation in their hiring processes.[44] A **2024 study by the University of Washington** found intersectional bias in three AI models used for recruitment, where names with female connotations were chosen far less frequently. These large language models (LLMs) selected male-associated names **85% of the time**, compared to only **11%** for female-associated names. Similarly, a **study by Northeastern University** revealed bias in Facebook job advertisements: technical and forestry positions were more often shown to white men, while cleaning jobs were disproportionately targeted toward Black women.[45]

The above indicates that predictive AI systems can restrict women's access to decent employment. Therefore, **algorithmic audits and regulations** in recruitment are necessary to ensure fairness and to prevent gender-based discrimination.

b. Representation and Identity

The representation of women in media and digital content can be distorted by AI bias in natural language processing (NLP) as well as computer vision. This makes women vulnerable to being placed within traditional stereotypes or even misidentified by automated systems. A **2024 UNESCO study** found that language models display regressive gender-stereotypical patterns: female figures were portrayed **four times more often in domestic roles**, while men were more frequently associated with careers and high-ranking positions.[46] A similar pattern occurs when large language models (LLMs) are asked to generate stories: prestigious professions are often associated with men, while women tend to be given servant roles. NLP systems and machine translation are also prone to bias. For instance, translation systems may alter gender-neutral sentences into sexist ones, as seen in the case of **Google Translate**, which rendered the English sentence *"the man has to clean the kitchen"* into German in a way that implied *"the woman has to clean the kitchen."*[47]

In addition, facial recognition technology raises significant concerns regarding women's identity, particularly for women of color. The **Gender Shades study** revealed a striking accuracy gap: commercial facial recognition systems performed with high reliability for light-skinned men (misidentification rate of around **0.8%**), but their accuracy dropped drastically

[42]Dastin, J. (2018, October 11). "Amazon scraps secret AI recruiting tool that showed bias against women." *Reuters*. Retrieved from https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/.

[43]Dastin, J. (2018, October 11). "Amazon scraps secret AI recruiting tool that showed bias against women." *Reuters*. Retrieved from https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/

[44]Milne, S. (2024, October 31). "AI tools show biases in ranking job applicants' names according to perceived race and gender." *UW News*, University of Washington. Retrieved from https://www.washington.edu/news/2024/10/31/ai-bias-resume-screening-race-gender/

[45]Mello-Klein, C. (2022, October 25). "Facebook's ad delivery algorithm is discriminating based on race, gender and age in photos, Northeastern researchers find." *Northeastern Global News*. Retrieved from https://news.northeastern.edu/2022/10/25/facebook-algorithm-discrimination/

[46]UNESCO. (2024, March 7). *Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes*. Retrieved from https://www.unesco.org/en/articles/generativeaiunescostudyrevealsalarmingevidenceregressivegenderstereotypes.

[47]Kayser-Bril, N. (2021, March 29). "Automated translation is hopelessly sexist, but don't blame the algorithm or the training data." *AlgorithmWatch*. Retrieved from https://algorithmwatch.org/en/automated-translation-sexist/

for dark-skinned women (error rates as high as **34.7%**)[48] This means that women, particularly from minority groups, are often poorly represented in the training data, leading to **contextual bias**, where datasets reflect majority groups while AI systems misrecognize or misidentify them. Such bias poses serious risks, as misidentification can result in false accusations or denial of access. Human rights organizations have warned that this kind of technology effectively **"automates discrimination"** by reinforcing existing racial and gender biases.[49]

Such distortions and failures of identification have serious consequences for gender equality. They can reinforce domestic stereotypes, exclude women from professional spaces, and threaten women's rights to identity and dignity in real life.

c. Privacy and Protection from Digital Violence

The rapid adoption of AI in content moderation, data tracking, and surveillance technologies raises serious concerns about women's privacy and the risks of gender-based digital violence. AI-powered surveillance, for example, can systematically threaten women's privacy. A stark case is Iran, which has deployed aerial drones, CCTV surveillance, and facial recognition technologies in public spaces to identify women who do not comply with hijab laws.[50] Besides, Iran has deployed the **"Nazer" mobile application**, which enables citizens and police officers to report women not wearing the hijab by uploading details such as the location, time, and even the license plate number of the car at the time of the violation. This form of digital repression clearly sacrifices women's privacy and freedom; they are left feeling constantly monitored and intimidated by state-controlled AI systems. This extreme example illustrates that without strict regulation, AI in the hands of authorities can become a tool of social control that discriminates against women, restricting not only their right to privacy but also their freedom of expression.

Advances in generative AI can be exploited to create new forms of digital violence. One example is image-based abuse in the form of deepfake pornography, which has become increasingly widespread as a means of harassing and threatening women. Deepfakes are generated using AI deep learning algorithms, a branch of machine learning that simulates neural networks trained on large-scale datasets to produce fake videos of real individuals. These algorithms are trained to recognize data patterns, human facial movements, and expressions, and can even match voices to mimic a person's authentic speech and gestures. As a result, deepfake technology can convincingly fabricate harmful content that violates women's dignity, autonomy, and safety in the digital space.[51] Such deepfakes are weaponized to shame, discredit, or blackmail women victims, effectively extending the pattern of **revenge porn** through more advanced technology. The traumatic impact of **non-consensual deepfakes** is comparable to, and in many cases compounded by, the harm caused by the distribution of real intimate images without consent. Both forms violate women's dignity, erode their sense of safety, and can cause lasting psychological, social, and professional harm.[52] Numerous cases have already emerged, such as women politicians in various countries facing harassing deepfake attacks designed to challenge, control, and undermine their presence in public life. Clearly,

[48]Buolamwini, J., & Gebru, T. (2018). "Gender Shades: Intersectional accuracy disparities in commercial gender classification." In *Proceedings of Machine Learning Research*, vol. 81, *Conference on Fairness, Accountability, and Transparency*, pp. 1–15. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[49]Fergus, R. (2024, February 29). "Biased Technology: The Automated Discrimination of Facial Recognition." *ACLUMN*. Retrieved from https://www.aclu-mn.org/en/news/biased-technology-automated-discrimination-facial-recognition.

