United Nations A/79/170

General Assembly

Distr.: General
17 July 2024

Original: English

Seventy-ninth session
Item 71 (b) of the provisional agenda*
Promotion and protection of human rights: human rights questions, including alternative approaches for improving the effective enjoyment of human rights and fundamental freedoms

Human rights and international solidarity

Note by the Secretary-General

The Secretary-General has the honour to transmit to the General Assembly the report of the Independent Expert on human rights and international solidarity, Cecilia M. Bailliet, in accordance with Human Rights Council resolution 53/5.

* A/79/150.

24-13137 (E) 090824

Report of the Independent Expert on human rights and international solidarity, Cecilia M. Bailliet

Artificial intelligence and international solidarity – towards human-centred artificial intelligence international solidarity by design

Summary

The present report is the first report prepared for the General Assembly by the Independent Expert on human rights and international solidarity, Cecilia M. Bailliet. It is submitted pursuant to Human Rights Council resolution 53/5. The Independent Expert addresses the current challenge of increased global inequality and explains how an artificial intelligence international solidarity design strategy can elucidate the duties of States, corporations and civil society to strengthen equal access to technology and the non-discriminatory inclusion of groups and individuals in vulnerable situations. In the report, the Independent Expert proposes that artificial intelligence governance should support State and corporate due diligence procedural mechanisms to involve direct and indirect stakeholders within data processing and decision-making in the life cycle of artificial intelligence.

I. Introduction

1. International solidarity is a principle of international law and a universal value that has captured the attention of policymakers who are concerned with solving global challenges, such as those relating to pandemics, climate change, inequality and clean energy.[1] There is a new consciousness emerging worldwide that is geared towards a transformative shift away from neoliberal inequality towards stewardship, solidarity and collective action driven by a values-based revolution.[2] It is imperative that artificial intelligence (AI) be used to unite and not divide humanity. In accordance with human rights, AI should promote collective international solidarity actions across borders.[3] In 2022, the United Nations Educational, Scientific and Cultural Organization (UNESCO) issued a Recommendation on the Ethics of Artificial Intelligence that recognized the need to strengthen solidarity to facilitate fair access to AI technologies and address the challenges that they bring to diversity and interconnectivity of cultures and ethical systems. In addition, solidarity can support the mitigation of potential misuse of AI, help to realize the full potential of AI and help to ensure that national AI strategies are guided by ethical principles.

2. As noted in one study, AI is increasingly being incorporated within societies in which "smart cities" are being created to facilitate the tracking and monitoring of citizens, potentially in every aspect of their lives, in settings where there is increased use of biometric registration, and as a tool within international development and humanitarian responsive actions to displacement and other crises.[4] The use of AI in government institutions and outsourced institutions (including companies and civil society organizations) to process and conduct decision-making in cases and claims, utilizing personal data, is expected to increase significantly around the world. AI surveillance particularly affects the poor, as State institutions utilize AI to identify the misuse of social welfare benefits, as a tool of "over-policing" marginalized neighbourhoods, and in the context of tracking irregular migration.[5] The importance of recognizing intersectoral vulnerabilities to AI discrimination, including race, ethnicity, religion, gender, location, nationality and socioeconomic status cannot be overstated.[6]

3. Moreover, the *Artificial Intelligence Index Report 2024* concluded that "AI developers lack transparency, especially regarding the disclosure of training data and methodologies. This lack of openness hinders efforts to further understand the robustness and safety of AI systems".[7] Further, information and communications technology (ICT) companies are more engaged with civil society organizations from the global North that may weaken solidarity with civil society organizations in the global South.[8] Efforts to redress that imbalance are complicated by the limited access for groups in the global South to participate in networks that address AI accountability for human rights violations.

---

[1] See Global Nation, *Global Solidarity Report 2023* (September 2023).

[2] Kurt April, "AI-Induced Solidarity Economy: The Need for Stewardship Orientation" in *Effective Executive*, vol. 26, No. 3 (2023).

[3] Input from the Internet Governance Forum. See also https://intgovforum.org/en/content/pnai-report.

[4] Linnet Taylor, "What is data justice? The case for connecting digital rights and freedoms globally", *Big Data and Society*, vol. 4, No. 2 (July–December 2017).

[5] Ibid.

[6] Ibid.

[7] Institute for Human-Centered AI, Stanford University, *Artificial Intelligence Index Report 2024* (Stanford University, California, 2024).

[8] See https://www.business-humanrights.org/en/from-us/briefings/dismantling-the-facade-a-global-south-perspective-on-the-state-of-engagement-with-tech-companies/dismantling-the-facade-a-global-south-perspective-on-the-state-of-engagement-with-tech-companies/.

4. In preparing the report, the Independent Expert sought contributions from Member States, civil society organizations, companies and academics. A consultation of civil society stakeholders was held in July 2024. There was a review of academic literature and of reports of United Nations treaty bodies and Charter bodies, and an analysis of international, regional and national legal standards.

II. Responsible artificial intelligence solidarity criteria: transparency, fairness, non-discrimination and inclusion

5. Given that data processing and decision-making powers are increasingly being delegated to AI, there is a fundamental need for transparency to ensure fairness, non-discrimination and inclusion. The *Artificial Intelligence Index Report 2024* defines fairness as: "Creating algorithms that are equitable, avoiding bias or discrimination, and considering the diverse needs and circumstances of all stakeholders, thereby aligning with broader societal standards of equity."[9] In the report, it is further stated that: "Results show that while most companies have fully implemented at least one fairness measure, comprehensive integration is still lacking. The global average for adopted fairness measures stands at 1.97 out of five measures" among the measures for which data were gathered as part of the Global State of Responsible AI survey. The World Benchmarking Alliance noted that as of 2023, only a quarter of the 200 most influential technology companies in the world are meeting a minimum standard of disclosure of the adoption of ethical AI principles on digital inclusion.[10]

6. In the "Rome Call for AI Ethics" inclusion is underscored as a key principle, and at the AI Ethics for Peace event held in Hiroshima in July 2024, the President of the Abu Dhabi Forum for Peace, Sheikh Abdallah Bin Bayyah, underscored that "Cooperation, solidarity and joint work are necessary to deal with the developments of artificial intelligence, in which interests, harms and benefits are mixed, to ensure that its systems and products are not only technically advanced but also morally sound".[11] The draft of the Global Digital Compact, as at 26 June 2024, sets forth a dedicated inclusive aim: "Our cooperation will close the digital divides within and between States and advance a digital environment that promotes and enables diversity".[12]

7. In accordance with the Guiding Principles on Business and Human Rights, technology companies should be expected to conduct human rights impact assessments and identify risks as part of their quality control processes, as well as engage with external stakeholders as part of the human rights assessments.[13] It may be suggested that an AI international solidarity approach would place a focus on the need for concrete establishment of procedural approaches to address discrimination and strengthen inclusion. It is recommended that intergenerational solidarity be incorporated within AI regulations.[14]

---

[9] *Artificial Intelligence Index Report 2024*.

[10] See https://www.worldbenchmarkingalliance.org/impact/investor-statement-for-ethical-ai-2024/.

[11] See https://www.romecall.org/ and https://www.romecall.org/ai-ethics-for-peace-hiroshima-july-9th-2024/.

[12] See the draft of 26 June, available at https://www.un.org/techenvoy/global-digital-compact.

[13] Kate Jones, *AI Governance and Human Rights: Resetting the relationship*, Research Paper, International Law Programme (London, Royal Institute of International Affairs, 10 January 2023).

[14] Sébastien Fassiaux, "Preserving consumer autonomy through European Union regulation of artificial intelligence: a long-term approach", *European Journal of Risk Regulation*, vol. 14, special issue No. 4 (December 2023). See also Jon Truby and others, "A sandbox approach to regulating high-risk artificial intelligence applications", *European Journal of Risk Regulation*, vol. 13, No. 2 (June 2022), recognizing solidarity as a principle within European Union AI regulation.

