Munich Convention on Artificial Intelligence, Data and Human Rights

(Draft for Public Consultation)

Foreword:

The fundamental premise of international human rights law set forth in the Universal Declaration of Human Rights (UHDR) is that "all human beings are born free and equal in dignity and rights" (UDHR; Art. 1) and that human rights form "a common standard of achievement for of peoples and nations" as stated in the preamble of this document. In the decades following the Second World War, the international community established instruments, regimes and enforcement mechanisms to realize the aspiration articulated in the Universal Declaration. The existing landscape of international law includes the International Covenant on Civil and Political Rights as well as the International Covenant on Economic, Social and Cultural Rights, the Convention on the Prevention of Genocide (1948), coinciding with the evolution of international humanitarian law and protection of civilians in international conflicts. Furthermore, the past decades have witnessed the rise of international human rights treaties protecting the rights of specific populations and groups, including but not limited to the Convention on the Rights of the Child, Convention on the Elimination of All Forms of Discrimination Against Women, Convention on the Rights of Persons with Disabilities or the Convention on the Rights of Indigenous Peoples.

Artificial Intelligence (AI) presents indeed significant opportunities to address global challenges and advance the purposes set forth in the UN Charter and international human rights law, including maintaining international peace and security, developing friendly relations among nations, and promoting social progress and higher living standards. AI has the potential to accelerate progress towards achieving the Sustainable Development Goals, particularly in areas such as ending poverty in all its forms, combating inequalities within and among countries, protecting the planet from degradation, and ensuring that all human beings can enjoy prosperous and fulfilling lives.

However, the introduction of AI in different domains and contexts, including critical infrastructures, education, public administration, human resources, robotics, immersive realities, neurosciences, healthcare, warfare and many more, raises concerns about potential and actual adverse impacts of this technology on human rights. These concerns include issues related to accessibility, algorithmic bias, discrimination, unequal treatment, its potential for misuse against human rights, its dependence on personal data processing, risks related to the accuracy and safety of AI systems, intellectual property rights violations including unauthorized use of copyrighted material and questions of AI-generated content ownership, the complex issue of human-machine interaction and algorithmic accountability, and the potential for these issues to coexist or reinforce one another.

Furthermore, the use of AI can also undermine freedom of expression and access to information, especially in the context of AI-driven content moderation and information dissemination on social media platforms, as well as democratic processes through AI-powered misinformation campaigns or manipulation of public opinion. Moreover, the use of AI can result also in group-specific effects, with particular attention needed for vulnerable populations such as children and individuals with disabilities and reinforce existing biases in the labor markets, education systems, finance or the allocation of public goods leading to exacerbation of existing societal inequalities or the creation of new patterns of discrimination. Human rights experts have also identified adverse impacts within the international AI value chain, such as labor rights abuses, the mental health and psychological implications of content moderators exposed to AI, and the consequences of the exploitation of natural resources essential for the physical components of AI systems, such as gold, lithium, and rare earth elements.

These issues highlight the need for a comprehensive approach to the entire lifecycle of AI systems, as they are intricately linked to human rights, but also to foundational principles of international law and global justice. Furthermore, AI systems need to be seen in a larger context, including the generation, use and curation of data, which is often linked to the exploitation of individuals. Given the different implications and applications of AI as a family of technologies, AI systems have the potential to adversely affect **all** human rights articulated and codified in international human rights law.

