

# Human Rights Council – Research Report II

## Advancing a human rights-based approach to AI

### Introduction to the topic:

Artificial Intelligence (AI) has become a transformative force in our society, driving advancements in several industries worldwide. From automating everyday tasks to enabling advancements in healthcare, AI holds the potential to address several global challenges. However, its rapid development raises significant human rights concerns.

These concerns are vast, ranging from disinformation and infringements on privacy to algorithmic bias and systemic discrimination. The table below presents several AI legal issues and the corresponding human rights principles affected.

**Table 1**  
Issues and affected human rights

AI legal issue	Human rights principles that might be affected
Lack of algorithmic transparency	fair trial and due process; effective remedies; social rights and access to public services; rights to free elections
Cybersecurity vulnerabilities	the right to privacy; freedom of expression and the free flow of information
Unfairness, bias and discrimination	elimination of all forms of discrimination against women; equal rights of men and women; enjoyment of children's rights without discrimination; equality before the law, equal protection of the law without discrimination; enjoyment of prescribed rights without discrimination; non-discrimination, right to life of migrant workers; right to liberty and security of the person; prohibition of discrimination on the basis of disability; right to fair trial; right to freedom from discrimination
Lack of contestability	right to an effective remedy; access to justice
Legal personhood, subjecthood, moral agency	right to recognition everywhere as a person before the law; right to equality; elimination of all forms of discrimination
Intellectual property issues	right to own property alone or in association with others; right to freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits; right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which s/he is the author.
Adverse effects on workers	right to social security; prohibition of discrimination in relation to the enjoyment of rights to work, to free choice of employment, to just and favourable conditions of work, to protection against unemployment, to equal pay for equal work, to just and favourable remuneration; right to work, including the right of everyone to the opportunity to gain his living by work which s/he freely chooses or accepts); right of persons with disabilities to work, on an equal basis with others
Privacy and data protection issues	migrant's right to privacy; respect for privacy of person with disabilities; right to respect for private and family life; right to privacy and data protection; children's privacy; protection of the integrity of older persons and their privacy and intimacy
Liability issues related to damage caused	right to life; right to effective remedies
Lack of accountability for harms	right to life; right to effective remedies

**Figure 1 - Table comparing AI legal issues to specific human rights principles affected (Journal of Responsible Technology, 2020)**

As highlighted in figure 1, one of the key challenges in AI is the issue of unfairness, bias, and discrimination. An illustrative example is facial recognition technology (a technology that uses AI

to recognize human faces), which has been shown to exhibit significant bias towards minority groups. The “Gender Shades” study by MIT Media Lab (2018) revealed that facial recognition systems from leading tech companies had error rates as high as 34.7% for darker-skinned women compared to only 0.8% for lighter-skinned men. Biases of this nature have severe real-world implications. In the U.S, for example, there have been at least seven confirmed cases of misidentification due to facial recognition technology, six of which involved Black people who have been wrongfully accused. The origins of such bias often lie in the datasets used to train these systems (algorithmic bias). When training data disproportionately represents certain demographics or excludes others, the resulting models inherit and amplify these imbalances. Consequently, societal biases end up being embedded in the data itself, reflecting historical inequities and systemic exclusion.

Another significant concern is the negative impact of AI on workers, particularly in areas such as recruitment bias and job opportunities. A 2023 study on *Ethics and discrimination in artificial intelligence-enabled recruitment practices* concluded that AI-systems used to recruit employees showed bias in gender, race, color and personality. According to the study, these biases stem from historical data and the personal preferences of algorithm engineers. A practical example is Amazon’s experience in 2018, when the company shut down its AI recruiting tool after discovering it discriminated against women. Designed to search the web and identify potential candidates, the algorithm systematically downgraded women’s CV’s for technical positions. This shows how unchecked AI systems can perpetuate discrimination, violating fundamental human rights to equality and fair treatment in the workplace.

