

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

JEDDAH DECLARATION ON THE GUIDING PRINCIPLES OF ARTIFICIAL INTELLIGENCE GOVERNANCE AND HUMAN RIGHTS PROTECTION

23rd Regular Session of the OIC-IPHRC

Jeddah, 4th July 2024: The Independent Permanent Human Rights Commission (IPHRC) of the Organization of Islamic Cooperation (OIC), during its 23rd Regular Session, held a thematic debate on the subject of *'Impact of Artificial Intelligence on Human Rights: Challenges and Opportunities.'* The OIC Secretary General, Secretary General of the Gulf Cooperation Council, OIC-IPHRC Chairperson, and the IPHRC Executive Director jointly inaugurated the event. The Minister of Justice of the Republic of Gabon, the Assistant Secretary General of the League of Arab States, made opening statements. Team leader of Human Rights and Digital Technology in the UN Office of the High Commissioner for Human Rights (OHCHR); Director of International Cooperation at the International Islamic Fiqh Academy (IIFA); Academic Staff Member of Civil Law at Al Manoufia University of Egypt; Director of Training and Technical Cooperation in SESRIC¹; Chairman of the Association of the Artificial Intelligence of Things Society in the State of Kuwait; Head of the Digital Foresight Sector from the Digital Cooperation Organization in Saudi Arabia; Expert from the Department of Legal Affairs and International Standard in ICESCO², President of Family Watch International, Secretary General of the Family Affairs Council in Riyadh; Deputy Secretary General of Technology Development in Ministry of Science, Technology and Innovation in Malaysia; Commissioner from the African Commission on Human and Peoples' Rights; Director of Human Rights Department in the League of Arab States, and Deputy Director of the National Center for Human Rights of the Republic of Uzbekistan, participated as expert panelists. OIC Member and Observer States and their National Human Rights Institutions (NHRIs) actively participated and contributed to the discussions.

Based on the comprehensive discussion, the Commission:

1. **Stressed** the importance of knowledge, learning, and critical thinking, which can be applied to the field of AI. The core Islamic concepts of *'Free Will'* and *'Intelligence'* that Allah has gifted to mankind as His *'Viceregent'* on this earth for *"promotion of virtue and prevention of vice"* are key principles to prescribe specific guidelines to regulate the evolution of AI;
2. **Recalled** that Islamic history is replete with a glorious scientific legacy bequeathed by Muslims. It was recalled that great Muslim scholars throughout history have been at the forefront of building knowledge that has formed the cornerstone of today's AI progress, which includes the foundations of the scientific method of *"Algorithm"* which remains central to the emergence of AI;

¹ The Statistical, Economic and Social Research and Training Centre for Islamic Countries

² The Islamic World Educational, Scientific and Cultural Organization

3. **Underscored** that AI is not inherently incompatible with Islam and holds the potential to transform various aspects of human life but cannot replace human beings in form and nature. The intersection of AI, transhumanism, and Islam prompts a nuanced exploration of its implications. Despite its potential benefits, there are concerns about the misuse of AI, especially in instilling values contrary to Islamic principles. A critical approach is essential to ensure alignment with ethical moral and universal human rights principles. Accordingly, all relevant stakeholders should commit to an open, accessible, secure, inclusive, and human-centered digital future by ensuring that the emerging AI technologies reflect universal human values where human rights and fundamental freedoms are respected, protected, and promoted throughout the life cycle of AI systems;
4. **Considered** that, in the absence of an agreed-upon definition of AI, it could be described as “*a science and a set of computational technologies that are inspired by, but typically operate quite differently from, the ways people use their nervous systems and bodies to sense, learn, reason, and take action.*”³

Further considered that emerging AI technologies could be categorized into systems that: a) think like humans; b) act like humans; c) think rationally; and d) act rationally⁴. The broad scope of such systems requires that robust ethical standards, governance, and accountability mechanisms must be established to protect human rights and fundamental freedoms based on the universal human rights principles enshrined in international human rights law, UN charter and OIC charter.