8. Social solidarity can be invoked to share the benefits and costs of AI deployment, promote diversity in AI trajectories and promote transparency and compliance in order to correct AI information asymmetries.[15] The Independent Expert highlights the importance of incorporating solidarity as a perspective when assessing AI design and implementation.

A. Inclusion and algorithmic non-discrimination as artificial intelligence solidarity aims

9. The need for regulation of AI to correct bias and discrimination and to ensure security is essential. AI can be emancipatory through its facilitation of education in remote locations, it can provide language accessibility through translation functions and it can be used to fight stereotyping and hate speech, hence there is a continuing need to support research on the social impact of AI. Institutions that are designing AI within decision-making processes (including in the judicial and administrative fields) should be inclusive in seeking input from vulnerable groups and democracy-oriented civic actors, throughout all stages from planning to application, in order to pursue the prevention of human rights violations and the mitigation of harm. Moreover, it is important to create an independent oversight mechanism to address data protection, one that is capable of regulating and issuing guidelines on collecting and processing personal data during different phases of the planning and deployment of AI. The Independent Expert supports the model of conditional demographic disparity that serves as a standard set of statistical evidence for automated discrimination cases.[16]

10. One submission for the present report sets forth the vision of an inclusive AI that fulfils the principle of international solidarity by creating platforms that amplify the perspectives of traditionally underrepresented populations in global dialogues, such as AI translation, social media platforms used by activists for networking, and by empowering civil society to hold Governments accountable, such as in cases of corruption.[17]

B. National standards on inclusion and non-discrimination

11. At the national level, there are numerous standards (the majority are aspirational, non-binding or in draft form) that declare allegiance to the principles of equality, inclusion and non-discrimination, however there is often an absence of separate institutional competence and concrete procedural mechanisms to ensure compliance and that provide remedies for violations.

12. Australia has voluntary, non-binding AI ethics principles that recognize that AI systems should benefit individuals, society and the environment and respect human rights, diversity and the autonomy of individuals.[18] Similarly, China has interim measures for the management of generative artificial intelligence services that call for the prevention of discrimination based on ethnic origin, religion, nationality, geographical origin, gender, age, occupation and health, as well as the adoption of

[15] Juan C. Mateos-García, "The Complex Economics of Artificial Intelligence" (2 December, 2018) (unpublished working paper).

[16] Sandra Wachter, Brent Mittelstadt and Chris Russell, "Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI", *Computer Law and Security Review*, vol. 41 (2021).

[17] Input from Jake Okechukwu Effoduh, Assistant Professor, Lincoln Alexander School of Law, Toronto Metropolitan University, Canada.

[18] See https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles.

measures to increase the diversity and non-discriminatory content of training data.[19] India has a national strategy for artificial intelligence that addresses equality and non-discrimination.[20] It sets forth that AI systems must treat individuals under the same circumstances relevant to the decision equally, that AI systems should not deny opportunity to a qualified person on the basis of their identity, and that it should not deepen the harmful historic and social divisions based on religion, race, caste, sex, descent, place of birth or residence in such matters as education, employment or access to public spaces. Furthermore, the strategy sets forth the aim that it should strive to prevent unfair exclusion from services or benefits. Brazil has a proposal for an AI regulation containing a non-discrimination standard that includes the right to challenge decisions and request human intervention and the correction of direct, indirect, illegal or abusive discriminatory biases.[21] Spain established an independent authority for equal treatment and non-discrimination that monitors and promotes the use of ethical, trustworthy and fundamental rights-compliant AI.[22]

13. The United States of America has a "Blueprint for an AI bill of rights: making automated systems work for the American people" that supports an Executive Order that calls for independent evaluation and reporting to ensure non-discrimination.[23] Kenya has a draft AI code of practice that actually calls for the creation of a mechanism to implement non-discrimination standards, including through the documentation of fairness evaluations, the recording of steps taken to address bias and the documentation of policies against discriminatory outcomes.[24]

C. Risks of biased applications of artificial intelligence

14. It has been recognized that "affordable and knowledgeable access to the Internet has become a fundamental need to fully realize all human rights and fundamental freedoms, democracy, development and social justice."[25] This has resulted in a range of private digital rights initiatives that seek to address Internet infrastructure, application and usage from the perspective of increasing accessibility. AI bias subjects individuals or groups to stereotyping or prejudice, by comparing them to others through the use of AI algorithms. AI bias appears within the design of an AI system, the collection and interpretation of data and through direct and indirect stakeholder interaction. The Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance confirmed the ongoing challenge of "data problems, algorithm design issues, the intentionally discriminatory use of artificial intelligence and accountability issues" (see A/HRC/56/68). A 2019 report by UNESCO on AI identifies solidarity as an element of the creation of a "knowledge society" and explains the urgency of addressing AI discrimination

[19] See https://www.chinalawtranslate.com/en/generative-ai-interim/.

[20] See https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf; https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf; and https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf.

[21] See draft law No. 2338 of 2023. Available at https://www25.senado.leg.br/web/atividade/materias/-/materia/157233.

[22] Input from Spain.

[23] See https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf; see also https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

[24] See https://www.dataguidance.com/sites/default/files/kebs-tc_094_n66_public_review_kenya_standard_dks_3007_ai_code_of_practice.pdf.

[25] George A. Walker, "Technology Law, Rights and Ethics – One Choice, One Future", *The International Lawyer*, vol. 56, No. 1 (2023).

resulting from intentional or unintentional discriminatory programing, bias in training data for machine learning algorithms, or other factors:[26]

> Many types of discrimination can be indirect; for example, an algorithm that uses mobile phone usage patterns to determine credit worthiness of a person is discriminatory if it assigns high credit risk to women in communities that (i) have low mobile phone usage or (ii) do not own mobile phones. The condition applied may appear to be equal and fair, but it disadvantages a particular group. Algorithms can cause and exacerbate these multiple forms of discrimination. Existing social and political biases are being systemized in machine learning algorithms in many ways. Furthermore, it is worth investigation into the potential new forms of discrimination that AI may bring about, such as exclusions decided based on statistical correlations that do not necessarily correspond to socially salient characteristics, but that are nonetheless strongly linked to one's personal identity.

15. The *Artificial Intelligence Index Report 2024* describes how the tokenization of language (in which languages are broken down into components for analysis) affects non-Western languages negatively, as AI will perform less effectively than with Western languages.[27] There are also civil society organizations that directly address algorithmic bias, exclusion, and discrimination against racial, ethnic or religious minorities, as well as on the basis of gender or other identity, and that work towards improving accessibility.[28] Amnesty International and the National Association for the Advancement of Colored People have raised concerns about predictive policing tools,[29] the use of automated systems to determine access to health care and social services, the surveillance of refugee and migrant movements, and the biased impact of AI tools used for facial recognition and fraud detection on racialized communities.[30] Persons subjected to these AI processes may be unaware and uninformed about the use of AI, and they often lack any remedies for their discriminatory treatment by State institutions. Civil society groups emphasize the need to include stakeholders in the design of AI regulations. It has been suggested that the European Union maintains the use of AI, despite criticism, in discriminatory surveillance systems such as risk assessment systems and predictive analytics to facilitate pushbacks,[31] while in the context of migration, AI is being developed for different purposes: (a) identity verification through the development of facial recognition technology; (b) risk assessment; (c) detention assessments; and (d) surveillance. The risk of minimization of human control has been highlighted in such contexts.[32]

---

[26] United Nations Educational, Scientific and Cultural Organization (UNESCO), *Steering AI and Advanced ICTs for Knowledge Societies: A Rights, Openness, Access and Multi-stakeholder Perspective* (Paris, 2019), p. 63.

[27] *Artificial Intelligence Index Report 2024*.