Beyond recruitment, AI also threatens millions of jobs worldwide. According to the World Economic Forum, 85 million jobs are projected to be replaced by AI. While this displacement is significant, the same report predicts that 97 million new roles could emerge through automation and augmentation. However, these newly created opportunities often require advanced skills and training, leaving many displaced workers at risk of being excluded from the evolving job market. Hence, without necessary adjustments to education and workforce development, this shift could lead to widespread unemployment, increased income inequality and diminishing workers’ rights.

Moving beyond the consequences of this issue, it is also crucial to understand its underlying causes. One major factor is the policy and its legal shortfall. The rapid advancement of AI technologies has outpaced the development of legal frameworks, resulting in a fragmented

regulatory environment. Existing regulations, such as the European Union's Artificial Intelligence Act (AI Act), aim to address these challenges but still have notable limitations. Critics argue that the AI Act's risk-based approach may not adequately cover all AI applications, leaving certain high-risk systems (such as AI's use in media and academia) under-regulated. Additionally, the Act's broad definitions, exceptions and complex compliance requirements can create legal ambiguities, potentially preventing effective enforcement. Another concern is the fact that the act only applies to members of the EU, so, even it were fully effective, it would still only solve the issue in 27 countries out of the entire world.

Another problem is the lack of alignment between different stakeholders. Different governments, private companies and other entities all bring different priorities to the table, resulting in conflicting goals and regulatory gaps. A notable comparison is the United States vs. the European approach to AI regulation. While both agree on the conceptual basis of AI regulation, the European Union adopts a stricter, centralized approach, while the U.S follows a decentralized strategy, relying on individual agencies and states to develop specific guidelines and penalties, with no overarching federal framework. Corporate interests add to this conflict of interests due to the profit-driven priorities of tech companies, which often overshadow ethical considerations. In October 2024, for example, Google, Meta, X and LinkedIn were compelled to pause or delay AI projects within the EU due to regulatory intervention. This regulatory pressure reignited tensions between Big Tech and European authorities, highlighting the ongoing conflict between rapid AI advancement and compliance with regulations.

Technical shortfalls also play an important role, as the design and development of technology lay the groundwork for addressing legal and ethical challenges. Creating well-designed AI systems from the beginning could significantly reduce the occurrence of biases, errors, and ethical problems.

Ultimately, without effective regulation, AI poses significant risks to human rights on a global scale. Special attention must be paid to vulnerable and underrepresented groups, particularly in developing nations, which often lack the resources to regulate or benefit from AI equitably. This leaves their populations disproportionately affected by AI misuse. To address these challenges, capacity-building initiatives, such as providing technical support and fostering international cooperation, are essential. As emphasized by organizations like UNESCO, ensuring equitable

access to AI while safeguarding marginalized communities is crucial to harnessing its potential for the benefit of everyone.

## **Background information:**

The origins of AI go back to the mid-20th century, when scientists began exploring the concept of machine learning. Alan Turing, a British mathematician and logician often regarded as the father of AI, introduced several foundational concepts, including the Turing Test, in 1950, designed to evaluate a machine's ability to mimic human behaviour. The formalization of AI as a field occurred at the Dartmouth Conference, in 1956, where the term "artificial intelligence" was first used. It was also in the 1950s that the first AI program was developed. However, the field faced significant challenges during the "AI Winters" of the 1970s and 1980s due to technological limitations and reduced funding. Nevertheless, by the 1990s, AI experienced a revival driven by advances in computational power and algorithm development. A significant milestone was IBM's Deep Blue, a program which made headlines, in 1997, for defeating world chess champion Garry Kasparov, showcasing AI's ability to tackle complex and human-dominated tasks.

From the 2000s to today, the field of AI has witnessed several breakthroughs, from deep learning techniques to transformers. One of the most significant developments has been the introduction of generative models, such as OpenAI's ChatGPT, in 2022. Using transformers to develop large language models, ChatGPT made conversational AI accessible to everyone.

As AI advances, regulatory frameworks have emerged to address ethical and societal concerns. Common examples include the EU's AI Act (2021), the OECD AI Principles and UNESCO's Ethical AI guidelines.