The pertinence of such AI-driven human rights violations has been increasingly acknowledged by organizations including, but not limited to, the United Nations General Assembly (UNGA), the African Commission on Human and Peoples' Rights (ACHPR), UNESCO, the Office of the High Commissioner for Human Rights (OHCHR), the Organization for Economic Co-operation and Development (OECD), the Council of Europe (CoE), and the UN Global Compact (UN GC). As stated in the Resolution of the UN General Assembly on A/78/L.49, the global dissemination and implementation of AI has raised alarms on the impact of AI on all types of human rights. While international human rights law provides a fundamental framework for the global regulation of AI, the technology introduces complexities transcending the scope of existing human rights instruments that necessitate new approaches to accountability, transparency, and governance, especially when it comes to ensuring meaningful human control over automated systems. This year (2024), the Council of Europe has engaged with the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, a human rights treaty specifying the meaning of human rights and related principles such as democracy or the rule of law in the context of AI. The work of the Council of Europe follows the trend of integrating human rights principles in domestic and supranational law, for instance in Brazil, the European Union, and the United States.
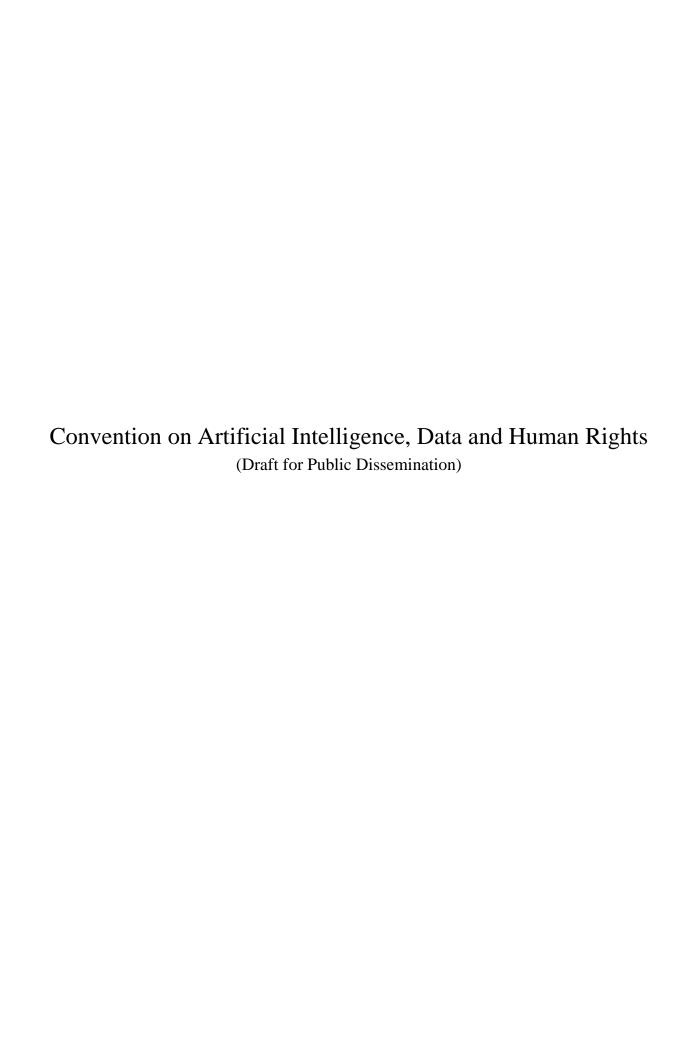
Nevertheless, most UN member states have so far not engaged in a comprehensive legislative response to AI nationally. As witnessed by earlier debates, particularly around business and human rights, such gaps and weaknesses in the enforcement and protection of human rights, for instance in terms of occupational safety or data privacy, tend to be exploited by organizations, particularly when they operate in less regulated jurisdictions. These precedents have given rise to integrating human rights into organizational practice, particularly the UN Guiding Principles on Business and Human Rights and the UN Global Compact.

Furthermore, the international context characterized by different levels of digital literacy, infrastructures, and developmental, political, and economic conditions additionally complicates the mitigation and prevention of adverse human rights impacts created by AI systems. The lack of an international framework makes it therefore difficult to hold actors in the AI lifecycle accountable for adverse human rights impacts in international contexts. The asymmetric response on the international level has already contributed to the emergence of international disparities in the protection of human rights in the AI context through legal instruments. The continuation of this trend, if not prevented by multi-stakeholder approaches to AI governance, will likely reinforce these and result in further legal loopholes.

The international asymmetries in AI regulation, as well as the potential of AI for misuse, necessitate, therefore an international convention, which creates a cross-jurisdictional standard for actors involved in different stages of the AI life cycle, including but not limited to developers, deployers and users of AI and operating in different legal environments. The most recent endorsement of the UN Guiding Principles on Business and Human Rights by the UN General Assembly in Resolution A/78/L.49 is, therefore, a positive signal for integrating human rights, in particular human rights due diligence, in global AI governance. Owing to the specificities of AI and the international fragmentation in terms of AI regulation, there is a need for a binding framework on an international level which enforces human rights obligations on UN member states to engage in a legal response to human rights risks posed by AI.