5. **Emphasized** that the use of international human rights norms and its well-developed universal standards and institutions to examine AI systems can provide a universal vocabulary and principles to develop guiding principles and normative frameworks to govern AI technologies across a wide spectrum of diverse cultures. To this end, took note of the discussions around the Global Digital Compact and Summit of the Future 2024 and urged that an international human rights normative framework for AI should ensure that such technologies are developed and used in ways that respect, protect, and fulfill human rights with due regard to specific cultural specificities;
6. **Welcomed** the adoption by the UN General Assembly resolution (A/78/L.49) adopted by consensus on the promotion of “*safe, secure, and trustworthy*” AI systems that respect, protect, and promote human rights in the pre-design, design, development, evaluation, testing, deployment, use, sale, procurement, operation, and decommissioning phases. Such systems must be human-centric, reliable, explainable, ethical, inclusive, and fully respected for promoting and protecting human rights and international law, privacy-preserving, sustainable development-oriented, and responsible⁵. Also, it welcomed the unanimous adoption of the same resolution by the United Nations Human Rights Council, a resolution that called on the

³ Stanford Study on AI: Artificial Intelligence and Life in 2030:

https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016fnl_singles.pdf

⁴ Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach (3rd Edition)* (Essex, England: Pearson, 2009)

⁵ UNGA Resolution A/78/L.49 of March 2024:

<https://undocs.org/Home/Mobile?FinalSymbol=A%2F78%2FL.49&Language=E&DeviceType=Desktop&LangRequested=False>

international community to take preventive and supervisory measures regarding AI. The resolution also called for enhancing the transparency of the AI systems while ensuring that data allocated to such technology is collected, used, shared, stored, and deleted in such methods that are consistent with human rights;⁶

7. **Upheld** that the Charter of the OIC highlights the need to “promote cooperation among Member States to achieve sustained socioeconomic development for effective integration in the global economy, in conformity with the principles of partnership and equality” and also to “preserve and promote all aspects related to environment for present and future generations”;
8. **Acknowledged** that AI can be an enabler for realizing human rights, as it offers many opportunities across various sectors with extraordinary potential for scientific discoveries that expand the bounds of human knowledge to develop tools that optimize finite resources and assist people in everyday tasks. It can help transform public health and leapfrog challenges of access in education across the world. By unlocking new potentials in productivity, creativity, and problem-solving, AI significantly contributes to economic growth and societal advancement;
9. **Further acknowledged** that AI indeed poses significant challenges to human rights. It can reinforce biases or expand surveillance; automated decision-making can blur the accountability of public officials. Privacy is a major concern, as AI-driven surveillance technologies, including facial recognition, can lead to unprecedented monitoring and data collection levels, infringing on individuals' right to privacy. The potential for bias in AI algorithms also threatens equality and non-discrimination, as these systems can perpetuate and even exacerbate existing prejudices, affecting decisions in areas like hiring, lending, and law enforcement.
10. **Expressed concern** that AI systems employed in the criminal justice sector to predict future criminal behavior have already demonstrated tendencies to reinforce discrimination and undermine rights, including the presumption of innocence. AI can bolster the development of lethal autonomous weapons to be used indiscriminately to target the civilians. Autonomous weapons and AI in military applications could “have serious consequences for global peace and security.” AI systems' speed, autonomy, and opacity challenge traditional regulation models as ever more powerful systems are developed, deployed, and used. These technologies can potentially cause horrific levels of threats and gross violations to the fundamental right to life and right to privacy on an unimaginable scale requiring robust frameworks to ensure that technological advancements do not come at the expense of fundamental freedoms and protections;
11. **Recognized** that these opportunities and risks of AI for people and society are evident and have seized public interest. They also manifest globally with geostrategic tensions over access to data, computational technologies, and human resources that help develop AI, leading to an imminent AI cold war. Also, the benefits and risks are not equitably distributed with ever-increasing digital divide against the nations and disadvantaged segments. There is a real danger: even if humanity harnesses the positive aspects of AI, then those will be limited to a club of the rich only. Today's AI benefits are accruing mainly to few States and companies;