[50]https://www.theguardian.com/global-development/2025/mar/24/iran-police-women-surveillance-hijab-drones-dress-code-law?utm_source=chatgpt.com

[51]https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf

[52]**Artificial Intelligence-Altered Videos (Deepfakes), Image-Based Sexual Abuse, and Data Privacy Concerns**Okolie C*Journal of International Women's Studies (2023) 25(2) 11*

AI in this domain is being misused to amplify the **scale and intensity of digital violence against women**, while moderation and protective measures often lag far behind.

d. Health

In the health sector, AI bias can result in medical services that are less responsive to women's needs. The use of training data and algorithms in healthcare may reinforce inequities, especially when groups such as women and ethnic minorities are underrepresented in the datasets. As a result, AI systems may produce misdiagnoses or downplay women's conditions.[53] Historically, medical research has been centered on the male body aged 23–35 as the reference standard. According to data-driven research from the **University of Leeds** (reported by the **British Heart Foundation**), women are **50% more likely** to receive a misdiagnosis after experiencing a heart attack compared to men. This is due to the fact that many clinical protocols and historical research data have focused on the male body as the diagnostic norm. Consequently, women's atypical heart attack symptoms are often overlooked or misinterpreted.[54]

Digital health applications and AI-based monitoring tools also exhibit bias. **Wearable devices** such as fitness trackers or heart-rate monitors are often calibrated using male physiology (average heart rate, male activity patterns), making the readings less accurate for women. These devices frequently fail to account for hormonal fluctuations, pregnancy, or menstrual cycles.[55] Thus, while AI has great potential to improve women's healthcare, if implemented without addressing social and cultural biases, it risks **exacerbating gender inequality** in the health domain instead of reducing it.[56]

e. Participation in Public and Political Life

The **"Safer Scrolling" (2024) study** by UCL and Kent found that TikTok's algorithm quickly steers users toward increasingly extreme misogynistic content. In an experiment with a teenage male account, within just five days the share of misogynistic videos recommended rose **fourfold**, from 13% to 56%.[57] As a result of this exposure, the report observed that narratives of **toxic masculinity and harassment of women** spilled over into offline interactions (such as in schools), shaping youth culture. This phenomenon demonstrates how **algorithmic filter bubbles** can intensify gender bias—spreading demeaning views of women more rapidly and broadly, and ultimately creating a public discourse climate that is harmful to women.

As a further consequence, women are often intimidated or attacked online, which ultimately limits their involvement in public and political debate. The **United Nations** has noted that online gender-based violence has silenced women's voices in digital spaces and reduced their participation in public life, democratic processes, and leadership positions.[58] For example, women politicians globally face far more intense personal and sexualized attacks than their male counterparts. An analysis of the **2020 U.S. elections** found that female candidates received significantly more online harassment; on **Facebook**, Democratic women candidates were subjected to **10 times more abusive comments** than male candidates from the same

[53]Norori, N., Hu, Q., & Tzovara, A. (2021). "Addressing bias in big data and AI for health care: A call for open science." *Patterns, 2*(10), 100347. https://doi.org/10.1016/j.patter.2021.100347

[54]BHF Press Office. (2016, August 30). "Women are 50 per cent more likely than men to be given incorrect diagnosis following a heart attack." *British Heart Foundation*. Retrieved from https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2016/august/women-are-50-per-cent-more-likely-than-men-to-be-given-incorrect-diagnosis-following-a-heart-attack

[55]Benjeaa, Y., & Geysels, Y. (2020, August 13). "Gender bias in the clinical evaluation of drugs." *Applied Clinical Trials*.

[56]Joshi, A. (2024, October 16). *Big data and AI for gender equality in health: Bias is a big challenge. Frontiers in Big Data, 7*, 1436019. https://doi.org/10.3389/fdata.2024.1436019

[57]Regehr, K., Shaughnessy, C., Zhao, M., & Shaughnessy, N. (2024). *Safer Scrolling: How algorithms popularise and gamify online hate and misogyny for young people*. Association of School and College Leaders. Retrieved from https://www.ascl.org.uk/ASCL/media/ASCL/Help%20and%20advice/Inclusion/Safer-scrolling.pdf

[58]https://unric.org/en/how-technology-facilitated-gender-based-violence-impacts-women-and-girls/

party.[59] A similar trend can be seen in many countries. An **Amnesty International survey in India** found that **1 in 7 tweets** mentioning women politicians was offensive or abusive—with even higher intensity directed at **Muslim women** or those from marginalized castes .[60] Such tactics, including the creation of doctored images or deepfakes that sexualize or discredit women, are intended to block women from holding positions of power by undermining public trust. The combined effect of harassment and algorithms' failure to address it results in **"digital expulsion"**: many women leaders, journalists, and activists become reluctant to speak out, limit their postings, or even withdraw from public office for safety reasons.[61] This represents a profound loss for **democracy and human rights**, as women's perspectives are silenced by fear, diminishing diversity and equality in public discourse.

International Legal Framework for Addressing Gender Discrimination in AI

Normative Basis

The impacts of AI described above demonstrate how AI can exacerbate gender inequality. This situation not only creates social problems but also threatens fundamental rights guaranteed under international law. For instance, the case of **Amazon's recruitment algorithm**, which downgraded résumés mentioning the word *"women's"*, or **Google's ad system**, which more frequently displayed high-paying job opportunities to men, can be categorized as **indirect discrimination** because they create harmful effects for women. Such outcomes stand in conflict with the principle of **substantive equality** guaranteed under **Articles 2 and 3 of the ICCPR and ICESCR** The **International Bill of Human Rights**, comprising 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)*, affirm the principle of equality and non-discrimination based on gender. This principle constitutes a **jus cogens norm** in international law, binding on all states and leaving no room for derogation. [62]

a. The Right to Equality and Non-Discrimination

The opening article of the **Universal Declaration of Human Rights (UDHR)** explicitly affirms that all human beings are born free and equal in dignity and rights. These rights must be enjoyed without any form of discrimination, including on the basis of gender.[63] The same principle is guaranteed in **Articles 2 and 3 of the ICCPR** and **Articles 2 and 3 of the ICESCR**, which place women and men on an equal footing in the enjoyment of all human rights. The **Convention on the Elimination of All Forms of Discrimination against Women (CEDAW)** goes further by specifically obligating states to eliminate all forms of discrimination against women in both public and private spheres.[64] And states are obligated to guarantee women's right to participate in public and political life.[65] In addition, **Article 6 of**

[59]Carnegie Endowment for International Peace. (2020, November 30). *Tackling online abuse and disinformation targeting women in politics*. Retrieved from https://carnegieendowment.org/research/2020/11/tackling-online-abuse-and-disinformation-targeting-women-in-politics

[60]**Amnesty International USA.** (2020, January 23). *New study shows shocking scale of abuse on Twitter against women politicians in India*. Retrieved from https://www.amnestyusa.org/press-releases/shocking-scale-of-abuse-on-twitter-against-women-politicians-in-india/.