This Draft Convention on AI, Data and Human Rights is, therefore, a call for international organizations and intergovernmental bodies such as the UN Human Rights Council to engage in an international effort to establish a common ground for AI and data governance aligned with human rights. The document which has been initially drafted by the Institute of Ethics in Artificial Intelligence (IEAI) at Technical University of Munich is the result of discussions during the International Artificial Intelligence and Human Rights Summit in Munich organized by the Institute of Ethics in Artificial Intelligence at the Technical University of Munich, in collaboration with Globethics. The IEAI hosted the International AI and Human Rights Summit inviting more than 50 international policy makers, representatives of NGOs, representatives of business enterprises, and scholars from different pertinent fields to discuss the design and implications of the Convention.

While not an authoritative document in international law, the document provided seeks to ignite the debate on the need for an international convention and provide concrete suggestions with relevant definitions on AI, AI life cycle, actors in the AI lifecycle, deepfakes, data, personal data, as well as domains critical to the enjoyment of human rights. Based on discussions with and the input of experts on AI and human rights and the incorporation of feedback from academia, policymakers, civil society organizations, business representatives, and individuals working for and with international organizations between July and September 2024, the goal of this draft is to start a concrete debate within relevant international bodies, including the UN Human Rights Council, on an international Convention on AI, Data and Human Rights and to define potential areas that warrant further discussion. This includes the definition of domains of AI use critical to the enjoyment of human rights, prohibited practices, and specific rules for global AI operations as well as an enforcement mechanisms which guarantees the realization of the Convention's content.

Convention on Artificial Intelligence, Data and Human Rights
(Draft for Public Dissemination)

Preamble

The Parties to this convention

Reaffirming the international law, in particular the Charter of the United Nations and recalling the Universal Declaration of Human Rights,

Recalling the resolution 70/1 of 25 September 2015, entitled "Transforming our world: the 2030 Agenda for Sustainable Development" and the resolution of the General Assembly 78/L.49 11 March 2024 entitled "Seizing the opportunities of safe, secure, and trustworthy artificial intelligence systems for sustainable development".

Recalling also resolution 37/8 of the Human Rights Council's resolution on "Human rights and the environment", the Human Rights Council's resolution on "The right to privacy in the digital age" (A/HRC/RES/42/15) (2019); and the Human Rights Council's resolution on "New and emerging digital technologies and human rights" (A/HRC/RES/41/11) (2019),

Recalling also the resolution 17/4 of the Human Rights Council on "Human rights and transnational corporations and other business enterprises".

Taking note of the efforts of the International Telecommunication Union to convene the Artificial Intelligence for Good platform as well as the United Nations Educational, Scientific and Cultural Organization of its Recommendation on the Ethics of Artificial Intelligence of 23 November 2021, as well as taking note of the Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework, as endorsed by the Human Rights Council in its resolution 17/4 of 16 June 2011;8 and the work of the Office of the United Nations High Commissioner for Human Rights regarding artificial intelligence.

Taking note also of the Council of Europe's Framework Convention on Artificial Intelligence, Human Rights, Rule of Law and Democracy, as well as many existing national policies, other frameworks and initiatives elaborated by relevant United Nations entities, intergovernmental organizations, including regional organizations, as well as those by the private sector, professional organizations, standards organizations, non-governmental organizations, and the scientific community, related to the ethics, governance and regulation of AI technologies,

Recognizing the profound and dynamic impact of artificial intelligence systems in fields critical to human rights and international peace, in particular the military use of AI, the use of AI in social and societal contexts such as social media, as well as the growing deployment and use of AI in the private and public sectors,

Recognizing also the potential of artificial intelligence systems in promoting human rights and realizing sustainable development,

Recognizing also the need to ensure that the Rule of Law is applied in the adoption and maintenance of appropriate measures to regulate artificial intelligence,

