Version 1

Human Rights – Universal Standards for a World with Artificial Intelligence (A.I.)

How can the Human Rights Council (HRC) promote a human rights-based approach to A.I. and help avoid a further deepening of the digital divide?

Contents

Non-paper presented by the co-facilitators and Permanent Representatives of The Gambia, Luxembourg and the Republic of Korea to the President of the Human Rights Council – April 2024

Introduction

Artificial intelligence confronts humankind with a serious conundrum and with ambiguous feelings about, on the one hand, the promise of undreamt-of opportunities for the betterment of human life and development and, on the other hand, frightening perspectives, up to threats to our very existence, as a species.

Today technological development is happening at an ever-higher speed, driven not only by readily available know-how and natural resources, but also to a large extent by sky-rocketing market value and profit margins of those private sector actors that have taken the lead and the economic lion share in the digital realm with little consideration, responsibility and accountability for the consequences of inequity, widening divides and risks to privacy, ethics, human rights and dignity for all.

On that backdrop, and while, in the past, digital developers of A.I. or other new technologies were used to work and innovate with relatively little or no legal and administrative restrictions, oversight or requirements of accountability, for some time now and increasingly, requests are voiced for tighter, more responsible and accountable regulation, within a pragmatic multilateral governance architecture and framework (MGAF), with human rights and human dignity at its core. This MGAF must be guided by new mechanisms, including digital rapporteurs, experts and investigators. Such mechanisms can facilitate evidence-based oversight to enhance trust and transparency and establish conduct and standards, in order to put in place smart guardrails for a humanist, trusting, safer approach to human rights compliant algorithms. Such an approach shall enable A.I development, platforms, tools and services, with related activities and products, that will respect and be accountable for the privacy, ownership, access and use of data for all users everywhere. It will also ensure and guarantee the human rights and dignity of all users, while adhering and being compliant to international law, existing protocols and conventions, that are adapted, responsive and applicable to the digital space. Given the rapid advancements and complexities in this field, calls for action and response in the context of multilateralism are getting louder from the private sector, CSO'S, academia and other actors.

However, since the technological progress in the digital field advances with unprecedented swiftness, regulators and lawmakers, both at national and international levels, are struggling to keep up with the constantly changing environment. How are they supposed to deal effectively with both the opportunities (I) and the risks (II) inherent to A.I. for human rights and human dignity?

In addition, the ever-deepening digital divide (III) between those who have access to A.I. and those who do not, is putting additional pressure on the universal validity of human rights and human dignity. Indeed, the challenge to equitably manage the up- and downsides of A.I. is exacerbated by the global, yet uneven consequences of accelerating A.I. development and deployment throughout all of humankind's activities.

Confronted with this new multifaceted and fast-changing reality, what role is there to play for multilateral bodies and for the Human Rights Council, building on its existing initiatives, in particular (IV)?

I. Opportunities for Human Rights

A.I. development and deployment aiming at the good for the individual and at the collective global public good, have the potential to help better promote and protect human rights, for example:

- the right to health, through A.I. triggered breakthroughs in medical research, more accurate disease detection and diagnosis, data-based decision making, more personalized treatment in real time and the use of surgical robotics;
- the right to food, through A.I. empowered progress in agriculture, data science-based prediction models; plant disease detection; smart weather stations; wildlife monitoring;
- the right to education, through increased accessibility and availability of education, especially for women and girls, as well as persons in vulnerable situations; through enabling more personalized learning, quicker and more detailed feedback on performance, enhanced efficiency for teachers and access to vast knowledge resources;
- the right to participate in government and free elections, through data-based decision making for policy and legislative processes; predictive data-based simulations, improved connections for citizens with public administrations and infrastructures and better access to public services;
- the rights of persons with disabilities, through enabling them to identify accessible routes around their location; through enhancing the personal mobility of visually impaired persons with navigation tools powered by A.I.;
- the right to seek, receive and impart information, through allowing individuals to collect and receive information much more efficiently and through providing opportunities for human rights-based responses to disinformation and misinformation;
- the rights of human rights defenders (HRD), through allowing broader and more effective monitoring of violations and abuses of human rights, and helping them gather and analyze information at a much quicker rate;
- The right to adequate housing, through improving disaster risk management, facilitating in-home, residential and other community support services and through improving accessibility of the housing environment.

Initiatives like the ITU's "A.I. for Good – Global Summit" (30-31 May 2024) document positive examples of beneficial A.I. use.