[61]**UN Regional Information Centre (UNRIC).** (2023, November 29). *How technology-facilitated gender-based violence impacts women and girls*. Retrieved from https://unric.org/en/how-technology-facilitated-gender-based-violence-impacts-women-and-girls/.

[62]Wallace, R. M. M. (1994). *International law* (2nd ed.). London: Sweet & Maxwell, p. 33.; Hossain, K. (2005). "The concept of jus cogens and the obligation under the U.N. Charter." *Santa Clara Journal of International Law, 3*(1), 73–98.

[63]United Nations. (1948). *Universal Declaration of Human Rights* (UDHR), Article 2.

[64]United Nations. (1979). *Convention on the Elimination of All Forms of Discrimination against Women* (CEDAW), Articles 2 and 3.

[65]United Nations. (1979). *Convention on the Elimination of All Forms of Discrimination against Women* (CEDAW), Article 7.

the ICESCR guarantees the right of every individual to work, while **Article 11 of CEDAW** requires the elimination of discrimination against women in the field of employment.

Thus, algorithmic bias that excludes women from employment opportunities, disregards their health needs, or reinforces gender stereotypes is not merely a technical issue, but also a violation of fundamental human rights norms such as the rights to equality, non-discrimination, and work. A state that allows AI systems to operate without oversight, audits, and regulation fails to meet its obligations to guarantee equality and non-discrimination as mandated by these international instruments. This means that state responsibility extends beyond the level of technology policy, it also requires ensuring that the development of AI does not conflict with the core principles of human dignity and gender equality.

b. The Right to Privacy

The right to privacy is a human right that protects individuals from arbitrary interference with their private life, family, home, correspondence, as well as their reputation and honour. This right is guaranteed in **Article 12 of the UDHR** and later adopted in **Article 17 of the ICCPR**. [66] Privacy does not only mean *"data confidentiality"* but also includes the right to control one's personal information,[67] freedom from excessive surveillance,[68] as well as the guarantee of human dignity.[69] This formulation positions privacy as a protective boundary against the abuse of power, including through technological instruments.

AI in mass surveillance and facial recognition often displays both gender and racial bias. Women are more frequently misidentified, which can lead to wrongful criminalization, harassment, or denial of access to public services. In the context of big data, the right to privacy is also at risk if women's data is processed without consent or used for discriminatory purposes.[70] A state's failure to regulate AI to prevent such risks may amount to a violation of its obligations under the **ICCPR**, specifically **Article 17** on privacy and **Article 26** on equality and non-discrimination.

c. The Right to Participation in Public and Political Life

The right to participate in public and political life plays a vital role in promoting democratic governance, the rule of law, social inclusion, and economic development, as well as in advancing all human rights. Normatively, this right is guaranteed under Article 25 of the ICCPR, which affirms every citizen's right to take part in the conduct of public affairs, to vote in elections, and to access public office without discrimination. Furthermore, Article 7 of CEDAW specifically obliges states to eliminate discriminatory barriers that restrict women's participation in political and public life. However, this right is increasingly threatened by a biased digital ecosystem. Social media algorithms have been shown to amplify misogynistic content, creating online spaces that are hostile to women and even discouraging them from actively voicing their opinions.[71] Gender-based digital violence, such as trolling, doxxing, and deepfakes, restricts women's presence in politics and democratic processes.[72] From the

[66]United Nations. (1948). *Universal Declaration of Human Rights* (UDHR), Article 12. United Nations. (1966). *International Covenant on Civil and Political Rights* (ICCPR), Article 17.

[67]Solove, D. J. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press, pp. 24–25.

[68]United Nations Human Rights Committee. (1988). *General Comment No. 16 on Article 17 (Right to Privacy)*, para. 10.

[69]De Hert, P., & Gutwirth, S. (2006). "Privacy, data protection and law enforcement: Opacity of the individual and transparency of power." In *Privacy and the Criminal Law* (p. 75). Antwerp: Intersentia.

[70]Buolamwini, J., & Gebru, T. (2018). "Gender Shades: Intersectional accuracy disparities in commercial gender classification." In *Proceedings of Machine Learning Research*, vol. 81, *Conference on Fairness, Accountability, and Transparency*, pp. 1–15. Retrieved from https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[71]UN Regional Information Centre (UNRIC). (2023, November 29). *How technology-facilitated gender-based violence impacts women and girls*. Retrieved from https://unric.org/en/how-technology-facilitated-gender-based-violence-impacts-women-and-girls/

[72]UN Regional Information Centre (UNRIC). (2023, November 29). *How technology-facilitated gender-based violence impacts women and girls*. Retrieved from https://unric.org/en/how-technology-facilitated-gender-based-violence-impacts-women-and-girls/

perspective of international law, this phenomenon amounts to a violation of **Article 25 of the ICCPR**, as women's political participation is restricted not by law but by algorithms and unchecked digital violence. It also represents a breach of states' **positive obligations under Article 7 of CEDAW**, since governments have failed to take effective measures to protect women from gender-based barriers in public life.

d. The Right to Health

Article 12 of the International Covenant on Economic, Social, and Cultural Rights (ICESCR) affirms the obligation of States to ensure the highest attainable standard of physical and mental health, including universal access to health care services. Specifically, Article 12 of CEDAW underscores states' duty to eliminate discrimination in the field of health by guaranteeing women equal access to medical services, including those related to reproductive health. However, the development of AI-based medical technologies introduces new risks to the fulfillment of women's right to health.