⁶ A/HRC/53/L.27/Rev.1 adopted on 12 July 2023

12. **Expressed** concerns that many AI applications pose serious challenges to human rights norms and standards because AI could elude and transgress human leadership, effective regulation, and legal accountability. In this regard, AI requires governance, not merely to address these challenges and risks but also to harness its potential in ways that leave no one behind. As such, solutions to these complex global challenges call for cross- and trans-disciplinary collaboration among all stakeholders and a robust interface between science, policy, culture and society in order to build trust in such technologies;
13. **Highlighted** that the emergence and growing influence of AI, most of which is being developed by private stakeholders, presents significant challenges to national and international governance systems, with far-reaching implications over different social, political, environmental, and legal sectors with “profound risks to society and humanity.” Therefore, there is a risk that this emerging technological reality will likely hinder the effective implementation of current human rights legal standards, requiring new appropriate AI mechanisms to mitigate the potential threats to the realization of a wide range of human rights;
14. **Called upon** all governments and, where applicable, other stakeholders to refrain from or cease the use of AI systems that are non-compliant with international human rights law or that pose undue risks to the enjoyment of human rights, especially of those who are in vulnerable situations, and **reaffirmed** that the same rights that people have offline must also be protected online, including throughout the life cycle of AI systems.
15. **Noted** that disinformation and propaganda are serious threats to the fundamental right to freedom of expression as a key human right, which is vital for the development of stable, peaceful and progressive democratic societies. While it is not a new phenomenon, misuse of digital technologies, however, have revolutionized the way false and manipulated information is created, spread, and amplified, often for unjust political, ideological, or commercial motives;
16. **Further stressed** that human rights challenges of AI also include bias and discrimination, including against minorities, women and girls, and people with disabilities. Widely acknowledged as an inherent human right risk in AI models and systems, biases manifest themselves in AI algorithms and the data used to train these algorithms. These are also inherent in the societal contexts in which these systems are developed and/or used. Muslim minorities in several Western societies have already been victims of biases in AI-empowered systems, particularly in countries where Muslims have already been socially profiled and disproportionately discriminated against;
17. **Called upon** the international community to expand its ongoing collaborative efforts towards strengthening global digital governance for current and future generations. In particular, to address the potential impact of AI on human rights based on a proactive and sustainable approach, able to promote an open, free, secure, and human-centered digital future deeply rooted in universal human rights while promoting the achievement of sustainable development goals within the new emerging opportunities provided by AI;

18. **Expressed** concerns that the digital world is currently characterized by many divisions and inequalities, including in areas of access to information, affordability, and digital skills and data collection. Hence stressed that safe, secure, and trustworthy AI is necessary to promote digital transformation while bridging the digital divide by ensuring equitable access to advances in new and emerging AI technologies. In this regard, it called for the promotion of a strong, efficient, effective, and robust multilateral system to enhance cooperation, strengthen capacity building and technical and financial assistance to developing countries to close digital divides between and within countries, including intra and inter OIC group, and support developing countries' effective, equitable, and meaningful, participation and representation in international processes and forums on the governance of AI systems;
19. **Called upon** all governments to formulate policies, including national legislations and local regulations, to enhance the transparency of AI systems while ensuring that data allocated to such technology is collected, used, shared, stored, and deleted in such methods that are consistent with human rights. In this regard, the Commission **urged** the technological community, which is driving the AI systems development, to integrate the human rights framework, as well as cultural particularities, in their business activities to develop positive societal externalities;
20. **Affirmed** that the States have the responsibility to create enabling legislative and legal frameworks, including a mix of voluntary and mandatory policy and regulatory measures, in which the responsibility of AI developers to respect and protect human rights is guaranteed. Also, effective judicial and non-judicial remedial instruments and mechanisms should be in place to address the grievances of the victims or to investigate the allegations of human rights abuses that result from AI applications by relevant stakeholders;
21. **Further reaffirmed** that States have responsibilities to a) protect human rights from harms related to all business activities, including AI applications, and ensure that the public is informed, involved, and given access to remedies when rights are violated: (b) cooperate internationally to address transboundary impacts of activities of AI technologies, developers and users and accordingly safeguard human rights during mitigation activities.
22. **Underlined** that contribution of science, technology, and innovation to sustainable development is a critical source of economic growth and achievement of 2030 Agenda for Sustainable Development (SDGs). AI has a great potential to improve real incomes, enable faster and broader deployment of innovative solutions, support more inclusive forms of participation and more sustainable modes of production, and provide policymakers with powerful planning tools. In this regard, affirmed that human rights principles must inform the development of AI in a way that paves the way for the implementation of the 2030 Agenda across the world, with a focus on the least developed countries to reduce the digital divide in line with the commitments of right to development of nations and peoples.
23. **Urged** the OIC and all its subsidiary organs to engage in a comprehensive interaction at the strategic level to formulate a comprehensive AI ethical strategy aligning Islamic values and global ethical standards. This strategy focuses on ensuring that AI technologies are developed and deployed in ways that respect human dignity, promote fairness, uphold privacy, and

emphasize the need for inclusivity and equity across Member States. Also, it may guide and inform the Member States during negotiations for the finalization of the Global Digital Compact to create a unified framework that respects the peculiar socio-cultural sensitivities of the OIC Group;