II. Risks for Human Rights

Just like for the potential opportunities inherent to A.I., the following list of risks linked to A.I. is by no means an exhaustive one. However, not taking a human rights-based approach to the development, the deployment and the use of A.I. can entail serious threats to the enjoyment of a large number of human rights, such as:

- the freedom from discrimination, through bias in hiring practices, in law enforcement and in criminal justice and access to services, thereby violating the right to non-discrimination.
- the right to own property, through misuse of personal data and inadequate landownership registration;

- the freedom from interference with privacy and the right to peaceful assembly and association, through mass surveillance; through A.I. and new technologies that will infringe upon human rights, human dignity and threaten democratic principles and violate international law; through A.I. and new technologies and use for indiscriminate and targeted surveillance and facial recognition technologies that may be weaponized; through the potential misuse or unauthorized access to user data;
- the freedom of expression and information, through algorithms used by social media platforms and search engines that may influence and augment substantially the information and viewpoints to which users are exposed; through targeted disinformation and deep fakes; through potentially restricting the right to freedom of expression and access to diverse sources of information; through censorship;
- the right to participate in government and free elections, through hacking and election meddling.
- the right to desirable work and the right to an adequate standard of living, through automation and job displacement (increasing automation of tasks through AI and new technologies has the potential to displace workers, leading to job loss and economic insecurity); through biased job descriptions;
- the right to life itself, through unregulated and unsupervised use of Legal Autonomous Weapon Systems (LAWS);
- Risk of inequality of arms in judicial adjudication between the State and accused persons;
- Unethical use of A.I.: Looming ethical risks may take the form of lack of transparency, accountability and oversight mechanisms needed to address ethical concerns related to A.I.;
- Discrimination in access to Education, Digital Skills and STEM Development: As AI and new technologies reshape labour markets and the future of work, we must ensure that individuals have access to meaningful advanced digital literacy and competence, STEM education and training programs. The development of such competencies needs adequate learning infrastructure that equip learners to thrive in a digital economy, both at consumer and producer levels, through productive capacities and competencies for meaningful value contribution and participation in the global digital market economy ecosystems and value chains. This is essential for upholding the right to education and the right to work with readiness and competencies required by the future of work.

III. Closing the Digital Divide

The digital divide is first and foremost threatening the right to equality and the freedom from discrimination, as well as the right to development. As AI and new technologies become more prevalent, disparities in access and digital literacy and skills could widen, exacerbating existing inequalities and limiting the enjoyment of human rights for marginalized populations.

The digital divide is a cause of division and reflects and amplifies inequalities, as it segregates in terms of:

- the access to infrastructure, data and investment that are indispensable for the use of modern technologies and A.I.;
- the use of these technologies, through variable levels of digital literacy and skills;
- the different levels of quality of use (quality of the technologies, quality of the information and data available).

The digital divide comes in different disguises. It presents geographic, social and generational aspects. It creates and exacerbates differences between countries from the North and the South, between urban and rural areas, between rich and poor, young and old, and it is often gender based. And what may seem to be reason for hope through the access to new technologies and A.I. in one part of the world may very well be perceived as a risk through the absence thereof in other parts.

As such, the ever-deepening digital divide is a barrier to the equal enjoyment of i.a.:

- the right to education,
- the right to a clean, healthy and sustainable environment (climate change and global warming),
- the right to desirable work,
- the right to social security,
- the right to seek, receive and impart information.

To bridge that digital divide or, at least to stop its constant widening, will take an extraordinary collective effort of many stakeholders, public and private, to share the benefits of the new technologies and A.I. and to jointly mitigate the risks of their misuse. Indeed, the danger of collective inaction is very high.

The ultimate, almost apocalyptic form of digital divide would oppose humankind, on the one side, and machines empowered by artificial general intelligence (AGI), on the other. Some argue that the danger inherent to the non-alignment of A.I. to the shared values and goals of humanity may well amount to a threat to the right to life, in case of the occurrence of a singularity, when technological progress and growth spring out of human control and turn against humanity itself and threaten human dignity. In order to prevent the realization of such a dark perspective, calls for containment and harnessing of potentially dangerous autonomous A.I. are getting louder.

IV. A role for the Human Rights Council

Global challenges are best addressed by the international community collectively and on the basis of a universally accepted set of references. Human rights, as one of the three pillars of the UN Charter, constitute such a body of references. Furthermore, human rights are overarching in nature, insofar that the international community's political and operational efforts in the fields of peace, security and development aim at respecting and protecting human rights and assuring human dignity.

Confronted, at the global level, with both the opportunities and the risks that new technologies and A.I. present for human rights and human dignity, the Human Rights Council (HRC), as the main intergovernmental body within the United Nations responsible for human rights, is well advised to examine and discuss these eminently important issues, and take adequate action; all the more that the digital divide, in its many forms, is exacerbating the negative impact of new technologies and A.I. on human rights and human dignity.

As a matter of fact, the HRC has already begun taking action to address human rights implications of A.I., notably through the consensual adoption of the resolution on new and emerging digital technologies (53/29) and other relevant resolutions, such as the recent adoption of the resolution on the right to privacy. In addition, the HRC's special procedures have increasingly drawn attention to various opportunities and areas of concern in the use of A.I., as it relates to their respective

mandates, including with regard to law enforcement, content moderation, the exacerbation of discrimination and marginalization of vulnerable populations and the violation of privacy rights.

Against the backdrop of various international initiatives in this area, such as the ongoing negotiations on the Global Digital Compact and the upcoming WSIS+20 process, it would be appropriate for the HRC to deliver more, on that basis, for a human rights-based approach to the accelerating and potentially harmful development and deployment of digital technologies, including A.I. – The HRC should strengthen its capacity to do so.