[28] See https://digitalrightsfoundation.pk/wp-content/uploads/2021/03/Policy-brief-2.pdf; https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content-instagram-and; https://www.apc.org/en/member/7amleh-arab-center-social-media-advancement; https://smex.org/; https://www.derechosdigitales.org/; and https://www.accessnow.org/.

[29] On predictive policing as leading to disproportional surveillance and policing of black communities, see: https://naacp.org/resources/artificial-intelligence-predictive-policing-issue-brief.

[30] See https://www.amnesty.org/en/latest/campaigns/2024/01/the-urgent-but-difficult-task-of-regulating-artificial-intelligence/.

[31] Input from Indira Boutier, Glasgow Caledonian University.

[32] Ibid. See also Petra Molnar and Lex Gill, *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System* (International Human Rights Program (Faculty of Law, University of Toronto) and Citizen Lab (Munk School of Global Affairs and Public Policy, University of Toronto), 2018), pp. 31–34.

16. Given that State institutions and companies utilizing AI for data processing and decision-making are not transparent in explaining its development, training, use or application, individuals or groups claiming human rights violations are often unable to produce the evidence connecting AI with the violation. The violation may first be marked by the speedy feedback loop between different State systems connected through AI, resulting in harmful treatment or decision-making by the State institution. Criminals are also using AI to fraudulently appropriate themselves of remittances sent by migrants to their families, or money sent by family members to migrants, in both the country of origin and the country of residency.

17. The General Assembly has recognized the need to combat algorithmic discrimination,[33] identify vulnerabilities, improve accessibility and provide remedies for human rights violations.[34] The use of AI within State institutions that are responsible for addressing migration, crime and health care and elderly care has received attention from the United Nations human rights treaty bodies and special procedures. The Committee on the Elimination of Racial Discrimination issued concluding observations in which it expressed concern about the discriminatory impact of the use of AI in the context of asylum.[35] The Committee on the Elimination of Discrimination against Women has issued concluding observations calling for adequate safeguards to prevent gender stereotyping associated with the biometric, surveillance and algorithmic profiling systems used by law enforcement authorities in crime prevention and investigation, and to adopt measures to eliminate algorithmic bias relating to artificial intelligence and algorithmic services.[36] The Independent Expert on the enjoyment of all human rights by older persons expressed her appreciation for the opportunity provided during her visit to learn about how artificial intelligence can be useful in relation to ageing, care and health services, but she also recommended that the use of data collection be reviewed, in order to ensure that the use of such data upholds non-discrimination obligations in relation to older persons of diverse backgrounds, thereby implying an AI solidarity perspective.[37]

18. Civil society groups underscored the negative impact that AI had on the ability of LGBTQI activists to engage in solidarity actions.

III. Regional initiatives to address institutional artificial intelligence processes

19. The Organisation for Economic Co-operation and Development (OECD) principles on artificial intelligence were updated in 2024 and underscore the importance of advancing inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and promoting diversity, fairness, social justice, transparency, human agency and oversight.[38] A problem with the increased interest of State institutions in using AI for case processing and decision-making is that many of these systems are built by corporations, hence there is a risk of a lack of transparency. State institutions may be unaware of the risk of AI problems such as “hallucinations”, whereby AI creates content that is not based on content, and

[33] General Assembly resolution 78/265, para. 6 (h).
[34] General Assembly resolution 78/213.
[35] CERD/C/DEU/CO/23-26, paras. 45–46.
[36] CEDAW/C/ITA/CO/8, para. 26. See also similar concluding observations in document CEDAW/C/DEU/CO/9, paras. 27–28. The Committee expressed its concern regarding AI and gender violence in concluding observations contained in document CEDAW/C/ESP/CO/9, paras. 21, 23, 31 and 33.
[37] A/HRC/45/14/Add.1, para. 93.
[38] See https://oecd.ai/en/ai-principles; see also https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

the essential need for human control to monitor the use of information and identify algorithmic bias. The OECD catalogue of tools and metrics for trustworthy AI provides examples of a variety of technical, educational and procedural approaches for pursuing AI solidarity.[39]

20. The Artificial Intelligence Act of the European Union addresses international solidarity through its identification of prohibited AI practices, in article 5 (1) (b), (c) (g) of the Act. The European Commission has proposed a Liability Directive that is intended to be implemented by national courts, and it thus sets forth evidentiary standards.[40]The European Commission adopted the European Ethical Charter on the use of artificial intelligence in judicial systems and their environment in 2018, in which it called for external audits of data processing in order to screen for discrimination.[41] The Commission recognized that the use of AI in migration, asylum and border control management and in the administration of justice and democratic proceedings carries a high risk and hence adopted the Artificial Intelligence Act in 2024. The Act calls for the creation of a European Artificial Intelligence Office (AI Office) within the Commission to monitor the effective implementation and compliance with the Act by providers of general-purpose AI models. The European Artificial Intelligence Board will be composed of representatives of member States and corporations. The AI Office may invite providers of general-purpose AI models and relevant national competent authorities to participate in creating codes of practice, while civil society, industry, academia, downstream providers and independent experts may support the process.[42] That standard appears weak in terms of its vagueness regarding the ability of civil society to have a direct input on impact assessments.

21. The European Court of Justice identified discrimination risks regarding its future use of AI in a strategy document: “One of the main risks associated with the adoption of A.I. technology is the possibility that involuntary biases may be introduced during the training of A.I. models, resulting in unintentional discrimination.” [43] The European Parliament Subcommittee on Human Rights conducted an analysis of AI that underscored the risk of entrenching structural discrimination: “AI systems can perpetuate and amplify biases, leading to discrimination in various sectors, including employment, law enforcement and credit scoring. There is substantial evidence proving that AI can entrench socioeconomic disparities by reflecting the prejudices present in their training data or design.”[44]

22. The Committee of Ministers of the Council of Europe adopted a Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law on 17 May 2024, which will be opened for signature in September 2024. Article 10 of the Framework Convention, on equality and non-discrimination, presents the idea of a life cycle approach to identify and correct bias systematically; AI systems

[39] See https://oecd.ai/en/.

[40] See https://artificialintelligenceact.eu/the-act/. See also the proposal for a directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive). Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52022PC0496.

[41] See https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment.

[42] See https://artificialintelligenceact.eu/high-level-summary/.

[43] See https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf.

[44] See https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf; see also Ahmet Bilal Aytekin, “Algorithmic bias in the context of European Union anti-discrimination directives”, paper presented at the European Workshop on Algorithmic Fairness, 7–9 June 2023, Winterthur, Switzerland; and https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf.

should be reviewed for discrimination throughout the period of use in accordance with the applicable law.[45]

> 1. Each Party shall adopt or maintain measures with a view to ensuring that activities within the lifecycle of artificial intelligence systems respect equality, including gender equality, and the prohibition of discrimination, as provided under applicable international and domestic law.
>
> 2. Each Party undertakes to adopt or maintain measures aimed at overcoming inequalities to achieve fair, just and equitable outcomes, in line with its applicable domestic and international human rights obligations, in relation to activities within the lifecycle of artificial intelligence systems.

23. There have also been other regional initiatives to promote digital inclusion, such as resolution 580 (LXXVIII) 2024 on Internet shutdowns and elections in Africa, adopted by the African Commission on Human and Peoples' Rights.[46] The Inter-American Court of Human Rights held a dialogue with judges in November 2023, during which they discussed the UNESCO *Global Toolkit on AI and the Rule of Law for the Judiciary.* The Toolkit describes issues of exclusion:

> The problem is that […] data might be infused with bias[…] For instance, clinical trials often exclude women and people of colour, leading to inadequate data representation. This could have severe consequences if algorithms trained using such data are used to analyse skin images or prioritize care for patients. As a result, it is crucial to ensure that AI algorithms are trained using representative data to avoid such biases and ensure equitable outcomes for all.[47]

24. The Toolkit also addresses the issue of AI representativeness:

> Digital divides in many Global South countries have led to "data invisibility," which is likely to impact historically marginalized groups like women, castes, tribal communities, religious and linguistic minorities, and migrant labour. The usefulness and validity of AI algorithms developed on readily available data may be constrained by biases perpetuated by data invisibility. This underlines the requirements for algorithmic transparency and accountability.