In reality, many medical algorithms are trained on population data biased toward men, historically rooted in clinical research that has treated the male body as the medical "standard".[73] This bias has serious implications when AI systems fail to recognize women's specific symptoms, for example, in heart attack diagnosis, where studies show that women are "overlooked" **50–60% more often** than men due to differences in clinical manifestations.[74] A similar issue occurs with digital health applications in the form of **wearable devices**, which are often calibrated based on male physiology, leading to less accurate readings for female users.

This situation has the potential to violate the right to health on an equal and non-discriminatory basis as stipulated in the **ICESCR** and **CEDAW**. Women's access to quality healthcare is not equal to that of men, which is inconsistent with **Article 2(2) and Article 12 of the ICESCR** on the obligation of non-discrimination, as well as a violation of states' **positive obligations under Article 12 of CEDAW**, due to their failure to take effective measures to eliminate gender-based barriers in health systems, including those created by technology.

e. The Right to Representation and Identity

In addition to the rights already mentioned, rights related to the use of AI are also guaranteed under Article 6 of the UDHR and Article 16 of the ICCPR, which affirm the recognition of legal identity. Every individual has the right to be acknowledged as a legal person without discrimination, including on the basis of gender, race, or other status. This means that states must ensure that women and men alike enjoy equal legal recognition and protection. Algorithmic bias that results in misidentification or the erasure of gender identity can be interpreted as a violation of the right guaranteed under Article 6 of the UDHR. Everyone, including women and vulnerable groups, must be guaranteed recognition as legal persons. In the context of AI, this requires regulation to ensure that digital systems do not erase, misidentify, or discriminate against women's identities.

The bias identified by **UNESCO (2024)** shows that generative language systems reproduce and reinforce regressive gender stereotypes by placing women in domestic roles while associating men with prestigious careers.[75] This clearly contradicts **Article 6 of the UDHR** and **Article 16 of the ICCPR**, which affirm that every person has the right to be

[73] *Ibid*

[74] BHF Press Office. (2016, August 30). *Women are 50 per cent more likely than men to be given incorrect diagnosis following a heart attack*. *British Heart Foundation*. Retrieved from https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2016/august/women-are-50-per-cent-more-likely-than-men-to-be-given-incorrect-diagnosis-following-a-heart-attack

[75] UNESCO. (2024, March 7). *Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes*. Retrieved from https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes

recognized as an equal legal subject, without being reduced to gender-based stereotypes. [76] The case of **Google Translate** illustrates how NLP systems can transform neutral sentences into sexist ones, further demonstrating how algorithmic bias undermines equal recognition and representation.

Another problem arises with facial recognition technology, which undermines the recognition of women's identities, particularly women of colour. This is largely due to the underrepresentation of minority groups in training datasets, resulting in contextual bias and less accurate facial recognition for women. Such inaccuracies threaten the right to legal identity (Article 6 of the UDHR and Article 16 of the ICCPR), as well as the right to non-discrimination (Articles 2 and 26 of the ICCPR, and CEDAW). This demonstrates how technological bias can directly erode fundamental human rights protections.

Integrating a Gender Perspective into International AI Regulation

Without strong **ethical safeguards**, AI risks reinforcing existing biases and discrimination, particularly against vulnerable groups. For this reason, **UNESCO adopted the Recommendation on the Ethics of AI in November 2021**, which stands as the first global normative instrument on AI ethics.[77] This Recommendation places the protection of **human rights and dignity** as its central foundation, emphasizing core principles such as **transparency, fairness, and the importance of human oversight** in AI systems. At the regional level, regulatory frameworks have also evolved, including the **OECD AI Principles (2019)** and the **EU AI Act (2024)**.

a. UNESCO Recommendation on the Ethics of AI (2021)

UNESCO has taken the lead on the issue of gender and AI through its **Recommendation on the Ethics of AI**, which specifically includes a dedicated policy section on gender. For example, UNESCO emphasizes **digital-era skills and STEM-based curricula** as key mechanisms to advance gender equality.[78] Thus, education is seen not merely as a pathway for individual mobility but also as a **preventive instrument** to address the root causes of women's marginalization in technology. In addition, states are obliged to ensure that **gender equality and women's rights and freedoms** are not violated at any stage of the AI lifecycle. Dedicated budgets must be allocated to **gender-responsive schemes** to guarantee that women are not left behind in the AI-driven digital economy. [79] AI systems must not reproduce gender stereotypes or discriminatory bias. In line with this, states have an active obligation to ensure that AI systems are designed and monitored in such a way that any emerging bias is identified and eliminated.[80]

Beyond technical solutions, achieving gender equality in AI requires structural transformation by creating an inclusive AI ecosystem. This includes fostering workplaces free from harassment, implementing policies that promote diversity throughout the AI lifecycle, and ensuring women's representation in research, academia, and top leadership positions within the AI sector. UNESCO adopts a human rights–based normative approach, guided by principles of fairness, non-discrimination, human oversight, transparency, accountability, as well as awareness and multi-stakeholder engagement.

[76]KayserBril, N. (2021, March 29). *Automated translation is hopelessly sexist, but don't blame the algorithm or the training data*. *AlgorithmWatch*. Retrieved from https://algorithmwatch.org/en/automated-translation-sexist/.

[77]UNESCO. (2022; updated September 26, 2024). *Recommendation on the Ethics of Artificial Intelligence*. Retrieved from https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.

[78]UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*, Policy Area 6: Gender. Member States "shall ensure that the potential of AI systems to contribute to achieving gender equality is fully maximized, and further, they must also ensure that the human rights and fundamental freedoms of girls and women, and their safety and integrity are not violated at any stage of an AI system life cycle" (Para. …). See also UNESCO Women4Ethical AI platform for implementation guidance.

[79]Ibid

[80]Ibid

As an implementation response, UNESCO complements its Recommendation with supporting instruments. The Readiness Assessment Methodology (RAM), for instance, is offered as a tool for states to evaluate their AI governance readiness with a focus on gender and inclusivity. Meanwhile, the Women for Ethical AI (W4EAI) platform was established to expand women's participation, bridge best practices, and strengthen representation within the global AI ecosystem. The presence of these two initiatives marks a shift in UNESCO's role, from merely setting ethical norms to taking concrete action in promoting women's representation in the field of AI.

