"Governing AI for Humanity": APC's feedback to the interim report of the UN Advisory Body on Artificial Intelligence

Association for Progressive Communications

MAY 2024

Table of contents

Introduction

This text presents comments and observations provided by the Association for Progressive Communications (APC) in response to the interim report draft published by the United Nations' Advisory Body on Artificial Intelligence.[1]

APC understands there is a global governance deficit in relation to artificial intelligence (AI) and firmly reiterates that human rights should be at the centre of any improvements looking forward. In particular, the rights of women and gender non-conforming individuals, as well as those of marginalised communities and groups, should lead decision making in this area, so that the benefits of AI are proportionally distributed, and its risks do not immoderately target specific individuals and groups.

With the above in mind, we propose the following comments for the *Interim Report: Governing AI for Humanity* to better address, with a gender perspective, how we can harness the potential benefits of AI for humanity, the risks and challenges that AI presents now and in the foreseeable future, and the current global governance deficit regarding AI.

First, we provide remarks on the general report, followed by feedback to the three main specific sections ("Opportunities and Enablers", "Risks and Opportunities" and "Guiding Principles and Institutional Functions for AI Global Governance").

Remarks on the general report

- The report describes how AI is necessitating the design and implementation of governance, "not merely to address the challenges and risks but to ensure we harness its potential in ways that leave no one behind" (paragraph 4). We consider that this captures the premise that must knit the whole report and upcoming strategies together.

- We applaud the overall balance of potential positive and negative impacts in the report structure.

---

[1] https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf

- Discussions have brought up the possibility of creating a dedicated organ within the UN to tend to AI governance. While we agree that having an organ specifically dedicated to governance matters is important, and that the UN would constitute an adequate system in which to embed such an organ, we are concerned about creating more bureaucracy and about budgetary constrictions. Attention should be given to assessing existing processes and mechanisms within the UN that already develop related functions, including to address their ongoing shortcomings (which, in our view, are too often related to lack of political commitment and funding). Creating another layer of governance without addressing existing challenges will only promote additional decentralisation of resources and efforts, also generating legal uncertainty. Improved governance for AI is needed, but possibly, related functions could be taken over by some existing bodies; better coordination among relevant bodies is urgently needed.

On the section "Opportunities and Enablers"

- Elaborate clearly on the possibility of investing in governance and regulations that allow us to guide the development of AI so as to make equity, diversity, rights and responsible ethics the mainstream norm for upcoming and emerging AI products, services and solutions, and hence of social reality.

- Explicitly include and elaborate on the role of women, girls and people of diverse genders and sexualities not only as receivers of impacts, effects and outputs of AI, but as fundamental contributors to the development of AI solutions to specific problems. Women and girls, in all their diversity, must be involved in the creation, development, implementation and governance of AI systems and technology.

- The report describes an example of AI's potential to aid in the processing of "citizen-reported data on effects of hyperlocal climate change" (Box 1); we recommend insisting on the priority of participative processes for both citizen input and input organised by civil society for AI technologies in order to, like the report posits, "reflect the lived experience of community members and local decision makers" (Box 1) and avoid "technical silos" (Box 1). These concerns, needs and demands must be constantly revised and updated.

- Elaborate on the opportunities to contribute to progressive communication and interconnection through the potential of AI to enhance accessibility to individuals and groups that have been structurally and historically marginalised and consequently excluded from the full exercise of their rights, aggravated by

emerging and ongoing technological developments for which, as the report indicates, "the pace of change is only going to increase" (paragraph 5), with the due implications for AI governance.

- Extend the analysis on the basic material foundations of a digital economy related to the AI divide (paragraph 18) to consider the implications for infrastructure and access in conflict zones and complex settings, such as humanitarian missions, refugee camps, war areas and migration hubs.

- Promote the importance of open source resources for AI development, presenting it not only as a budget and accessibility alternative (paragraph 19) but as a vital network and platform for inclusive collaboration that enables the full enjoyment of digital rights for all.

- Include considerations for cross-sector collaboration and intersecting goals for AI development and application, going beyond the current paragraphs that contain a comparative analysis of different sectors.

- Incorporate the concept of human-centred design to guide AI development and governance approaches with a focus on user needs and rights, and which set diverse, plural and real experiences as the main axis of solution-planning.