The Toolkit highlights the problem of proxy discrimination, such as the use of postal codes, educational level or income, by banks assessing loan applications that may be associated with ethnic or racial minorities, which is a method with an inherent risk of perpetuating intersectional discriminatory policies and practices. It suggests that AI may help to identify patterns of bias against marginalized groups and individuals by public and private actors, including the judiciary.

25. The Inter-American dialogue with judges also discussed the UNESCO Ethical Impact Assessment tool for AI, which calls for the consideration of "the diversity of the AI project team, especially in terms of – but not limited to – gender, age, race, colour, descent, language, religion, national origin, ethnic origin, social origin, economic or social condition, disability, and sexual orientation, including how this reflects the complexity and diversity of expected user population, and how this could introduce biases".[48]

[45] See https://rm.coe.int/1680afae3c.

[46] See https://achpr.au.int/en/adopted-resolutions/580-internet-shutdowns-elections-africa-achprres580-lxxvii.

[47] See https://nataliazuazo.com/2023/11/20/ai-and-the-rule-of-law-at-the-inter-american-court-of-human-rights/. See also https://unesdoc.unesco.org/ark:/48223/pf0000387331.

[48] See https://unesdoc.unesco.org/ark:/48223/pf0000386276.

A. Solidarity for digital literacy against disinformation

26. There are increased calls for regulation in "extremist-triggered election misinformation campaigns".[49] The improved quality of AI-generated content makes it difficult for people to identify disinformation and false content, and there are increased reports of Internet shutdowns by States, which have a disproportional impact on access to information. There is a need for an international solidarity strategy, incorporating legislative, technical and educational cooperation in order to raise digital literacy.[50] Responsive solidarity has been invoked as a means to handle emergencies linked to the malicious use of AI, such as "when thousands of deep fake videos with ethnic violence circulate on a day of elections in a country with a history of genocide".[51]

27. The European Union adopted a Digital Services Act in 2022, which entered into force in 2024, calling for the use of codes of conduct by large online platforms and search engines, including the implementation of responsive measures to disinformation (including false issue and political adverts) that would be strictly necessary and targeted in application.[52] The Act supports the Strengthened Code of Practice on Disinformation 2022 (applicable to self-regulating entities), which includes guidance on demonetizing the dissemination of disinformation, guaranteeing transparency in political advertising, enhancing cooperation with fact-checkers and facilitating researchers' access to data.[53] Nonetheless, there is a critique that "while these developments may have a unifying effect, the continental approach to disinformation is currently fragmented",[54] Moreover, there is concern that there may be potential for conflict between the codes of conduct and the rulings of the European Court of Human Rights.[55]

28. The Artificial Intelligence Act of the European Union also contains regulations to curb the influence and impact of deep fakes. According to article 50 (4), the deployer, i.e. the person using an AI system, must generally disclose that the content was generated by AI. However, the law itself contains exceptions to this disclosure obligation, meaning that its effectiveness must be scrutinized.

29. The African Union Development Agency calls for legal protection and regulatory frameworks to combat algorithmic discrimination.[56] The African Union Executive Council supports the creation of a continental strategy based on the

[49] Input from from Jake Okechukwu Effoduh.

[50] Ibid.

[51] Miguel Luengo-Oroz, "Solidarity should be a core ethical principle of AI", *Nature Machine Intelligence*, vol. 1 (November 2019). See also Patrik Hummel and Matthias Braun, "Just data? Solidarity and justice in data-driven medicine", *Life Sciences, Society and Policy*, vol. 16, No. 8 (2020).

[52] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act).

[53] See https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation.

[54] Input from Dorothy Andersen, Human Rights Fellow in the International and Comparative Law Program at the Law School of George Washington University, Washington, D.C.

[55] Paolo Cavaliere, "The Truth in Fake News: How Disinformation Laws Are Reframing the Concepts of Truth and Accuracy on Digital Platforms", *European Convention on Human Rights Law Review*, vol. 3, No. 4 (November 2022). See also Ethan Shattock, "Fake News in Strasbourg: Electoral Disinformation and Freedom of Expression in the European Court of Human Rights", *European Journal of Law and Technology*, vol. 13, No. 1 (2022). See also https://edmo.eu/wp-content/uploads/2022/01/Case-law-for-policy-making-Report-2022.pdf.

[56] See https://www.nepad.org/publication/ai-and-future-of-work-africa-white-paper.

conceptual framework on AI adopted by the Executive Council.[57] Within the region, there is a trend towards "technological content" regulation that involves censoring or blocking access to content in order to prevent fake news and/or hate speech and "legal content" regulation that includes discussing, drafting and passing bills to regulate fake news and hate speech. Both measures run a risk of overly broad application or being targeted towards political opposition and thereby weakening the independence of elections.

30. Saudi Arabia informed the Independent Expert about the creation of the Saudi Data and Artificial Intelligence Authority as a significant development for the Arabic large language model (known as ALLaM). The initiative is aimed at leveraging AI technologies and digital applications to foster cultural diversity and benefit all of humanity. One of the projects, SauTech, is focused on localizing AI technology and preserving local culture, and is able to recognize and transform speech to text across a wide range of Arabic dialects in order to find solutions tailored to local contexts.[58] The Saudi Data and Artificial Intelligence Authority raises awareness of AI through training programmes. Similarly, Colombia provided input that described AI training at different levels.[59]

B. Gender solidarity in artificial intelligence

31. There is also increased attention to the potential of an international solidarity approach to combat gender discrimination in relation to AI.[60] The *Artificial Intelligence Index Report 2024* found that European countries reported more male than female graduates in information technology-related studies, with slow progress in narrowing the gender gap.[61] The Committee on the Elimination of Discrimination against Women issued concluding observations to a number of States parties' reports in which it expressed its concern regarding the underrepresentation of girls and women in artificial intelligence, and the absence of concrete measures to prevent the integration of gender stereotypes and algorithmic gender discrimination and violence in AI coding, calling for the establishment of effective safeguards.[62]

32. In its input, Saudi Arabia reported that women constituted over 50 per cent of the technical workforce of the Saudi Data and Artificial Intelligence Authority.[63] The country's International Centre of AI Research and Ethics launched the Elevate Initiative in collaboration with Google Cloud, aiming to train 25,000 women in the area of data and AI. The Institute of Democracy and Human Rights in Peru reported that the President of Peru had supported a training programme for business projects developed with AI. Citizens registered since May 2024 had the opportunity to apply

[57] See https://au.int/en/newsevents/20240419/multistakeholder-consultative-sessions-development-continental-strategy.

[58] Input from Saudi Arabia

[59] Input from Colombia.

[60] Keratso Georgiadou, "Solidarity, Gender, Dialogue in the AI era", in *Kritische Pädagogik und Bildungsforschung: Anschlüsse an Paulo Freire*, Wassilios Baros, Rita Braches-Chyrek, Solvejg Jobst and Joachim Schroeder, eds. (Wiesbaden, Germany, Springer Fachmedien, 2024), pp. 477–489.

[61] *Artificial Intelligence Index Report 2024*.

[62] See CEDAW/C/ITA/CO/8; CEDAW/C/DEU/CO/9, paras. 27 and 28; CEDAW/C/ESP/CO/9, paras. 21, 23, 31 and 33; CEDAW/C/TJK/CO/7, paras. 47, 48 and 55. See also concluding observations contained in CEDAW/C/TUR/CO/8, para. 46; CEDAW/C/CRI/CO/8, paras. 27, 28 and 37.