The **UNESCO Recommendation** represents a progressive step that demands the integration of human rights and gender equality at every stage of the AI lifecycle, offering not only guiding principles but also implementation instruments. However, the **soft law nature** of this recommendation means that its implementation depends entirely on the **political will of member states**, as there is no binding enforcement or monitoring mechanism in place..

b. OECD AI Principles (2019)

The **Organisation for Economic Co-operation and Development (OECD) AI Principles** represent the first global framework widely adopted by developed and emerging countries for the development and deployment of AI, aiming to **maximize its benefits while minimizing its risks**.[81] These principles emphasize five core values: AI should benefit people and the planet (inclusive growth, sustainable development, and well-being); respect human rights, democracy, and the rule of law; ensure transparency and explainability; and promote robustness, security, and accountability.

Unlike UNESCO, which explicitly designates gender as a separate policy area, the principle of inclusive growth, sustainable development, and well-being implicitly requires the integration of a gender perspective. Inclusive growth means that AI must be directed toward supporting inclusive development and reducing inequalities—including gender equality—recognizing this as a prerequisite for sustainable development.

Thus, if AI merely reinforces the gender gap in digital economy participation or reproduces algorithmic bias, it stands in contradiction to the **OECD Principles**. In addition, the OECD emphasizes **fairness and non-discrimination** throughout the AI lifecycle. In other words, every AI actor carries an **ethical and normative responsibility** to ensure that technological innovation does not harm women or other vulnerable groups—whether through biased data, discriminatory algorithmic design, or failures in transparency.

Although the OECD framework is substantively intended for all countries, **most OECD members are developed nations**.[82] As such, its formulation reflects the perspective of states with relatively strong regulatory capacity, advanced digital infrastructure, and more developed AI research and innovation ecosystems. This inevitably makes the OECD standards difficult to apply effectively in countries that still lag behind in these areas.

c. EU AI Act (2024)

The **EU AI Act (Regulation (EU) 2024/1689)** is the first comprehensive legal framework on AI, addressing AI-related risks and positioning Europe to play a leading role globally. The regulation classifies AI systems into four distinct risk levels: **unacceptable, high, limited, and minimal risk**. Each category carries different regulatory obligations and requirements for organizations that develop or deploy AI systems.

The **EU AI Act** does not explicitly set out provisions solely dedicated to gender equality. However, its general principles of **fairness and non-discrimination** implicitly

[81]OECD. (n.d.). *AI principles*. Retrieved from https://www.oecd.org/en/topics/ai-principles.html

[82]World Population Review. (2025). *Organisation for Economic Cooperation (OECD) Countries 2025*. Retrieved from https://worldpopulationreview.com/country-rankings/oecd-countries.

provide protection for gender equality. For example, in its recitals, the Act states that the principles of *"diversity, non-discrimination, and fairness"* require AI to be developed in a way that *"includes diverse actors and promotes equal access, gender equality, and cultural diversity."* This positions the right to gender equality as part of the fundamental rights safeguarded under the regulation[83]

Although gender is not singled out in a standalone provision, discriminatory mechanisms related to gender are addressed within the Act's operational provisions. For instance, under **risk management and data governance requirements**, providers of **high-risk AI systems** are obliged to implement a risk management system throughout the AI lifecycle. This includes a duty to conduct bias analysis of training data that may cause discrimination (including **gender bias**) and to adopt appropriate mitigation measures.[84]

Article 10 stipulates that training datasets must be evaluated for biases that *"are likely to result in negative impacts on fundamental rights or discrimination prohibited by Union law."* It also requires the implementation of *"measures to detect, prevent, and mitigate identified bias."*.[85] This provision implicitly covers gender bias, since discrimination based on gender is prohibited under EU law. Regarding transparency and human oversight, the regulation requires that AI systems clearly inform users when they are interacting with AI, uphold the right to explanation, and involve human oversight to prevent misuse. Although gender is not explicitly mentioned, the principles of fairness and fundamental rights protection (including non-discrimination) are enforced through certification procedures and post-market monitoring. Authorities responsible for safeguarding non-discrimination rights are granted the power to request documentation to ensure that AI systems do not infringe on fundamental rights—for example, provisions referring to *"authorities responsible for the protection of fundamental rights."*[86]

Inclusive Design and Voluntary Codes of Conduct: EU commissioners and AI bodies encourage the development of voluntary codes of conduct that prioritize inclusive design. The Commission's guidelines recommend that such codes include targets for **"inclusive and diverse development teams,"** with attention to **gender balance**, as well as the involvement of a broad range of stakeholders (academia, civil society, and vulnerable groups) in the design process.[87]

The three instruments position gender equality within different frameworks. The UNESCO Recommendation is highly explicit in addressing gender issues, incorporating concrete actions (such as funding, inclusive policies, and educational programs) to empower women in the AI ecosystem. UNESCO's approach is broad and social: AI must actively promote gender equality, eliminate gender bias, and involve women at all levels, research, policy, and practice.

By contrast, the OECD Principles include gender as one of the inequalities that must be reduced. They emphasize the values of fairness and non-discrimination ("equality"), as well as inclusive growth, which explicitly entails reducing gender inequality as part of the broader pursuit of sustainable and equitable development.

These principles function as a guiding framework, applying normative pressure on states and industry to bear social responsibility for AI, but without offering the kind of specific

[83]European Parliament & Council. (2024, June 13). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, Official Journal L 1689, 12 July 2024. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

[84]European Parliament & Council. (2024, April 12). *Artificial Intelligence Act* (Regulation (EU) 2024/1689), *Official Journal of the European Union*, L 1689. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

[85]Ibid

[86]Ibid

[87]ibid

operational instructions that UNESCO provides. In contrast, the EU AI Act (2024), as a binding technical regulation, is focused on preventing all forms of discrimination through compliance obligations such as risk management, data audits, and certification. Here, "gender equality "appears as a general principle in the recitals and in voluntary codes of conduct, but its implementation is embedded within the broader obligations of non-discrimination and bias mitigation.[88] Its approach is **technological and process-oriented**: ensuring representative datasets, fostering diverse development teams (including gender balance), and requiring audits and documentation to prevent bias.