Recognizing also the discourse in academia, research institutions, business enterprises, civil society, policymakers, regulatory bodies, and standards organizations on potential and actual adverse impacts created by or linked to artificial intelligence systems, including algorithmic bias, algorithmic opacity, data use and safety, but also existential risks for humankind,

Recognizing also the integral link between artificial intelligence and data usage,

Concerned by the dual-use potential and misuse of artificial intelligence systems and opposing the use of such systems for repressive purposes, in violation of international human rights law and the Charter of the United Nations,

Respecting the desire and right of some individuals or groups not to take part in all directions of technological progress,

Acknowledging the necessity of a global approach based on international law, human rights and dignity to effectively prevent and mitigate adverse human rights impacts created by or linked to AI,

Mindful of applicable international human rights instruments, such as the 1948 Universal Declaration of Human Rights, the 1966 United Nations International Covenant on Civil and Political Rights, the 1966 United Nations International Covenant on Economic, Social and Cultural Rights,

Have agreed as follows:

Chapter I: General Provisions

Article I: Purpose of the Convention

1. The aim of the Convention is to ensure that activities within the life cycle of artificial intelligence systems are fully consistent with the Universal Declaration of Human Rights and the principles of international law as enshrined in the Charter of the United Nations.
2. Each Party shall adopt or maintain appropriate legislative, administrative or other measures to give effect to the provisions set out in this Convention. These measures shall be graduated and differentiated, as may be necessary in view of violations of international human rights law and principles of international law set forth in the Charter of the United Nations, and shall apply throughout the lifecycle of artificial intelligence systems.

Article II: Definitions

1. For the purposes of the Convention, "artificial intelligence systems" are defined as information-processing technologies that integrate models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction, content, recommendations and decision-making in material and virtual environments. Artificial intelligence systems are designed to operate with varying degrees of autonomy by means of knowledge modeling and representation and by exploiting data and calculating correlations.
2. For the purposes of the Convention, the "life cycle of an artificial intelligence system" pertains to their pre-design, design, development, evaluation, testing, deployment, use, sale, procurement, operation and decommissioning.
3. For the purposes of the Convention, "data" means any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audio-visual recording.
4. For the purposes of the convention, "personal data" is defined as any information relating to an identified or identifiable natural person.
5. For the purposes of the convention, a "deepfake" is defined as AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful.
6. For the purposes of the Convention, the "impacts of artificial intelligence systems" encompass the direct and indirect consequences of artificial intelligence systems for humans on an individual and societal level. This encompasses the generation, collection, curation and evaluation of data used by these systems, the procurement of physical components necessary for their operation, the sourcing and processing of rare materials for manufacturing these components, as well as the intergenerational impact of the development, deployment, use and decommissioning of artificial intelligence systems with its physical components on the environment and the global climate.
7. For the purposes of the Convention, "actors" involved in the AI life cycle include developers, vendors, deployers and users of artificial intelligence systems as well as actors involved in the provision, curation or labeling of data used by artificial intelligence systems.
8. For the purposes of the Convention, "domains critical for enjoyment of human rights" ["critical domains"] are sectors, contexts and use cases of artificial intelligence systems that deserve specific protection under the international human rights law, as a minimum healthcare, occupational safety, employment, education, jurisprudence, law enforcement, public administration and critical infrastructures.

Article III: Scope

Each Party shall adopt or maintain legislative, administrative or other measures to ensure that the conduct of actors involved in the life cycle of artificial intelligence systems is consistent international human rights law and the foundational principles of international law enshrined in the UN Charter with the aim to preserve international peace.

Chapter II: Measures to be Adopted or Maintained by The Parties

Article IV: Alignment of Measures with Human Rights

1. Each Party shall adopt or maintain legislative, administrative or other measures to foster alignment of the development, deployment and use of artificial intelligence systems with the promotion and protection of human dignity, human rights and international law throughout their life cycle.
2. Each Party shall harmonize existing measures dedicated to supporting the research and innovation of artificial intelligence systems as well as accessible and affordable data required for the enhancement of artificial intelligence system conducive to the goals articulated in the UN Charter, particularly the resolution of international problems of an economic, social, cultural or humanitarian character and the promotion of human rights and fundamental freedoms.
3. Each Party shall adopt or maintain legislative, administrative or other measures to promote, support and fund the research, development, deployment and use of artificial intelligence and other digital technologies utilizing data in areas conducive to the promotion of human rights as well as aims articulated by the international community with particular emphasis on the United Nations Sustainable Development Goals. This includes, in particular, the development, deployment and use of artificial intelligence systems in order to address issues the realization of sustainable development goals and overcoming digital divides between and within countries.