## **Recent developments:**

Recent years have seen significant progress in adopting human rights principles into AI governance, with key developments worldwide. The Organization for Economic Co-operation and Development (OECD) has played a foundational role, establishing its AI Principles in 2019, which

were updated in May 2024. These principles provide a strong framework for countries to shape their AI policies and foster global interoperability. Notably, the OECD's definition of an AI system – "a machine-based system that can, for a given set of objectives, make predictions, recommendations, or decisions influencing real or virtual environments" – is still used by the United Nations and other diplomatic bodies as a cornerstone in legislative and regulatory actions.

The European Union (EU) has also helped align AI development with human rights through its Artificial Intelligence Act (2021). This regulatory framework categorizes AI systems by risk level, ensuring scrutiny for high-risk applications such as biometric surveillance and decision-making in critical sectors. By prioritizing the protection of fundamental rights, the EU has set a global benchmark for ethical AI governance.

In addition to broad regulatory frameworks, targeted initiatives have addressed sector-specific applications of AI. For instance, the use of AI in prison and probation services has come under scrutiny due to its potential to infringe on human rights. Predictive analytics, commonly employed for risk assessments and parole decisions, offer operational efficiency but raise ethical concerns. Recognizing these challenges, the Council of Europe issued recommendations, in late 2024, to its 46 member states. These guidelines emphasize transparency, compliance with international legal standards, and a focus on rehabilitation and reintegration to reduce recidivism. This initiative, grounded in frameworks like the Council of Europe Framework Convention on Artificial Intelligence, highlights the importance of ethical oversight when deploying AI in justice systems.

These developments demonstrate growing international involvement in developing human rights-based AI policies.

### **Focus of the debate:**

The focus of the debate will be to develop guidelines and strategies to tackle the violations of human rights by AI. Delegates should aim to balance the diverse interests of different stakeholders with the need to uphold human rights. Additionally, delegates are encouraged to explore ways of providing restitution or "repayment" to individuals affected by AI-related human rights infringements. Beyond addressing these challenges, delegates should also consider how AI

can be proactively used to promote and advance human rights. Ultimately, a realistic and sustainable consensus should be reached so as to ensure long-term protection and promotion of human rights in an ever-evolving industry.

### **Possible solutions:**

- Provide funding and technical assistance to developing nations to enhance their regulatory capacity and ensure equitable access to AI benefits
- Develop reskilling programs to help displaced workers transition into AI-driven roles
- Promote the use of AI tools to enhance accessibility and opportunities for marginalized populations, such as the development of assistive technologies for people with disabilities
- Establish human rights impact assessments for AI systems before deployment

### **Significant Parties:**

**European Union (EU):** The EU has positioned itself as a global leader in ethical AI governance. It develops and enforces regulations such as the General Data Protection Regulation (GDPR) and the proposed AI Act, which incorporates human rights considerations. By promoting "trustworthy AI," the EU influences both its member states and international discussions on AI policy.

**United States:** A major global player in AI development, the U.S. emphasizes innovation and competitiveness. Through initiatives like the National AI Initiative Act of 2021, it aims to foster collaboration among government, academia, and industry to align AI technologies with democratic values and human rights. However, criticism persists over the use of AI in areas like law enforcement and immigration, raising concerns about bias and discrimination.

**China:** As a technological powerhouse, China's AI policies focus on national security, economic growth, and technological advancement. However, its use of AI for surveillance and control — particularly towards the Uighur Muslim minority in Xinjiang — has raised significant international human rights concerns.

**Africa:** Many African nations are developing AI strategies within the context of the African Union's Data Policy Framework and Guidelines (2022). These efforts aim to ensure equitable

access to AI technologies while addressing regional human rights challenges, such as digital inclusion and data privacy.

**India:** India has made significant strides in leveraging AI to tackle pressing issues in healthcare, agriculture, and education, while also expanding access to AI technologies across its vast territory. However, its deployment of AI at borders, particularly in biometric screening and traveler profiling, has sparked concerns about ethical violations. Statements emphasizing a "zero-tolerance policy" from border authorities underscore the potential risks of AI misuse, drawing international scrutiny.