Strengthening Gender-Equality–Based AI Regulation

a. Positive Obligations of States under International Law

The **obligation to fulfil** means that states must take action to facilitate the enjoyment of basic human rights[89]. This implies that the absence of national regulation governing technologies with the potential to violate human rights amounts to a **human rights violation in itself**. International human rights instruments, including **CEDAW General Recommendation No. 28**, stress that states are not only prohibited from engaging in discrimination but are also required to **exercise oversight over the private sector**.[90] In the context of AI, this means that states have a duty to ensure that technology companies do not produce or deploy discriminatory systems. Therefore, AI regulation that exists only in the form of ethical guidelines or **soft law** is insufficient to guarantee the enjoyment of these rights. Binding **accountability mechanisms** are required, covering the entire AI lifecycle—from design to deployment and implementation—to ensure that AI systems align with fundamental principles of equality and non-discrimination.

In terms of AI regulation in Indonesia, the Personal Data Protection Law (UU PDP) and the Electronic Information and Transactions Law (UU ITE) are insufficient to address the complex challenges posed by AI, as these laws were not designed with the development and unique characteristics of AI in mind. While the PDP Law highlights data protection, there are still no specific AI regulations to ensure gender-bias audits or algorithmic transparency. This regulatory gap creates legal uncertainty, leaving room for discriminatory practices and weakening the protection of fundamental rights in the digital and AI ecosystem.

Given these international obligations, states need to develop **AI regulations that mandate audits, reporting, and legal sanctions** for violations of discrimination. Therefore, states must carry out their **positive obligations** by establishing **non-discrimination standards** in AI regulation, for example, mandatory labelling requirements and **gender impact assessments** before a system is deployed.[91] Strengthening law enforcement requires an **independent supervisory body** with the authority to sanction technology companies that violate equality principles. In addition, the legal framework should be complemented with a **dedicated AI law** that regulates comprehensive accountability, **from design to deployment**, in line with technological developments.

b. Global North–South Divide

In global practice, AI regulation is currently led largely by developed countries. In international forums and AI policymaking, the focus is placed on the interests of the **Global**

[88]Ibid

[89]Office of the High Commissioner for Human Rights (OHCHR). (n.d.). *Instruments and mechanisms of international human rights law*. Retrieved from https://www.ohchr.org/en/instruments-and-mechanisms/international-human-rights-law.

[90]United Nations Committee on the Elimination of Discrimination against Women. (2010). *General recommendation No. 28 on the core obligations of States parties under Article 2 of the Convention on the Elimination of All Forms of Discrimination against Women* (CEDAW).

[91]Ibid

North, while the voices and needs of the **Global South** are relatively overlooked.[92] This poses the risk of widening the **AI divide**. It is estimated that only about **3–8% of the global economic benefits of AI** are captured by regions such as Latin America, Africa, and Southeast Asia.[93] Meanwhile, the majority of AI research and capital is concentrated in the **United States, Europe, and China**.[94] The dominance of **Global North perspectives** in AI ethics and regulation can also result in local values in developing countries being left unacknowledged, and may even **exacerbate existing inequalities**.[95]

The **gender data gap** causes AI models to learn from an incomplete picture.[96] For example, a health chatbot project in **Nigeria** initially lacked data on certain groups of women due to the **digital access gap**.[97] Moreover, so-called **"gender-blind" algorithms** often overlook existing inequalities. A study in **Mexico** found that a women-specific credit model (which took gender into account) increased loan approval rates for women compared to a general model that ignored gender data.[98] Without gender awareness, AI can entrench such discriminatory practices. The organization **Women at the Table** emphasizes that if left unchecked, AI will instead create **"automated inequality"** by embedding historical biases into new systems.[99] Conversely, with **inclusive regulation and design**, AI has the potential to help identify gender gaps and address inequalities in areas such as healthcare, education, and financial services..[100]

However, the **unequal economic, infrastructural, and political structures** must be continuously addressed for the Global South to be equally empowered within the global AI ecosystem. Global attention must include the **empowerment of developing countries**, along with the protection of **human rights and gender equality**, to ensure that the benefits of AI can truly be enjoyed **fairly and equitably**.

c. The Need for Binding Regulation at the International Level

The **UNESCO Recommendation on the Ethics of AI (2021)** and the **OECD AI Principles (2019, updated in 2024)** set out values such as **respect for human rights, inclusion, and the reduction of gender inequality**.[101] However, both are **non-binding instruments**: their provisions are *"not mandatory"*, and their effectiveness depends on the goodwill of implementation. As a result, these standards are often not consistently applied in practice. For instance, the OECD calls for **AI accountability**, yet in reality, **algorithmic bias audits are rarely conducted**. At the national level, instruments such as anti-discrimination laws (e.g., the *disparate impact* doctrine) often remain a **patchwork of regulations** and are weakly enforced—making them insufficient to comprehensively address **gender-based algorithmic discrimination**.

On the other hand, the **EU AI Act** serves as a compelling example of adaptation to emerging AI challenges. This regulation sets a precedent for comprehensive AI governance, demonstrating

[92]Tony Blair Institute for Global Change. (2025, February 6). *How Leaders in the Global South Can Devise AI Regulation That Enables Innovation*. Retrieved from https://institute.global/insights/tech-and-digitalisation/how-leaders-in-the-global-south-can-devise-ai-regulation-that-enables-innovation

[93]Unger, N., & McLean, M. (2025, August 13). *An Open Door: AI Innovation in the Global South amid Geostrategic Competition*. Washington, DC: Center for Strategic and International Studies. Retrieved from CSIS website

[94]Ibid

[95]Ibid

[96]Smith, G. (2024, April 3). *How to Make AI Equitable in the Global South*. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/equitable-ai-in-the-global-south.