Article V: Rights Protecting Individuals from Adverse Human Rights Impacts

1. Each Party shall adopt or maintain legislative, administrative or other measures to protect the rights of individuals in all stages of the life cycle of the artificial intelligence system.
2. Each Party shall adopt or maintain measures to enable individuals to opt-out of using artificial intelligence and to utilize alternative ways of access in domains critical to the enjoyment of human rights and fundamental freedoms.
3. Each Party shall ensure that individuals have the right to an explanation if an artificial intelligence system makes a recommendation or decision in domains critical to the enjoyment of human rights and fundamental freedoms.
4. Each Party shall adopt or maintain measures to ensure that individuals have the right to fair compensation and reparation, remedy and redress if an artificial intelligence system makes an unjustified or erroneous decision that infringes upon their rights. Additionally, these measures must guarantee that individuals have an effective means to lodge complaints and contest decisions made by artificial intelligence systems through judicial remedies.
5. Each Party shall adopt or maintain measures that ensure that individuals are notified when interacting with an artificial intelligence system. Each Party shall adopt or maintain also measures that ensure that individuals shall have the right to be informed when they are subject to a decision when an artificial intelligence system has participated in the decision-making procedure or when the decision was made based on the output of an

artificial intelligence system. Each Party shall adopt or maintain measures that ensure the protection of intellectual property rights through the use of artificial intelligence systems.

Article VI: Definition of Prohibited Practices

1. Each Party shall adopt or maintain legislative, administrative or other measures to address existential risks linked to artificial intelligence systems that undermine human rights and fundamental freedoms as well as international peace.
2. Each Party shall adopt or maintain measures to define legally prohibited purposes of artificial intelligence systems that are irreconcilable with international human rights law, the Charter of the United Nations or international humanitarian law.
3. The Parties shall adopt or maintain measures to prevent the development, deployment and use of artificial intelligence systems that have the purpose of:

   a. impairing the autonomy and ability of individuals to make an informed decision through deceptive techniques,
   b. inciting hatred or violence targeting specific groups or individuals, including specific ethnic groups, women, individuals with disabilities and/or marginalized groups,
   c. undermining the access and integrity of information,
   d. enabling or supporting criminal activities, particularly cybercrime,
   e. creating or distributing deepfakes, that violate the dignity of the individuals represented.

4. Each Party shall adopt or maintain measures to prevent transnational cyberattacks facilitated or exacerbated by artificial intelligence systems within their own territory. The measures shall enhance national cybersecurity frameworks to detect, prevent, and respond to AI-driven cyber threats.

Article VII: Consideration of Different Types of Algorithmic Harm

1. The legislative, administrative or other measures adopted by each Party shall consider the different types of potential harm caused by or linked to artificial intelligence systems. This includes but is not limited to, the effects of automated decision-making, algorithmic bias, algorithmic opacity and the collection, use, curation and evaluation of data on human rights and international law. The measures adopted by each Party shall consider the current state of research with respect to the societal and ethical implications of artificial intelligence systems, as well as risks posed by emerging technologies and data utilization.
2. The legislative, administrative or other measures adopted by each Party shall consider the impact of artificial intelligence systems on different stakeholders, including but not limited to women, children, individuals with disabilities, the elderly, minorities, and other vulnerable groups.
3. The measures adopted by each Party shall be aligned with the principles of human dignity, individual autonomy, transparency, non-discrimination, equity, accessibility, explainability, reliability and safety.
4. The measures adopted by each Party shall encompass all organizations involved in the life cycle of artificial intelligence systems, including but not limited to developers, deployers and users of artificial intelligence systems as well as parties involved in the provision of data of or content moderation of artificial intelligence systems. When engaging in measures directed to entities in the private sector, each Party shall consider the alignment of measures with established frameworks of organizational governance, including the UN Guiding Principles on Business and Human Rights.