### **Past UN Actions:**

**-First United Nations HRC Resolution on the Right to Privacy in the Digital Age (2013):** This resolution emphasized the importance of protecting privacy amidst technological advancements. Subsequent resolutions have reiterated the need to safeguard human rights in the context of AI and other technologies that pose risks to privacy and freedom of expression.

**-International Telecommunication Union (ITU) AI for Good Global Summit (2017):** The ITU launched the AI for Good Global Summit to address the ethical implications of AI. This annual event gathers stakeholders from the UN, governments, academia, and the private sector to ensure that AI supports the public interest and aligns with human rights while advancing the UN's Sustainable Development Goals (SDGs)

**-Human Rights Report on the Impact of AI on Human Rights (2018):** The report highlighted the risks posed by AI technologies to human rights and proposed that AI systems should be designed and implemented to respect dignity, avoid discrimination, and safeguard privacy and freedom of expression.

**-UN Global Compact's Framework for AI and Human Rights (2019):** Framework calling for businesses to ensure AI technologies align with human rights principles. This framework urges companies to develop AI that respects privacy, equality, and inclusivity.

**-Global Partnership on AI, GPAI (2020):** International initiative that brings together governments, the private sector, and civil society to ensure that AI is developed and deployed in a manner that is inclusive, accountable, and aligned with human rights principles.

**-World Health Organization Digital Health Strategy (2020):** The WHO released the Global Strategy on Digital Health, which stresses the importance of AI in healthcare, while emphasizing the need to respect human rights, including privacy and non-discrimination. This strategy outlines how AI should be applied in the health sector with a focus on equity and ethical principles.

**-UN Secretary-General's Roadmap for Digital Cooperation (2020):** The roadmap emphasizes the importance of a human-centered approach to AI governance, ensuring that AI technologies are used for the public good and aligned with human rights principles such as non-discrimination and privacy protection.

**-UN Human Rights Council Resolution on the Impact of AI and Human Rights (2022):** Resolution focused specifically on the impact of AI on human rights. It urged member states to ensure that AI technologies are developed and deployed in ways that respect international human rights law, particularly regarding privacy, freedom of expression, and equality. It also called for the creation of a special rapporteur to monitor AI's impact on human rights.

**-UN Secretary-General's Digital Cooperation Report (2023):** The report emphasized the need for equitable access to AI technologies and for ensuring that these technologies respect the rights of individuals, particularly marginalized communities, while promoting digital inclusivity and addressing biases.

## **Glossary and Key Terms:**

**Generative AI:** AI systems, such as ChatGPT, that use machine learning to create new content, including text, images, and audio, based on input data.

**Artificial Intelligence (AI):** A machine-based system capable of performing tasks that typically require human intelligence, such as decision-making, problem-solving, and language understanding.



**Capacity Building:** Efforts to strengthen the abilities of countries, organizations, or communities to regulate, develop, and benefit from AI equitably.

**Machine Learning:** A subset of AI that enables systems to learn and improve from data without explicit programming.

**Deep Learning:** An advanced machine learning technique using neural networks with multiple layers to process complex data and learn patterns.

**Interoperability:** Refers to the ability of systems, organizations, or technologies to work together seamlessly and effectively.

**Recidivism:** The tendency of a convicted individual to re-offend or relapse into criminal behavior after having been released from incarceration, probation, or rehabilitation.

**High-Risk AI Systems:** AI applications with significant potential to impact human rights, safety, or critical societal functions, such as facial recognition or healthcare decision-making tools.