24. **Encouraged** all OIC organs to take an active role, within their respective mandates and resources, in forging closer links with national and multilateral science advisory bodies in the OIC Member States to optimally leverage science, technology, and innovation for the realization of the development priorities;
25. **Acknowledged** that the global share of OIC countries in the research and development in AI technologies and digital infrastructure is far below the actual potential which requires identification of priority areas for investment, improvement in access to digital technologies, enhance governmental and non-governmental collaborations to develop adequate regulations for particular national contexts to address the AI challenges and resist the ion from the industry that self-governance is sufficient.
26. **Called upon** OIC Member States, where appropriate, in line with their national priorities and circumstances while implementing their distinct national regulatory and governance approaches, to promote safe, secure, and trustworthy AI systems in an inclusive, non-discriminatory and equitable manner. Also, they may focus on the dissemination of societal and cultural values of Muslim societies that must be protected, especially in areas of education, health and culture, including the intellectual, ethical, and moral protection of children and families from exploitation and deception that they may be exposed to through various AI tools and methods. In this context, it commended the progress made by certain OIC Member States in using AI in public policy, including the government of Kingdom of Saudi Arabia which has gained top spot worldwide for the year 2023 in the Government Strategy Index for Artificial Intelligence, in the global AI classification. Also, applauded the formulation of comprehensive National guidelines on AI governance and ethics by the Government of Malaysia. Accordingly, urged OIC Member States to share their technical expertise in AI governance across OIC⁷.
27. **Highlighted** that some digital health technologies empowered by AI may present challenges to the ability of the State to enforce health regulations and preserve ethics at the national level. For instance, telehealth self-care abortions and prescription of drugs by AI systems may undermine the ability of the family and relevant national authorities to prevent illegal health practices that pose risks to the well-being of families, especially children, rural and undeserved communities. In this regard, family counseling and mental health support empowered by AI must be overseen and guided by ethical considerations and inclusivity, with a particular focus on reinforcing and nurturing culturally sensitive family values in line with the Islamic principles and full respect for sovereignty of individual states.
28. **Encouraged** the OIC Member States to enhance joint efforts for digital cooperation among multiple public and private stakeholders to harness the potential of science, technology, and innovation for the benefit of the Ummah. Particularly, to advance, protect, and preserve

⁷ <https://www.spa.gov.sa/en/N2085216>

linguistic and cultural diversity in line with the shared Islamic values and principles throughout the life cycle of the AI systems;

29. **Further encouraged** international specialized agencies, funds, programmes, other entities, bodies and offices, and related international and regional organizations, within their respective mandates and resources, to contribute to enhance the ability of developing countries, in particular the least developed countries, to address major structural impediments and lift obstacles to accessing the benefits of new and emerging technologies and AI innovation to achieve SDGs, including through scaling up the use of scientific sources, affordable technology, research, and development, and strengthened partnerships.
30. **Urged** the IIFA, within the limits of its mandate, to lead an initiative at the level of Member States to study the AI related jurisprudential aspects and associated implications, in such a manner that helps Muslim communities to avoid the risk of falling into jurisprudential violations, and thus realizes a balance between technological progress and life provisions for them.

31. **Recommendations:**

I- The Commission called upon the United Nations and international community to:

- a. consider coming up with ethical and human rights international frameworks to guide the Global Digital Compact to counter the serious challenges to human rights and govern AI's ethical development and use. These frameworks should prioritize human rights, based on the awareness about the risks involved in establishing AI systems without a legal framework, ensuring AI systems are transparent, accountable, and free from bias.
- b. establish national oversight bodies to monitor the application of AI ethics and safeguards and implementation of relevant laws, and encourage the private sector involved in the development of AI to establish self-regulatory bodies that develop ethical guidelines for AI developers.
- c. implement stringent data protection laws to safeguard individuals' personal information. AI systems should be designed with strong privacy protections to prevent unauthorized access and misuse of data;
- d. accelerate the diffusion of innovation and new technologies and associated know-how, and the transfer of AI technologies to developing countries on concessional and preferential terms, to facilitate access to assistive and generative technologies;
- e. address the major structural impediments, to accessing new and emerging technologies, including by scaling up the use of open science, affordable and open-source technology, research and development;
- f. enhance digital infrastructure connectivity and access to technological innovations through stronger international partnerships to help developing countries effectively participate throughout the life cycle of AI systems and accelerate its inclusive and positive contribution to society;
- g. establish a global fund for research and innovation to build capacity of developing countries to contribute to and benefit from emerging AI technologies and establish governance mechanisms provide technical expertise and means of implementation of AI to support developing countries in achieving the Sustainable Development Goals.