Article VIII: Consideration of Different Actors in the Life Cycle of Artificial Intelligence Systems

1. Each Party shall consider in their legislative, administrative or other measures the different functions, capabilities, relative power and domains of actors involved in the life cycle of artificial intelligence systems.
2. Each Party shall consider in their legislative, administrative or other measures the leverage that actors have when mitigating or preventing adverse impacts of artificial intelligence systems on the rights of individuals.
3. Each Party shall adopt or maintain specific measures to foster the deployment and use of artificial intelligence systems in the public sector consistent with human rights obligations, as enshrined in applicable international law and its domestic laws.

Chapter III: Specification of Measures to be Adopted or Maintained by the Parties

Article IX: Risk-based Approach

1. Each party shall align measures corresponding with the level of risks associated with artificial intelligence systems in terms of number of affected individuals, likelihood of occurrence of harm, irreversibility of harm and previous cases of adverse human rights impact created by or linked to particular artificial intelligence systems.
2. Each Party shall adopt or maintain measures that identify general principles which apply to all artificial intelligence systems and establish general principles for the different stages of the life cycle of artificial intelligence systems irrespective of the domain or context of use.
3. Each Party shall consider the specific risks resulting from artificial intelligence systems and other digital technologies in the context of frontier technologies such as neurosciences applications and their implications for the protection and realization of human rights, particularly the rights to privacy and human dignity, within measures to safeguard human rights.
4. Each Party shall adopt or maintain measures, including mechanisms to ensure human end control to address the use of artificial intelligence systems in situations that create irreversible consequences for individuals, particularly in the contexts of physical and mental health and safety, but also when AI solutions determine the access to or security of food and water systems or our environment.
5. Each Party shall adopt and maintain additional measures addressing risks posed by artificial intelligence systems situated in domains critical to the enjoyment of human rights and fundamental freedoms ["critical domains"].
6. Each Party shall adopt or maintain legislative, administrative or other measures to address risks associated with foundational models and large language models.

Article X: General Principles for Artificial Intelligence Systems

1. Each Party shall adopt or maintain measures to enhance accountability and liability for adverse human rights impacts arising from an organization's involvement in the life cycle of artificial intelligence systems and ensure the traceability of all such systems. These measures shall enable the assignment of accountability for algorithmic harm to specific organizations or individuals involved in the artificial intelligence system lifecycle, based on current state-of-the-art practices and technical feasibility.
2. Each Party shall adopt or maintain measures to protect the right to privacy in artificial intelligence systems that process the data of natural persons.

3. Each Party shall adopt or maintain measures to hold developers, deployers and users of artificial intelligence systems accountable for risks and harms caused by these systems.

4. Each Party shall actively collaborate to achieve a common understanding of accountability, cybersecurity, safety and traceability for artificial intelligence systems. This includes developing standards, metrics, indicators, codes of conduct, governance frameworks and identifying best practices.

Article XI: Principles for Artificial Intelligence Systems in Critical Domains

1. Each Party shall adopt and maintain specific measures requiring organizations to conduct impact assessments at pertinent stages of the life cycle of artificial intelligence systems deployed in domains critical to the enjoyment of human rights and fundamental freedoms. These assessments shall inform the basis for deriving specific mitigation and prevention measures to prevent adverse human rights impacts created by or linked to artificial intelligence systems.

2. Each Party shall adopt and maintain measures to ensure that organizations involved in the life cycle of artificial intelligence systems consider the effects of these systems on different groups and populations exposed to the artificial intelligence system, as well as consider the various contexts of its use.

3. Each Party shall adopt and maintain measures to address concerns related to algorithmic bias, discrimination and unequal treatment of individuals in critical areas.

4. Each Party shall adopt or maintain measures requiring organizations to demonstrate that the introduction of artificial intelligence systems does not result in asymmetric impacts on any relevant demographic group, leading to unequal treatment of individuals on the basis of protected criteria such as age, ethnicity, gender, sex, race, religion or political orientation.