## **Bibliography & Useful Links:**

AIPRM. (n.d.). 50+ AI Replacing Jobs Statistics 2024. <https://www.aiprm.com/ai-replacing-jobs-statistics/#:~:text=The%20World%20Economic%20Forum%20has,joint%20trend%20'automatio n%20and%20augmentation>

365 Data Science. (2024, October 4). *A Brief History of AI*. YouTube. <https://www.youtube.com/watch?v=eSj80Zr6TEE>

Baweja, S., & Singh, S. (2020, July 16). *Beginning of Artificial Intelligence, End of Human Rights*. LSE Human Rights. <https://blogs.lse.ac.uk/humanrights/2020/07/16/beginning-of-artificial-intelligence-end-of-human-rights/>

Buolamwini, J. (2017). *Gender Shades*. Gendershades.org. <http://gendershades.org/overview.html>

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, 81(81), 1–15. <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanit Soc Sci Commun* 10, 567. <https://doi.org/10.1057/s41599-023-02079-x>

Clark, S., & Haeck, P. (2024, October 9). *Europe's privacy patrol is spoiling Big Tech's AI party*. POLITICO. <https://www.politico.eu/article/europe-privacy-patrol-vengeance-block-ai-artificial-intelligence/>

Engler, A. (2023, April 25). *The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment*. Brookings. <https://www.brookings.edu/articles/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/#anchor5>

International Labour Organization. (2024, August 7). *Mind the Gap: Bridging the AI divide will ensure an equitable future for all*. International Labour Organization. <https://www.ilo.org/resource/news/mind-gap-bridging-ai-divide-will-ensure-equitable-future-all>

Lavanchy, M. (2018, November 18). *Amazon's sexist hiring algorithm could still be better than a human*. IMD. <https://www.imd.org/research-knowledge/digital/articles/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human/>

Leufer, D. (2024, September 24). *Why human rights must be at the core of AI governance - Access Now*. Access Now. <https://www.accessnow.org/human-rights-and-ai-governance/>

Michael. (2024, July 5). *Regional and International AI Regulations and Laws in 2024*. Keymakr's Blog Features the Latest News and Updates; Keymakr's Blog features the latest news and updates. <https://keymakr.com/blog/regional-and-international-ai-regulations-and-laws-in-2024/>

Régis, C., & Gaumont, E. (2023, January 27). *Assessing Impacts of AI on Human Rights: It's Not Solely About Privacy and Nondiscrimination*. Default. <https://www.lawfaremedia.org/article/assessing-impacts-of-ai-on-human-rights-it-s-not-solely-about-privacy-and-nondiscrimination>

Rodrigues, R. (2020). Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities. *Journal of Responsible Technology*, 4(100005), 1–12. ScienceDirect. <https://doi.org/10.1016/j.jrt.2020.100005>

Sanford, A. (2024, February 14). *Artificial Intelligence is Putting Innocent People at Risk of Being Incarcerated*. Innocence Project. <https://innocenceproject.org/artificial-intelligence-is-putting-innocent-people-at-risk-of-being-incarcerated/>

The Editors of Encyclopaedia Britannica. (2019). Turing test | Definition & Facts. In *Encyclopædia Britannica*. <https://www.britannica.com/technology/Turing-test>

Tuhu Nugraha. (2024, June 30). *The Urgency of AI Regulation for Developing Countries: Insights from Indonesia*. Modern Diplomacy. <https://moderndiplomacy.eu/2024/06/30/the-urgency-of-ai-regulation-for-developing-countries-insights-from-indonesia/>

United Nations Human Rights. (n.d.). *Taxonomy of Human Rights Risks Connected to Generative AI Supplement to B-Tech's Foundational Paper on the Responsible Development and Deployment of Generative AI*. <https://www.ohchr.org/sites/default/files/documents/issues/business/b-tech/taxonomy-GenAI-Human-Rights-Harms.pdf>

Wachter, S. (n.d.). Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond. *Yale Journal of Law & Technology*, 26(3). [https://yjolt.org/sites/default/files/wachter\\_26yalejltech671.pdf](https://yjolt.org/sites/default/files/wachter_26yalejltech671.pdf)

Wheeler, T. (2023, June 15). *The three challenges of AI regulation*. Brookings. <https://www.brookings.edu/articles/the-three-challenges-of-ai-regulation/>