II- The Commission further called upon the OIC Member States to undertake coordinated and accelerated actions in accordance with the commitments made in the international human rights treaties, revised OIC Charter, OIC 2nd Ten Year Plan of Action, SDGs, and UNGPs on Business and Human Rights to:

- a. mobilize high-level political support and commitment to devise laws, policy frameworks, development plans, and affirmative regulatory measures to address AI-related human rights impacts through advocacy, effective oversight, and accountability.
- b. actively participate in the ongoing negotiations for the Global Digital Compact to safeguard the unified OIC positions on sensitive issues such as opposition to sexual orientation and gender identity and protection of children from the potential adverse effects of Comprehensive Sexuality Education (CSE) being integrated into digital technologies, including via international health regulations and treaties that undermine national sovereignty of states to preserve the religious and cultural values of local communities.
- c. promote the importance of family values and the protection of traditional social structures in digital content and regulations. By maintaining a strong, collective presence in these negotiations, OIC Member States can advocate for policies that align with shared cultural and religious values, promoting a digital environment that respects ethical standards and protects future generations. To this end, they may build alliances with other nations that share similar values. This proactive approach will not only reinforce collective stance but also contribute to shaping a global digital framework that is inclusive, respectful, and mindful of the diverse values of all communities.
- d. invest in education and training programs to improve AI literacy among policymakers, developers, and the general public. This encourages in-depth discussion to foster the level of awareness of all stakeholders, in particular educators and public service personnel, about the effects of AI on human rights through interaction with various subject-matter experts and stakeholders in AI in such a manner as to share views and identify challenges related to human rights in this regard;
- e. consider establishing independent national oversight mechanisms to monitor AI applications and enforce compliance with human rights standards. These bodies should have the authority to audit AI systems, hold entities accountable for violations, and provide redress to the victims in an independent and impartial manner. To support the work of such national mechanisms on AI, OIC Member States may consider the possibility to establish a Center for Digital Technology Cooperation at the OIC level, to enhance multilateral cooperation and sharing of best practices, cultivate the benefits of AI and mitigate its risks across countries;
- f. support research initiatives, including human rights impact assessment studies focused on the ethical implications of AI, exploring ways to mitigate risks and enhance the positive impact of AI on human rights;

- g. focus on protecting the rights of vulnerable social segments, particularly, women, children, and persons with disabilities, who may be disproportionately affected by AI technologies. Special attention should be given to preventing exploitation and ensuring equitable access to AI benefits;
32. **Urged** the OIC General Secretariat to consider establishing a **Digital Advisory Unit** to coordinate the formulation of a comprehensive AI ethical and technical strategy in consultation with specialized institutions, including ICESCO, Islamic Development Bank, IPHRC, International Islamic Fiqh Academy and academic research institutions.
33. **Further urged the** AI developers and all stakeholders involved in the design, creation and implementation of AI technologies to respect human rights by ensuring that their models, operations, and practices do not contribute to human rights abuses or violations. This could be achieved through:
- a. **Human rights due diligence** to identify, assess, and mitigate all potential adverse impacts of AI systems on human rights. This involves understanding human rights risks associated with their operations, engaging with relevant stakeholders, and taking appropriate actions to prevent or address any negative impacts. Also, develop a comprehensive human rights-based approach to evaluating AI's impact on human rights situations, which requires implementing human rights standards to ensure that human rights guarantees are applied consistently across all relevant sectors using AI;
 - b. **Transparency and reporting** via the incorporation of feedback mechanisms to allow evidence-based discovery and reporting by end-users and third parties of technical vulnerabilities and, as appropriate, misuses of artificial intelligence systems and artificial intelligence incidents following their development, testing, and deployment to address them;
 - c. **Remedy and accountability** with effective grievance mechanisms to address human rights concerns raised by stakeholders resulting from AI, including employees, communities, and other affected parties;
 - d. **Continuous improvement and feedback** involve raising public awareness and understanding of the core functions, capabilities, limitations, and domains of appropriate civil use of artificial intelligence systems.

**Issued in the city of Jeddah, Kingdom of Saudi Arabia, on 28/12/1445,
corresponding to 4th July 2024.**