[63] Input from Saudi Arabia.

for study scholarships, of which 50 per cent were given to women in vulnerable situations.[64]

33. The Business and Human Rights Resource Center describes the ongoing challenges of AI promoting counter-solidarity against women:[65]

> Women and gender non-conforming persons continue to feel a disproportionate impact of poorly designed, developed and deployed technologies, as evidenced by the fact that the latest technological tools, including generative AI, are released with flaws and biases that amplify sexism and gender stereotypes, facilitate the social control of women, disempower women patients, discriminate against female job applicants, facilitate the targeting of women human rights defenders, and disadvantage women in their access to welfare benefits, amongst other harms.

34. Women are increasingly subject to AI hate speech, cyberbullying, non-consensual sexual content, revenge porn and stalking, all of which negatively affect their ability to enjoy self-fulfilment of personal and professional aspirations. Civil society organizations described an innovative initiative aimed at ending gender violence that included a digital security training, held at Kampala International University in 2023.[66] They also described the use of social media as essential in enabling international solidarity to demand the prosecution of femicide and legislative reform in Kenya, the declaration of femicide and violence against women as a national emergency, and the establishment of a commission specifically tasked with tackling these crimes and breaking the cycle of impunity.[67] They called for increased international collaboration between technology companies, non-governmental organizations (NGOs) and governments to develop effective and ethical AI-powered content moderation tools to identify and flag online violence against women, that are adaptable to different languages and cultural contexts.

C. Solidarity in artificial intelligence design

35. The second revision of the Global Digital Compact declares that it will be implemented in a spirit of global solidarity and specifically underscores inclusion by focusing on the need for digital technology companies and developers to engage with users of all backgrounds and abilities to incorporate their perspectives and needs into the life cycle of digital technologies, pursue the fair distribution of digital benefits and support digital accessibility and linguistic and cultural diversity in the digital space.[68] Moreover, the international community calls for inclusion of the needs of persons in vulnerable situations and those in underserved, rural and remote areas in the development and implementation of national and local digital connectivity strategies, and underlines the need to target and tailor capacity-building for women and girls, children and youth, as well as older persons, persons with disabilities and persons in vulnerable situations, and ensure their meaningful engagement in the design and implementation of programmes. In addition, it seeks to develop and conduct national digital inclusion surveys with systematic disaggregation of data by age, disability and gender, so as to identify learning gaps and inform priorities in specific contexts.

---

[64] Input from Instituto de Democracia y Derechos Humanos, Peru.

[65] See https://media.business-humanrights.org/media/documents/BHRRC_Submission__Gender_and_Role_of_Business_2023_j2HkLgP.pdf.

[66] Input from Our Voices Our Futures Consortium.

[67] Ibid.

[68] Available at https://www.un.org/techenvoy/global-digital-compact.

36. The Global Digital Compact supports the adoption of national AI legislation and effective oversight and remedy mechanisms. It also calls upon the tech companies and AI developers to recognize the obligation to respect human rights and implement due diligence and impact assessments. It calls upon tech companies and AI developers to co-develop industry accountability frameworks (in consultation with governments and other stakeholders) that increase transparency around their systems and processes, define responsibilities, and commit to standards and auditable public reports. It calls upon social media platforms to establish safe, secure and accessible reporting mechanisms for users and their advocates to report policy violations (including special reporting mechanisms adapted to children and persons with disabilities). It advocates that AI tech companies and social media platforms enhance the transparency and accountability of their systems, including their terms of service, content moderation and algorithms and their handling of users' personal data in local languages, in order to empower users to make informed choices and provide or withdraw informed consent.

37. UNESCO emphasizes inclusion in designing responsive AI stakeholder mechanisms as an element of transparency:[69]

- Does the government encourage participation by other stakeholders in national governance of AI?
- Are there active associations of AI professionals, consumers and other stakeholder communities?
- Does the government actively involve other stakeholder groups in developing policy towards global AI governance?

It calls for the implementation of indicators to ensure inclusion by measuring to what extent diverse stakeholder groups (women and gender-diverse persons, youth and marginalized groups) are involved in AI policymaking and participate in international and regional AI forums, processes and mechanisms.

38. It has been suggested that: "AI technologies within the sociotechnical system may then serve to either facilitate or shape social interactions and can therefore undermine or promote solidarity."[70] It is essential to include stakeholders in decision on AI design, noting that their relevance may change over time, hence there may be unforeseen solidarity issues. Human rights are the measure of solidarity in relation to AI.[71] One may consider that the use of AI may result in scenarios in which there are violations of the right to privacy due to facial recognition, or violations of non-discrimination of women or minorities in hiring, or denial or self-realization of life goals (or a life's project) such as denial of requests for a housing or educational loan. Human rights impact assessments should be considered central elements of AI solidarity design. Moreover, users' interests and needs should be identified and accounted for in the design process. Participatory design would invite users to discuss design choices and participate in the decision-making process, which should include collective decision-making that can enable solidarity within the group. However, a sole focus on users fails to acknowledge indirect stakeholders such as those who are unintentionally affected. An AI solidarity perspective would require going beyond a focus on users in order to focus on all affected stakeholders within the sociotechnical system. It would entail not only the assessment of harms and benefits overall, but also an assessment of how they are distributed across society and stakeholder groups. The fundamental issue concerns how to have tech companies accept the obligation of

[69] UNESCO, *Steering AI and Advanced ICTs for Knowledge Societies.*

[70] Catharina Rudschies, "Exploring the Concept of Solidarity in the Context of AI: An Ethics in Design Approach", *Digital Society*, vol. 2, No. 1 (2023).

[71] Ibid., at page 12.

solidarity and the sharing of the risks in addition to the sharing of the benefits. An AI solidarity approach would conduct a human rights risk assessment, identify the right of individuals to self-realize, and regulate for the distribution of responsibility for harm, in addition to distribution of the benefits of data processing and decision-making.[72] Hence, political solidarity to support collective decision-making, as well as a solidarity-based governance framework for AI, to deal with the implications of AI, will also be necessary.

39. An AI international solidarity approach may require the creation of regulations within different fields, such as health or energy. Dialogue in the design and development of AI should promote collaboration among users, AI technology developers, and other stakeholders to find AI solutions. There is a need for the creation of engagement pathways to support dialogue with AI use case developers, AI solution developers, or any other applicable AI/machine learning developer, to enable ongoing updates to address evolving risks and benefits of AI solution uses. The establishment of risk-based, tailored communications and engagement plans to enable customers to easily understand explanations about how the AI solution was developed, its performance and maintenance, and how it aligns with the latest best practices and regulatory requirement, is also necessary.

40. Recognition and respect should be considered elements of data solidarity that give "the concrete experiences of marginalization and experiences of injustice in data-intensive contexts a conceptual place."[73] Solidarity as an AI principle should imply sustainable equitable impact to share prosperity and burdens and prevent inequality.[74] There is an academic proposal of an overarching AI solidarity framework built on incentives to prevent inequality such as the payment of royalties each time an AI system trained with someone's data is used, rewards to doctors using an AI model to conduct diagnosis, or remuneration to persons producing text for an AI automatic text generator every time the robot writes a public article.[75] Others call for a solidarity approach that measures the external impact on society by advocating safeguards for individual control of data that are being processed.[76] An additional perspective is that a data solidarity framework that would make data processes visible for the public good.[77] They suggest that this can prompt corporate and public data stakeholders to share both the risks and the benefits of data access, production and sharing. Data solidarity is presented as promoting the visibility of data set processes in order to identify and correct marginalization-based discrimination.[78] They call for collective action, utilizing solidarity as a principle of data governance for the creation of publicly held datasets in order to start building trust and accountability. There is a view that purports that the international community is facing a turning point in which data solidarity should promote the creation of infrastructure and should be shaped according to democratic values, such as inclusion.[79] Further, this may require the establishment of an independent global governance entity with a membership composed of representatives of industry, States, civil society, international organizations and academia to promote human rights-based rules on AI.[80] One may

[72] Ibid., at page 13.