The HRC's mandate, flowing from A/RES/60/251, and the procedures defined in its Institution building package provide the legal basis that warrants such a proactive role of the HRC in the ongoing international efforts for a human rights-based approach to establish pragmatic guardrails with universally acceptable standards and HRC mechanisms to the accelerating and potentially harmful technological development.

The HRC and its President and Bureau, in close cooperation with the Office of the High Commissioner for Human Rights, can chose from a series of instruments, tools and formats to address the issue of new technologies, A.I. and the digital divide:

- by continuing to adopt resolutions, as well as appropriate, declarations, presidential statements and decisions, related to A.I. and other new and emerging digital technologies, and mandating the OHCHR to continue its valuable work in this field, including producing reports pertinent to this topic;
- by elaborating further on the constitutive elements of a human rights-based approach to the governance of new and emerging digital technologies, including A.I.;
- by providing authoritative guidance on applicable human rights norms, building upon the extensive existing work of the OHCHR and treaty bodies on this issues, as well as related various resolutions of the HRC on this topic;
- by making A.I. related recommendations to the General Assembly for the further development of international law in the field of human rights applicable to the digital space, A.I. and new technologies, including on the need to pragmatically regulate new technologies and A.I. with a human rights-based approach, as well as on human rights-based governance architecture and structures for new technologies, data and A.I.;
- by serving as a convener, a forum and a platform for constructive dialogues on thematic issues related to new and emerging digital technologies and human rights, as well as on ways to promote and elaborate further on a human rights-based approach to new and emerging digital technologies;
- by serving more specifically as a forum for dialogues on thematic issues on the need to strengthen the capacity of the HRC to promote and elaborate further on a human rights-based approach in order to address the implications of new technologies, including A.I. in a holistic, comprehensive and inclusive manner;
- by addressing A.I. related aspect in its work under its agenda item 10, most notably on access, data and digital public infrastructure as well as digital utilities and social services, capacities and investment;
- by including A.I., data and digital aspects into the UPR process, insofar that they are conducive to the implementation of member states' human rights obligation or represent an obstacle to do so;
- by mainstreaming A.I. related aspects into the mandates of special procedures, insofar that these aspects are helpful in fulfilling these mandates or represent an obstacle to do so. Given the increasing complexity, advancement and centrality of the digital and A.I. in human lives,

reshaping of societies, culture, laws and anticipated risks and economic implications, the HRC may intentionally consider appropriate response by introducing new digital mandate holders specific to data, new digital technologies and A.I., in the context of human rights;
- by defining taxonomies and ontologies to help integrating human rights considerations at the technical levels of developing new technologies and A.I.;
- by following up to the request for a feasibility study with a view to establishing an integrated and accessible digital information and knowledge management system related to the Human Rights Council's tools and activities, as expressed in Presidential statement A/HRC/PRST/OS/17/1.

As per its mandate, the HRC should promote the effective coordination and mainstreaming of HR in the UN system. In that respect, the HRC should welcome policy documents on A.I. by other organizations and entities that stress the centrality of human rights in their approaches, such as:

- the Zero Draft of the Global Digital Compact [27 mentions of human rights on 13 pages];
- the UNESCO Recommendation on Ethics of A.I. (adopted by acclamation by 193 member states in Nov. 2021);
- the report of the UNSG's High Level Advisory Board on A.I.;
- the European A.I. Act [soon to be adopted];
- the Council of Europe's framework Convention on A.I., Human Rights, Democracy and the Rule of Law [soon to be adopted];
- the OECD Principles on Artificial Intelligence ;
- the UNGA landmark resolution on « Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development" (March 21st, 2024).

Furthermore, the HRC, through its President and mandated co-facilitators, should reach out to other relevant organizations and entities with mandates, expertise and experiences in the fields of new technologies, A.I. and the digital divide, in their respective sectors, first and foremost:

- International Telecommunication Union (ITU);
- World Intellectual Property Organization (WIPO);
- World Health Organization (WHO);
- International Meteorological Organization (IMO);
- United Nations High Commissioner for Refugees (UNHCR);
- Office for the Coordination of Humanitarian Affairs (OCHA);
- International Organization for Migration (IOM);
- World Trade Organization (WTO);
- United Nations Conference on Trade and Development (UNCTAD);
- United Nations Economic Commission of Europe (UNECE);
- International Committee of the Red Cross (ICRC);
- International Federation of the Red Cross and the Red Crescent Societies (IFCR);
- Geneva Graduate Institute;
- ETH, Zurich;
- University of Geneva;
- University of Zurich;
- Diplo Foundation;
- Geneve Internet Platform;
- GESDA;

- IDAIR- International Digital Health and A.I. Research;
- Cyber Peace Institute;
- Microsoft, Geneva;
- UNESCO;
- World Bank;
- FAO;
- Regional Groups of the United Nations.

The HRC, through its President and members, should reach out to representatives from the private sector and the civil society.