[97]Ibid

[98]Ibid

[99]Women at the Table. (2025, May 27). *Multilateral Leadership in AI and Gender Equality*. Retrieved from https://www.womenatthetable.net/2025/05/27/multilateral-leadership-in-ai-and-gender-equality/

[100]Ibid

[101]Farhad, S. (2025, May 6). *Passengers in Flight: AI Governance Capacity in the Global South*. *Digital Society, 4*, Article 39. https://doi.org/10.1007/s44206-025-00195-6.

that **binding provisions** on risk management, bias audits, transparency, and human oversight can be formulated in a **technology-neutral** manner while remaining fully **compatible with human rights**.[102] However, its scope remains limited to the European level and it carries no binding force outside the EU. Therefore, there is a pressing need for a **similar binding regulation at the international level**—one that applies to all states

The **Office of the High Commissioner for Human Rights (OHCHR, 2021)** stressed that **high-risk AI practices must be suspended or prohibited until adequate human rights safeguards are in place**, underscoring that ethical standards alone are not sufficient. to guarantee **accountability and non-discrimination**.[103] In addition, from the perspective of international law, the principles of gender equality and non-discrimination are recognized as jus cogens norms, binding on all states. Thus, the failure to regulate AI that risks violating these principles can be interpreted as a form of state negligence in fulfilling its positive obligations, as stipulated under the ICCPR, ICESCR, and CEDAW.

When shaping a new framework for AI governance at the international level that binds all states, it is crucial to actively involve Global South countries and vulnerable communities. This ensures that the resulting regulation is inclusive, reflects local values, addresses the needs of marginalized groups, and incorporates perspectives of digital justice.

Accordingly, the need for binding international regulation is not merely a policy option, but a legal necessity to ensure that AI develops in harmony with the principles of gender equality and human rights protection.

CONCLUSION

Artificial Intelligence not only brings convenience to humans in various aspects of life but also poses challenges to social life. This technological development touches upon the most fundamental principles of human rights, including **gender equality and non-discrimination**. Gender bias and discrimination in AI outputs may arise from training data, algorithms, product design, corporate policies, user feedback, and social contexts rooted in historical and structural bias. This exacerbates existing injustices through **discrimination in recruitment, privacy violations, representation, healthcare services, and even digital surveillance**.

This phenomenon shows that the issue of AI is not simply a technical matter, but also a violation of rights guaranteed under international human rights instruments such as the UDHR, ICCPR, ICESCR, and CEDAW, particularly in relation to the rights to equality and non-discrimination, privacy, work, and public participation. To respond to these challenges, several international regulations have been introduced, such as the UNESCO Recommendation on the Ethics of Artificial Intelligence and the OECD Principles, which remain in the form of soft law. Another example is the EU AI Act, which is legally binding but still limited in scope to Europe. For this reason, and to meet states' positive obligations to prevent, monitor, and address discrimination, there is a clear need for a binding international regulation applicable to all countries—one that ensures accountability and guides AI development in line with the principles of gender equality and the protection of human rights.

Recommendations

To address gender bias and discrimination arising from the use of AI, there is a need for binding national regulations that set clear standards for bias auditing and require AI systems

[102] *The Act Texts*, EU Artificial Intelligence Act website, page providing access to the final version of the EU AI Act published in the Official Journal on 12 July 2024. Retrieved from https://artificialintelligenceact.eu/the-act/

[103] Office of the High Commissioner for Human Rights (OHCHR). (2021). *The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights*. Geneva: United Nations.

to operate transparently. Accountability must be enforced at every stage of the AI lifecycle—from design, data collection and processing, and model development, to implementation and post-deployment evaluation. Equally important is the active involvement of women at every stage of the AI lifecycle to ensure inclusivity and prevent the reproduction of structural bias. At the international level, a binding legal framework should be established to close the current regulatory gap. Such regulation must not be shaped solely by the perspectives and interests of developed countries but should also ensure the active participation of Global South nations, so that the resulting policies are more just, contextual, and applicable.

BIBLIOGRAPHY

Amnesty International USA. (2020, January 23). *New study shows shocking scale of abuse on Twitter against women politicians in India*. Retrieved from https://www.amnestyusa.org/press-releases/shocking-scale-of-abuse-on-twitter-against-women-politicians-in-india/

Benjamin, R. (2020). *Race after technology: Abolitionist tools for the new Jim Code*. Cambridge: Polity Press.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*(pp. 610–623). ACM. https://s10251.pcdn.co/pdf/2021-bender-parrots.pdf

Bendett, S., Boulègue, M., Connolly, R., Konaev, M., Podvig, P., & Zysk, K. (2021, September). *Advanced military technology in Russia: Capabilities and implications*. Chatham House. https://www.chathamhouse.org/sites/default/files/2021-09/2021-09-23-advanced-military-technology-in-russia-bendett-et-al.pdf

BHF Press Office. (2016, August 30). *Women are 50 per cent more likely than men to be given incorrect diagnosis following a heart attack. British Heart Foundation*. Retrieved from https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2016/august/women-are-50-per-cent-more-likely-than-men-to-be-given-incorrect-diagnosis-following-a-heart-attack

Bloomberg. (2024, February 29). Inside Project Maven: The US military's AI project. *Bloomberg*. https://www.bloomberg.com/news/newsletters/2024-02-29/inside-project-maven-the-us-military-s-ai-project

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Proceedings of Machine Learning Research*, vol. 81, *Conference on Fairness, Accountability, and Transparency* (pp. 1–15). Retrieved from https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Carnegie Endowment for International Peace. (2020, November 30). *Tackling online abuse and disinformation targeting women in politics*. Retrieved from https://carnegieendowment.org/research/2020/11/tackling-online-abuse-and-disinformation-targeting-women-in-politics

Chang, J.-P.-C., Cheng, S.-W., Chang, S. M.-J., & Su, K.-P. (2025). Navigating the digital maze: A review of AI bias, social media, and mental health in Generation Z. *AI, 6*(6), 118. https://doi.org/10.3390/ai6060118

De Hert, P., & Gutwirth, S. (2006). Privacy, data protection and law enforcement: Opacity of the individual and transparency of power. In *Privacy and the Criminal Law* (p. 75). Antwerp: Intersentia.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint*, arXiv:1702.08608.

European Parliament & Council. (2024, June 13). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, L 1689 (12 July 2024). Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

Farhad, S. (2025, May 6). Passengers in flight: AI governance capacity in the Global South. *Digital Society, 4*, Article 39. https://doi.org/10.1007/s44206-025-00195-6

Ferrara, E. (2023, April). Should ChatGPT be biased? Challenges and risks of bias in large language models. *SSRN Preprint*. https://doi.org/10.2139/ssrn.4627814

Ferrara, E. (n.d.). The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness. *SSRN preprint*.