[73] Hummel and Braun, "Just data? Solidarity and justice in data-driven medicine".

[74] Luengo-Oroz, "Solidarity should be a core ethical principle of AI".

[75] Ibid.

[76] Hummel and Braun, "Just data? Solidarity and justice in data-driven medicine".

[77] Mercedes Bunz and Photini Vrikki, "From Big to Democratic Data: Why the Rise of AI Needs Data Solidarity", in *Democratic Frontiers: Algorithms and Society*, Michael Filimowicz, ed. (London, Taylor & Francis, 2022).

[78] Ibid.

[79] Ibid.

[80] Ana Beduschi, "Human rights and the governance of artificial intelligence", research brief (Geneva Academy of International Humanitarian Law and Human Rights, 2020).

consider Anthropic's "Claude's Constitution", which includes principles that encourage the consideration of non-Western perspectives, in addition to human rights.[81]

41. Some tech companies gave input that described the implementation of AI international solidarity by design:

(a) Training sessions are designed in an accessible, culturally sensitive and inclusive manner;

(b) Algorithms are trained on diverse data sets to represent the people they serve;

(c) Regular audits, test across different demographic groups and implement fairness-aware machine-learning methods;

(d) Ethical AI guidelines that prioritize fairness, transparency and accountability in algorithm development and deployment;

(e) Continuous monitoring and adjustment of AI systems to prevent discrimination;

(f) Engagement with advocacy groups, NGOs and community representatives from vulnerable populations to understand concerns and gather feedback on the impact of technologies;

(g) Ongoing training to teams on algorithmic fairness and the ethical implications of AI;

(h) Content moderation policies that prohibit the dissemination of false information, hate speech and harmful content;

(i) Enforcement through automated tools and human moderators trained to identify and remove disinformation;

(j) Collaboration with reputable fact-checking organizations;

(k) Prioritize transparency in their algorithms to mitigate the unintentional amplification of disinformation;

(l) Algorithm-design to promote credible sources and reduce the visibility of misleading content;

(m) Educate users about identifying and reporting disinformation through informational campaigns and in-platform notifications;

(n) Work with governmental and non-governmental agencies and academic institutions to share insights and best practices for combating disinformation effectively.

IV. National orientation towards positive practice

42. Spain reported that it has a national artificial intelligence strategy and an agency for its supervision.[82] It stated that it will establish an ethical and regulatory framework that reinforces the protection of individual and collective rights, in order to guarantee inclusion and social welfare. Spain plans to develop a charter of digital rights and launch a model of national governance of AI ethics through an AI advisory council,

[81] See https://www.anthropic.com/news/claudes-constitution.

[82] See https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/ficheros/National-Strategy-on-AI.pdf.

in collaboration with a digital transformation advisory council. Article 23 of law No. 15/2022 of 12 July 2022 stipulates that measures must be taken to reduce biases and at the same time contribute to greater transparency and accountability when AI is used in decisions by public bodies. These measures include analysing the design and the training data and evaluating them for any discriminatory influence.[83] Moreover, it seeks to establish forums for dialogue, awareness-raising and national and international participation in the field of AI, to foster communication between government, the scientific community, the private sector and civil society.

43. Saudi Arabia reported that the Saudi Data and Artificial Intelligence Authority has a system to organize, develop and handle AI data and provide government services, seeking to reliably support digital transformation and data preservation through the inclusion of stakeholders in development and implementation of AI.[84] It promotes seven AI ethics principles, including fairness; privacy and security; humanity; social and environmental benefits; reliability and safety; transparency and explainability; and accountability and responsibility.[85] The Authority announced an incentive programme to help companies voluntarily comply with AI ethics. The process begins by identifying and evaluating all potential risks and the severity of their impact.

44. Malaysia provided input that described its proposed oversight mechanism, an AI coordination and implementation unit, that will act as the government body on all matters related to AI.[86] It plans to establish a foresight committee that will undertake horizon scanning, foresight and policy advocacy.

45. Germany utilizes AI as part of its Match'In programme to improve the integration of refugees within municipalities by consulting them to identify their preferences as to housing, employment/professional experience, hobbies, recreation, health care, family status, religious community and other matters.[87] That approach promotes recognition of the refugees' right to self-realization, in keeping with an international solidarity approach. It has been noted that the Match'In programme is the only programme that includes the perspectives of the refugees in the design programme.[88]

46. The Dominican Republic indicated that it has a national artificial intelligence strategy that prioritizes the use of AI in public administration in key sectors such as justice, health, education, environment and security, and as predictive analytical models to design services for the citizenship. It aims to have code that would include preventive actions, procedural guarantees and accountability mechanisms to ensure responsible implementation of AI. In addition, supervision mechanisms would be established to verify ethical compliance. The YoSoyFuturoRD Human Talent and Innovation HUB will give priority to vulnerable sectors. It proposed the creation of a strong and collaborative regional AI ecosystem to significantly contribute to technological advancement, economic and social development, and cooperation within the region. It states that it will establish supervision, compensation and recourse mechanisms against damage caused by AI systems, in order to safeguard the rights of its citizens.

47. The municipality of Rawson in Argentina uses AI to promote social assistance programmes and incentives, such as "Red de Economía Social y Solidaria", to

[83] Input from Spain.
[84] Input from Saudi Arabia.
[85] Similarly, Colombia notes in its input respect for principles such as transparency, privacy, human control and non-discrimination, ensuring that technologies are inclusive and fair.
[86] Input from Malaysia.
[87] See https://matchin-projekt.de/en/.
[88] See https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf.

generate sustainable food circuits framed in food sovereignty. It also has neighbourhood service centres for the municipality of Rawson (Municerca), which process requests for improvement of public spaces and streets, the initiation and consultation of cases, and the reception of municipal claims.

A. Health-care solidarity and artificial intelligence

48. Health-care solidarity should be at the foundation of regulations on the use of AI to ensure access and equity, but achieving that objective is rendered complex owing to the privatization of health care in many countries.[89] Pre-existing health disparities due to race, class, age, and other factors could be replicated within AI medical systems that do not address structural injustice. AI used for delivering health care needs to be designed with community health care in mind, and more specifically to be deployed in low-resource areas by facilitating health coverage in low-resource settings. Some have put forward a vision of AI health-care solidarity that would place individuals in a position to share or retract data by controlling data flows, by creating infrastructure for controllability and by focusing on output orientation in governance so as to prevent and mitigate injustice.[90] Brazil has a Bolsa Família programme that uses digital tools to distribute aid to low-income families. The programme has reduced poverty and promoted the development of human capital by linking cash transfers to specific conditions, such as school attendance and medical check-ups. Malaysia uses AI to assist medical diagnosis, develop personalized treatment, provide information through chatboxes and assist with predictive analytics in order to identify proactive measures.[91]

49. Saudi Arabia described its investment in AI health-care solidarity. It has a Centre of Excellence for Artificial Intelligence in the Health Sector. Saudi experts and engineers developed Eyenai to revolutionize diagnostic medicine on a regional level. Diabetic retinopathy is a disease caused by diabetes types 1 and 2; it is one of the leading causes of blindness in Saudi Arabia. Early diagnosis is vital for the possibility of alleviating later complications, and Eyenai facilitates this by providing accurate detection and diagnosis. An application developed by the Saudi Data and Artificial Intelligence Authority (Tawakkalna) received a United Nations Public Service Award in 2022 for enabling an innovative response to the coronavirus disease (COVID-19) pandemic.