On the section "Risks and Challenges"

- Incorporate specific considerations to address, prevent and respond to risks and challenges faced by individuals and groups with intersecting and compounded vulnerabilities to malevolent uses of AI, including women, people of diverse genders and sexualities, and racialised, socioeconomically and/or politically marginalised communities

  - These considerations must be developed in collaboration and consultancy with technical experts and civil society representatives and be centred in the lived experiences of victims and survivors of AI malevolent targeting.

- Elaborate on the point of risks to societal trust (paragraph 27), including the implications for evidence admissible in courts of law, particularly for cases in which the privacy, safety and integrity of a person are at risk, considering the

differentiated risks for women and girls, in all their diversity, as well as for people with reduced access to digital, technological, legal and/or political literacy.

● Include specific dispositions, or plans to promptly develop a strategy, to engage in multilevel dialogues to identify emerging biases and results of AI uses and applications (paragraphs 30-35 and Box 3), with special attention to differentiated impacts for distinct populations and diverse contexts globally, to develop mechanisms for early detection and responses to effects that may interfere with the rights of a person or group, including the right to equality, privacy, security, freedom and protections. Measures should consider the following:

  ○ AI and machine learning should be trained on thick data for diagnostics and analysis, rather than big data for deterministic correlations.
  ○ While AI offers potential for collective benefit, it currently replicates the inequalities and oppressions of our world. AI is presently designed by biased humans, using biased datasets that fail to represent the diversity of contexts and people, leading to discrimination and marginalisation based on racial, ethnic, religious, gender identity and other criteria. This paradigm must be fought and modified, harnessing the opportunities of AI as part of this change.

● Regarding how there can be no space for accountability deficits (paragraph 35) in developing AI governance at all levels, call for further exchange of ideas to promote clear, feasible and meaningful commitments on this point.

  ○ We applaud the clarity with which this is presented in the current report and encourage developing this issue along this line, including a call to the private sector and to governments to engage in transparency practices that align with human rights and are human-centred.

On the section "International Governance of AI"

General observations

● The need to educate the public on AI governance and how it interacts with individual and social rights will require further dialogue and development.

- As highlighted in the report, "governance is not an end but a means, a set of mechanisms intended to exercise control or direction of something that has the potential for good or ill" (paragraph 75). Governance architectures and mechanisms only make sense if the final objectives are well defined and guide decision-making. We call on companies and states to refrain from "techno-solutionism" approaches to AI, as if AI by itself had the potential to address systemic and historical global challenges without a clear and intentional framing that puts people, not technology, at the centre of any developments.

- The report states that "more inclusive engagement is needed [...] as many communities – particularly in the Global South or Global Majority – have been largely missing from these discussions, despite the potential impact on their lives. A more cohesive, inclusive, participatory, and coordinated approach is needed, involving diverse communities worldwide, especially those from the Global South or Global Majority" (paragraph 7); greater attention should be given to designing concrete mechanisms to ensure such decentralised participatory schemes.

- In its next iteration, the AI Advisory Body's report should more concretely identify which specific measures should be adopted to address the risks posed by the concentration of AI technological knowledge and power. Current developments at the edge of technological progress are only possible as a result of huge concentration in data collection, computational capacity and extremely high financial investments..

- AI is and will be disruptive for society and creative governance structures must be designed to tame "profit" as the main goal in corporate behaviour. Safety and profit should not merely "be aligned", as has been defended within the sector. Safety must take precedence, and the more diversity is brought to the sector, the more this can be addressed through decentralisation and a variety of actors able to influence and build the AI landscape.

- AI has been used to increase automation and amplification, and therefore to catalyse, online harmful behaviour, be it related to speech or other illegal activities. Despite being a new technology, and one whose future developments are possibly unforeseeable today, AI exacerbates old challenges that have a real impact on people's lives today. At the root of these threats is the exploitation of data collection schemes that allow for the accumulation of personal data with disregard for privacy and other user rights. This issue must be addressed by any AI governance proposed.

- Clear "do no harm", transparency, safety and responsibility criteria, built on human rights standards, should be the basis of governance functions. Not only compliance, but also clear and effective accountability schemes, must be developed. Institutional functions related to the governance of AI must:
  - Foresee developments and address long-term implications.
  - Promote harmonisation of norms and rules.
  - Promote international cooperation and solidarity for distributing AI benefits.
  - Contribute to the decentralisation of knowledge and power.
  - Monitor and address risks and coordinate responses (licensing and impact assessments are examples of tools in this regard), which may include mitigation or halting of overly risky technological developments.