Geva, M., Goldberg, Y., & Berant, J. (2019). Are we modeling the task or the annotator? An investigation of annotator bias in natural language understanding datasets. *arXiv preprint*, arXiv:1908.07898.

Hossain, K. (2005). The concept of jus cogens and the obligation under the U.N. Charter. *Santa Clara Journal of International Law, 3*(1), 73–98.

Joshi, A. (2024, October 16). Big data and AI for gender equality in health: Bias is a big challenge. *Frontiers in Big Data, 7*, 1436019. https://doi.org/10.3389/fdata.2024.1436019

Kayser-Bril, N. (2021, March 29). Automated translation is hopelessly sexist, but don't blame the algorithm or the training data. *AlgorithmWatch*. Retrieved from https://algorithmwatch.org/en/automated-translation-sexist/

Lacroix, C. (2020, September). Preventing discrimination caused by the use of artificial intelligence. *Council of Europe, Committee on Equality and Non-Discrimination*. https://assembly.coe.int/LifeRay/EGA/Pdf/TextesProvisoires/2020/20200915-PreventingDiscriminationAI-EN.pdf

Lee, R. S. T. (2020). *Artificial intelligence in daily life*. Singapore: Springer. https://doi.org/10.1007/978-981-15-7695-9

Manasi, A., Panchanadeswaran, S., & Sours, E. (2023, March 17). Addressing gender bias to achieve ethical AI. *The Global Observatory*. Retrieved from https://theglobalobservatory.org/2023/03/gender-bias-ethical-artificial-intelligence/

Munro, R., Bethard, S., Kuperman, V., Lai, V. T., Melnick, R., Potts, C., Schnoebelen, T., & Tily, H. (2010). Crowdsourcing and language studies: The new generation of linguistic data. In *NAACL Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 122–130). Association for Computational Linguistics.

Office of the High Commissioner for Human Rights (OHCHR). (2014). *Women's rights are human rights*. Geneva: United Nations. http://www.ohchr.org/Documents/Events/WHRD/WomenRightsAreHR.pdf

Office of the High Commissioner for Human Rights (OHCHR). (2021). *The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights*. Geneva: United Nations.

Pagan, N., et al. (n.d.). A classification of feedback loops and their relation to biases in automated decision-making systems. *arXiv preprint*.

Prates, M., Avelar, P., & Lamb, L. C. (n.d.). Assessing gender bias in machine translation: A case study with Google Translate. *arXiv preprint*.

Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology, 4*, 100005. https://doi.org/10.1016/j.jrt.2020.100005

Serhan, Y. (2024, December 18). How Israel uses AI in Gaza—and what it might mean for the future of warfare. *Time*. https://time.com/7202584/gaza-ukraine-ai-warfare

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: Definition and background. In *Mission AI. Research for Policy* (p. 15). Cham: Springer. https://doi.org/10.1007/978-3-031-21448-6_2

Smith, G. (2024, April 3). How to make AI equitable in the Global South. *Stanford Social Innovation Review*. https://ssir.org/articles/entry/equitable-ai-in-the-global-south

Solove, D. J. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press.

Tai, M. C.-T. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal, 32*(4), 339–343. https://doi.org/10.4103/tcmj.tcmj_71_20

Tony Blair Institute for Global Change. (2025, February 6). How leaders in the Global South can devise AI regulation that enables innovation. Retrieved from https://institute.global/insights/tech-and-digitalisation/how-leaders-in-the-global-south-can-devise-ai-regulation-that-enables-innovation

United Nations. (1948). *Universal Declaration of Human Rights* (UDHR), Articles 2 & 12.

United Nations. (1966). *International Covenant on Civil and Political Rights* (ICCPR), Article 17.

United Nations. (1979). *Convention on the Elimination of All Forms of Discrimination against Women* (CEDAW), Articles 2, 3, & 7.

United Nations Committee on the Elimination of Discrimination against Women. (2010). *General recommendation No. 28 on the core obligations of States parties under Article 2 of the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW)*.

United Nations Human Rights Committee. (1988). *General Comment No. 16 on Article 17 (Right to Privacy)*, para. 10.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. Paris: UNESCO.

UNESCO. (2024, March 7). Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes. Retrieved from https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes

UN Women. (2024, May 22). *Artificial intelligence and gender equality*. Retrieved from https://

www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality

Unger, N., & McLean, M. (2025, August 13). *An open door: AI innovation in the Global South amid geostrategic competition*. Washington, DC: Center for Strategic and International Studies.

Ünver, H. A. (2024). *Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights (PE 754.450)*. Policy Department for External Relations, Directorate General for External Policies of the Union, European Parliament.

USAID. (2024). *AI in global development playbook*. Washington, DC: United States Agency for International Development. https://www.usaid.gov/sites/default/files/2024-09/Artificial%20Intelligence%20in%20Global%20Development%20Playbook.pdf

Wallace, R. M. M. (1994). *International law* (2nd ed.). London: Sweet & Maxwell.

Wan, Y., Pu, G., Sun, J., Garimella, A., Chang, K.-W., & Peng, N. (2023). "Kelly is a warm person, Joseph is a role model": Gender biases in LLM-generated reference letters. *arXiv preprint*, arXiv:2310.07371 [v5]. https://doi.org/10.48550/arXiv.2310.09219

West, S. M., Whittaker, M., & Crawford, K. (2019). *Discriminating systems: Gender, race and power in AI*. New York: AI Now Institute. Retrieved from https://ainowinstitute.org/publications/discriminating-systems-gender-race-and-power-in-ai-2

Women at the Table. (2025, May 27). Multilateral leadership in AI and gender equality. Retrieved from https://www.womenatthetable.net/2025/05/27/multilateral-leadership-in-ai-and-gender-equality/

World Economic Forum. (2023). *Future of jobs report 2023 – New jobs to emerge, upskilling is key* [video]. Retrieved from https://www.weforum.org/videos/foj-job-market

World Population Review. (2025). Organisation for Economic Cooperation (OECD) countries 2025. Retrieved from https://worldpopulationreview.com/country-rankings/oecd-countries