B. Worker solidarity and artificial intelligence

50. The fragmentation of labour resulting from the use of AI is seen as an impediment to the ability of workers to organize in solidarity: “The ease with which a platform company can employ new workers, sometimes even from around the world, generates competition and isolation, and significantly hinders the formation of a collective identity. In short, platform-based work fragments workers’ collective identity and sabotages collective action, especially action based on solidarity and trust.”[92]

[89] Nicolas Terry, “Of Regulating Healthcare AI and Robots”, *Yale Journal of Law and Technology*, special issue No. 133 (2019).

[90] Hummel and Braun, “Just data? Solidarity and justice in data-driven medicine”.

[91] Input from Malaysia.

[92] Tammy Katsabian, “The rule of technology – how technology is used to disturb basic labor law protections”, *Lewis and Clark Law Review*, vol. 25, No. 3 (2021).

51. Nevertheless, it has been suggested that solidarity brokers can help workers in times of crisis.[93] A central query concerns whether AI can be re-imagined as an instrument for fostering worker solidarity.[94] The Social Security Bank of Uruguay created centralized, automated procedures for employers, promoting self-management through online services, a mobile application and multichannel assistance. A chatbot reportedly answered 97 per cent of queries with a 100 per cent satisfaction rate. The results included a significant 24.4 per cent reduction in the social security contribution evasion rate, with 57 per cent of employers registering online and 42 per cent making payments online.[95] The Office of the Superintendent of Occupational Risks of Argentina successfully implemented an AI chatbot called Julieta.[96] On the one hand there are new forms of horizontal solidarity that are distinctly more open and diverse. There is also the trend of "distributed discourse" "in which the bureaucratic barriers between activists and union officials were removed, transparency was enhanced, and everyday members had increased power to challenge and reformulate oligarchical union structures".[97] Moreover, there is the emergence of "accelerated pluralism", which has been defined as "the ongoing fragmentation of interest-based group politics in favour of more issue-based and fluid group politics and new safe spaces for democratic discussion among women and other workers who had previously been marginalised by traditional union structures".[98] One may consider the increased ability of women agricultural workers to access AI with predictive capacity to anticipate harvest and weather patterns to be an example of AI solidarity.[99] Another example is the anti-apartheid movement led by workers in the field of technology.[100] On the other hand, there is the adept engagement of the California Nurses' Association to demand AI safeguards to protect patients against deficient treatment.[101] Nevertheless, it has been observed that "employers have counter-mobilised in the online space and the uncertain legal status of social media speech has been observed to exert an additional chilling effect on online employee speech."[102]

52. Hence, the context of the type of trade union activity as well as the political and organizational differences of labour movements will affect the use of AI in relation

---

[93] Saiph Savage and Mohammad H. Jarrahi, "Solidarity and AI for transitioning to crowd work during COVID-19", paper produced for virtual symposium on The New Future of Work, August 2020. See also Kurt April, "AI-Induced Solidarity Economy: The Need for Stewardship Orientation".

[94] Frances Flanagan and Michael Walker, "How can unions use artificial intelligence to build power? The use of AI chatbots for labour organising in the US and Australia", *New Technology, Work and Employment,* vol. 36, No. 2 (2021).

[95] See https://www.issa.int/sites/default/files/documents/2024-06/2-AI%20in%20SecSoc%202024.pdf.

[96] Ibid.

[97] See Andy Hodder, David Houghton, "Union use of social media: a study of the University and College Union on Twitter", *New Technology, Work and Employment*, vol. 30, No. 3 (November 2015).

[98] See Anne-Marie Greene and Gill Kirton, "Possibilities for remote participation in trade unions: mobilising women activists", *Industrial Relations Journal*, vol. 34, No. 4 (October 2003).

[99] See CEDAW/C/OMN/CO/4.On AI and weather prediction, see: https://deepmind.google/discover/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/.

[100] See https://www.notechforapartheid.com/.

[101] See https://www.nationalnursesunited.org/press/cna-demand-patient-safeguards-against-artificial-intelligence.

[102] Flanagan and Walker, "How can unions use artificial intelligence to build power? The use of AI chatbots for labour organising in the US and Australia". See also Louise Thornthwaite, "Chilling times: Social media policies, labour law and employment relations", Asia Pacific Journal of Human Resources, vol. 54, No. 3 (August 2015).

to international solidarity in matters of contestation of power, reflecting group identity, internal cohesion, culture, strategy, governance and community.[103]

V. Due diligence complaint procedures

53. The growing concern about transparency is prompting claims for a review of corporate overreach in terms of access to and use of personal data. For example, Meta announced that it planned to implement a default setting whereby users' content serves to train artificial intelligence models. The Norwegian Consumer Council considered the opt-out process to be complex and filed a legal complaint against the company for violating the General Data Protection Regulation of the European Union.[104] The Guiding Principles on Business and Human Rights articulate a due diligence obligation of companies, set out in principle 15, that includes a call for preventive approaches as well as a remedial process; principle 25 sets out the framework for access to justice; and principle 29 calls for business enterprises to set up remedial mechanisms.

54. The United Nations Global Principles for Information Integrity call for the creation of oversight mechanisms and the commissioning of regular external and independent human rights audits that would cover, among other areas: terms of service, community standards, advertising policies, content moderation, complaints processes, data access for researchers, impact assessments on vulnerability and marginalization, on gender equality and on children's rights. Audits should be public, accessible and understandable for all users.[105]

55. In its Recommendation on the Ethics of Artificial Intelligence, UNESCO indicated that States are expected to create oversight mechanisms that would be inclusive of all stakeholders, and to encourage all stakeholders to develop human rights. In Directive (EU) 2024/1760 of the European Parliament and of the Council of 13 June 2024 on corporate sustainability due diligence, article 14 sets forth the obligation of States to ensure that companies to have a fair, publicly available, accessible, predictable and transparent complaint mechanism for persons, civil society organizations, human rights defenders, trade unions and others, in order to address human rights violations and the environmental impact of the companies, their subsidiaries, or other entities in the chain of activities.[106] These measures should uphold confidentiality standards and prevent retaliation against the person or entity that filed the complaint. The directive creates a framework that complements the communication mechanism called for in the revised draft declaration on the right to international solidarity (art. 6 (3)). Moreover, the revised draft declaration (art. 8 (3)) sets forth a flexible framework for the potential design and sharing of AI international solidarity policies and practices at the universal periodic review. According to article 8 (3) of the revised draft declaration States have the duty to take steps within their respective capacities to facilitate the protection of actual and virtual spaces of communication, including access to the Internet and infrastructure, in order to enable individuals and peoples to share solidarity ideas. These norms provide a foundation to support creative AI international solidarity oversight mechanisms that could promote greater participatory inclusion by civil society.

[103] Ibid.

[104] See https://www.forbrukerradet.no/side/legal-complaint-against-metas-use-of-personal-content-for-ai-training/.

[105] See https://www.un.org/sites/un2.un.org/files/un-global-principles-for-information-integrity-en.pdf.

[106] See https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401760.

VI. Conclusion

56. The Independent Expert agrees with the view that there is a need to create a global multistakeholder governance model to prevent and redress discrimination and other human rights violations within AI systems.[107] In terms of governance, the High-level Advisory Body on Artificial Intelligence recommends an interoperable approach that covers public and private actors, and extends across jurisdictions, including the international, regional and national levels, such as the Artificial Intelligence Act of the European Union.[108] Nevertheless, a key marker of the evolving system is the use of voluntary codes and agreements, such as the anti-deepfake agreement to combat election interference, signed by technology companies, AI developers and security firms.[109] At present, the international community lacks an equitable sharing of AI solidarity benefits and risks and there appears to be a fundamental lack of trust between civil society organizations of the global South and technology companies and AI developers from the global North. Moreover, the majority of AI ethics principles are drafted in the global North and may lack attention to context-specific concerns of the global South, even as the latter provides the minerals and energy that are essential for the development of AI data infrastructure.[110] There is a concentration of power among the technology companies and AI developers that heightens the risk that AI related anti-solidarity measures will be strengthened, thereby worsening the digital divide between and within countries and among different sectors of society. The Internet Governance Forum underscores the problem of the exclusion of representatives of the global South from participation in global consultations, such as lack of funding, precarious connectivity, and the prioritization of English as the language for dialogue.[111] AI international solidarity, by design, can form a key value for future AI development and for the implementation of oversight mechanisms to uphold procedural justice and inclusionary participation by all. In order to support a policy change towards the implementation of a sustainable AI international solidarity, the Independent Expert offers recommendations to be adopted urgently within the next five years.