On the guiding principles to guide the formation of new global governance institutions for AI

- Regarding how a "do no harm" principle and self-regulation are necessary but not sufficient (paragraph 49), include considerations to ensure decisive commitments by governments to incorporate gender-responsive digital policies and governance guidelines for the development and applications of AI, to harness the potential of AI to achieve gender justice at all levels and across all UN member states.

- Incorporate the concept of equality-by-design principles as priorities to consider for the design, development, application and review of AI systems.

- Regarding Principle 1:

  - Elaborate on technical analysis with an intersectional gender perspective. Datafication, automation and algorithmic mediation and responses have differentiated impacts on women, girls, and people of diverse genders and sexualities, impacts that can turn into online gender-based violence such as hate speech and gendered disinformation. Studying situated knowledge(s) and embodied experiences of women and girls across the globe will help understand algorithmic harms as a more complex phenomenon, instead of a straightforward experience.
  - Make clear calls for governments to demonstrate a commitment to development of gender-responsive national digital policies and frameworks that harness AI for gender equality goals.

- ○ Make clear plans for international financial institutions to create new mandates that enable and support developing countries to build digital capabilities, including AI-based, for domestic development and gender equality.
- ○ Elaborate on the role of public financing to guarantee that universal, equal and safe access to AI infrastructures is available to all women, considering the urgency to "to turn digital divides into inclusive digital opportunities" (paragraph 47) through feminist infrastructures and public innovation systems that provide affirmative and corrective measures for technological gaps, including access and capacity building (paragraph 47).

- ○ Elaborate on differentiated risks for women and feminist activists considering that:

  - ■ AI impacts privacy, security and freedoms, with implications of exclusion and bias. If designed for surveillance, it can threaten the safety and work of women human rights defenders.
  - ■ Disregard for and underinvestment in content moderation, human and algorithmic, has led to the proliferation of sexism, gendered disinformation and other forms of online gender-based violence.

- ○ The AI governance we build, together and shared, must be directed toward a digital realm for the well-being of people and the planet, based on the shared goals of a feminist internet where AI projects and tools can be assessed through values such as justice, dignity, intersectionality, agency, accountability, autonomy, non-binary identities, cooperation, decentralisation, consent, diversity, decoloniality, empathy and solidarity.

- ● Regarding Principle 5, emphasise and further elaborate on the following:

  - ○ While the report identifies the need for space for different regulatory approaches to co-exist reflecting the world's social and cultural diversity (paragraph 40), accountability and transparency must be demanded from corporations who build and sell AI, to ensure AI development and deployment are rooted in international human rights frameworks and do not erode democracy, rights and labour standards.

On the institutional functions that an international governance regime for AI should carry out

- Regarding Institutional Function 1:

  - In relation to the assessment of future development of AI, the report mentions that AI governance must "reflect qualities of the technology itself and its rapidly evolving uses – agile, networked, flexible – as well as being empowering and inclusive, for the benefit of all humanity" (paragraph 5). The report must be centred on the fact that such characteristics must be grounded in human rights and public interest principles to avoid diverting from the urgent need to address persistent challenges related to structural inequality and inequity online and offline.
  - The need for human impact and human rights due diligence should apply to all tech innovation, guided by the precautionary principle. Risky development beyond a pre-defined threshold should be halted.

- Regarding Institutional Function 4:

  - Elaborate further on the role of multistakeholderism and multilateralism in shaping AI governance. Multistakeholderism should be ensured through built-in mechanisms for participation, not merely as an add-on high level goal.
  - Elaborate on plans and guidelines to engage with different publics, including diverse forms of organised civil society and communities, to question who stands to benefit and to be harmed by different outcomes of AI development.

- Institutional Function 7: Compliance and accountability based on norms
  - The focus of this principle has to be on responsibility and accountability and not on the promotion of AI per se. Any technology development (including AI) should comply with certain standards, and actors (companies or governments) developing them should be accountable for eventual harms.
  - The precautionary principle should be considered. Mitigation should not be seen as the only response, as there are risks that are unacceptable.