5. Each Party shall adopt and main measures to foster the accessibility and inclusivity of AI systems in critical domains, particularly for demographic populations and groups protected by specific instruments of international human rights law including the Convention on the Rights of the Child, Convention on the Elimination of All Forms of Discrimination Against Women, Convention on the Rights of Persons with Disabilities.

6. Each Party shall adopt and maintain measures to ensure meaningful and appropriate human end control of and oversight over artificial intelligence systems in domains critical to the enjoyment of human rights and fundamental freedoms. Human end control and oversight shall be aligned with the convention's purpose in preventing and mitigating adverse human rights impacts created by or linked to artificial intelligence systems.

7. Each Party shall adopt and maintain measures to ensure the transparency of the artificial intelligence system in domains critical to the enjoyment of human rights and fundamental freedoms. The aim of embedding transparency in measures shall be aligned with the rights protecting individuals from adverse human rights impacts articulated in Article V of the Convention.

8. Each Party shall adopt and maintain measures to ensure the accuracy, robustness and reliability of the artificial intelligence system in critical domains. This includes the monitoring of the performance of these systems and the tracking of incidents created by or linked to artificial intelligence systems in sensitive areas.

9. The Parties shall adopt and maintain measures to foster the involvement of stakeholders in critical decisions and design choices throughout the life cycle of artificial intelligence systems in domains critical to the enjoyment of human rights and fundamental freedoms.

10. The Parties shall adopt and maintain measures to foster information-sharing between different organizations involved in the life cycle of the artificial intelligence system with the purpose of preventing foreseeable accidents and adverse human rights impacts in domains critical to the enjoyment of human rights and fundamental freedoms.

Article XII: Principles for Artificial Intelligence Systems Used in an International Context

1. Each Party shall adopt or maintain measures to foster the recognition of specific impacts of artificial intelligence systems in a global context by organizations involved in the life cycle of artificial intelligence systems. Each Party shall consider the implications of political, cultural and developmental contexts when developing, deploying or using artificial intelligence systems in an international context.
2. Each Party shall consider specific measures to prevent artificial intelligence systems from undermining the planet's linguistic and cultural diversity. Furthermore, the participatory, fair and consented representation of indigenous populations, ethnic, racial, gender, religious or linguistic minorities and cultures needs to be considered in the implementation of AI on an international level.
3. Each Party shall adopt or maintain measures to support efforts on the international level to reduce the global digital divide, promote digital literacy and contribute to the capacity building to inform individuals on the risks posed by artificial intelligence systems to human rights and their rights to remedy.

Chapter IV: Enforcement of the Convention

Article XIII: Enforcement Mechanism

The Parties shall implement effective enforcement mechanisms to uphold the convention's provisions and protect individuals from adverse human rights impacts facilitated by artificial intelligence systems. These enforcement mechanisms should aim to:

1. Provide a definition of existential risks created by or linked to artificial intelligence systems,
2. Design appropriate safeguards at an international level to prevent existential risks reinforced by AI systems,
3. Specify domains critical to the enjoyment of human rights and fundamental freedoms,
4. Track and investigate adverse human rights impacts created by or linked to artificial intelligence systems,
5. Establish mechanisms to report adverse human rights impacts to the Parties,
6. Establish best practices for human rights protection in the context of AI systems, with a particular focus on accountability and organizational governance as well as mechanisms for redress and restitution,
7. Develop mechanisms to address cybersecurity threats and criminal activities associated with AI systems,
8. Engage in a dialogue with stakeholders, including governments, international organizations, businesses, academia, civil society, standards organizations, and representatives of vulnerable groups,
9. Raise the awareness of the public on algorithmic harms and means to redress and restitution,
10. Provide a platform for education and capacity building for individuals to understand different types of harm posed by artificial intelligence, the meaning of human rights in the context of artificial intelligence systems, and instruments to redress and challenge adverse impacts created by or linked to artificial intelligence systems.
11. Propose amendments to the Convention in order to respond to emerging risks and technological change.
12. Establish the position of a special rapporteur on artificial intelligence, data and human rights with the authority, and competence required to investigate human rights violations in the context of artificial intelligence, data and other digital technologies.