VII. Recommendations

57. **The Independent Expert recommends that States, companies and civil society actors:**

(a) **Ensure that government institutions, technology companies and civil society groups implementing AI for data processing and decision-making include the active participation of all individuals and groups, irrespective of their race, colour, descent, gender, age, language, religion, political views, national or ethnic origin, social or economic background, disability, or any other factors, in accordance with the UNESCO Recommendation on the Ethics of Artificial Intelligence;**

[107] See https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality?gad_source=CjwKCAjw4f6zBhBVEiwATEHFVvzWNyAYvJV56epBlSUMbQTVZ4hm_tCsn_VJGGvIzoRMMXfutYebdhoCzqEQAvD_BwE.

[108] See https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf. See also https://oecd.ai/en/accountability.

[109] See https://www.techradar.com/pro/top-tech-companies-ai-developers-and-security-firms-sign-anti-deepfake-agreement-to-combat-election-interference.

[110] Input from the Internet Governance Forum.

[111] Ibid.

(b) **Adopt national regulations to promote AI international solidarity by design, applying a non-discriminatory approach by ensuring the inclusion of civil society and other relevant stakeholders in the planning, selection, design and implementation of AI in government institutions and technology companies that are involved in data processing and decision-making;**

(c) **Governments and companies should conduct continuous human rights impact assessments with direct input by civil society before the collection, processing or deployment of data and provide full transparency, including on the training of AI;**

(d) **Ensure that government institutions and technology companies implementing AI for data processing and decision-making establish independent, external and systematic transparency audits and human rights impact assessments (to identify biases in training, algorithms and decision-making) with international solidarity as a value and aim throughout the life cycle of AI;**[112]

(e) **Ensure the inclusion of independent, transparent, accessible, effective complaint and appeal mechanisms in national and regional legislation so as to provide for accountability for discrimination and/or exclusion through or by AI utilized in government institutions, technology companies or civil society groups involved in data processing or decision-making;**

(f) **Recognize that the burden of proof should be placed on the relevant government institution and/or technology company to show how the AI technology is utilized in cases involving discrimination or other violations of human rights due to AI data processing or decision-making;**

(g) **Adopt explicit regulation of AI in institutional data processing and decision-making impact assessment and rights protection mechanisms when the development involves sensitive data (e.g., health data) or massive amounts of data (e.g., national databases);**

(h) **Ensure that all individuals and groups in a vulnerable situation are informed by State institutions and technology companies or civil society groups in an easily understood manner that AI will be used in data processing or decision-making, and that consent will be requested beforehand without direct or indirect coercion;**

(i) **Ensure access to justice mechanisms that are independent in order to address AI-related discrimination or other human rights violations;**

(j) **Governments and technology companies should provide legal information in an explainable manner to groups in vulnerable situations as to their rights and mechanisms for redress when subjected to violations related to the use of AI;**

(k) **Invest in data literacy education for the general public to address AI deep fakes, disinformation and hate speech, so as to create more resilient societies;**

(l) **Recognize the right of all persons to pursue self-realization as a key criteria of an AI international solidarity by design approach, with processes addressing, inter alia, access to education, housing, employment, health care;**

(m) **Respect the autonomy of data subjects when collecting and processing data;**

[112] See https://digitalrights-check.bmz-digital.global/; see also https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai and https://ieeexplore.ieee.org/document/8058187.

(n) **Governments and tech companies are expected to inform data subjects of their data protection rights and make all efforts necessary to ensure that everyone has access to data protection rights, including the right to delete, right to access and right to opt out;**

(o) **Governments and tech companies processing data beyond the original purposes (secondary use of data) should seek free, prior and informed consent from individuals, regardless of how the data were obtained; security authorities should not target international solidarity activists with AI tools for counter-extremism or for counter-terrorism activities, in compliance with legal standards relating to necessity, proportionality, legality and non-discrimination;**

(p) **Establish a legal framework and systematic procedure to identify the scope of the public interest and categorize the urgency and necessity of the processing of data; persons should retain opt-out rights to protect their privacy;**

(q) **States and companies should ban the creation of synthetic or manipulated, non-consensual intimate deep fakes, in accordance with a gendered AI international solidarity.**

58. **The Independent Expert recommends that civil society:**

(a) **Continue to advocate inclusion in the creation of ethical AI-powered content moderation tools to prevent, identify and remove online violence and discrimination against women and persons subject to intersectoral discrimination;**

(b) **Continue to provide examples of algorithmic discrimination to the United Nations, national governments and technology companies or AI developers in order to provide recommendations to improve practice;**

(c) **Workers' unions should promote the inclusion and consultation of workers in AI co-design, training and risk assessment, in addition to access to digital literacy.**

59. **The Independent Expert recommends that companies:**

(a) **Establish concrete mechanisms to amplify the voices of traditionally marginalized groups so as to promote a safe, inclusive online environment that recognizes mutual respect for diversity as a principle, including consultations that are available in different languages, and the creation of grants to facilitate the participation of stakeholders from the global South;**

(b) **Technology companies and AI developers should adopt AI international solidarity aimed at respecting planetary boundaries by design; conduct independent, transparent, continuous environmental risk assessments; promote energy-efficient algorithms; support sustainable digital development by adopting the use of renewable energy; provide AI solidarity economy financing for the creation of data to support environmental and climate change knowledge of common areas accessible to the public (such as the seabed and the ocean); manage e-waste; and pursue circular economy principles that are inclusive of local communities;**

(c) **Funders of AI and AI developers should commit to AI international solidarity by conducting preventive and systematic human rights assessments to identify the risk of use algorithmic bias in violation of non-discrimination and equality, in the entire life cycle of data, as well as to require the inclusion of direct and indirect civil society stakeholders (including structurally silenced communities) in an independent oversight mechanism to identify and respond to anti-gender equality and anti-democratic agendas;**

(d) **Technology companies should create preventive and reactive solidarity mechanisms to address human rights violations related to the use of AI, such as disinformation and misinformation campaigns that result in societal violence or in the harassment, surveillance, discrimination or disproportional censorship of structurally silenced communities;**[113]

(e) **Technology companies should provide rapid response teams that are adequately resourced and have the decision-making power to provide timely response to complaints and offer solutions.**

60. **The Independent Expert recommends that the United Nations:**

(a) **Create a digital solidarity platform through which civil society organizations (from the global North and the global South) can exchange international solidarity ideas, and in addition, discuss AI inclusion and non-discrimination policies and practices, challenges and innovations with technology company representatives, AI developers and State officials, in line with the recommendations outlined in the Independent Expert's report on civil society and international solidarity;**[114]

(b) **Create a fund to support the establishment of large language models for data in all languages, so as to protect cultural diversity.**

[113] The recommendation correlates with the United Nations Global Principles for Information Integrity. "Elevate crisis response. Working with stakeholders operating in high-risk areas, establish early warning and escalation processes with accelerated and timely response rates in contexts of crisis and conflict. Establish mechanisms to enable prominent, timely access to reliable, accurate information that serves the public interest."

[114] A/HRC/56/57.