**Treaty on the Regulation of Artificial Intelligence Development and Deployment**

**Preamble**:

Acknowledging the rapid advancements in artificial intelligence (AI) and the profound impact it has on various aspects of human life, the participating nations, hereinafter referred to as the "Parties," recognize the need for a global treaty that addresses the ethical, legal, and societal challenges posed by AI development and deployment. Mindful of the potential risks associated with unchecked AI progress, the Parties aim to strike a delicate balance between fostering innovation and ensuring the responsible and accountable use of AI technologies.

Affirming the significance of international cooperation, the Parties recognize that effective regulation and governance of AI require collaborative efforts, shared knowledge, and the establishment of common norms. The Parties further acknowledge the importance of inclusivity, transparency, and public engagement in shaping the future of AI, to ensure that it upholds fundamental human rights, social justice, and the dignity of all individuals.

Building upon the principles of precaution, prudence, and foresight, the Parties seek to adopt a proactive approach to AI governance. They emphasize the importance of addressing risks and challenges before they become insurmountable, thus safeguarding the interests of humanity and future generations. Recognizing the transformative potential of AI, the Parties also acknowledge its capacity to contribute positively to global progress, economic growth, sustainable development, and the achievement of the United Nations Sustainable Development Goals.

Mindful of the need to protect fundamental human rights, including the right to privacy, freedom of expression, and non-discrimination, the Parties commit to ensuring that AI technologies are developed and deployed in a manner that respects human dignity, promotes fairness, and guards against bias and discrimination. The Parties acknowledge the importance of inclusive AI systems that reflect the diversity of society and ensure equitable access to benefits.

Recognizing the potential impact of AI on the labour market, the Parties acknowledge the need to address the social and economic consequences of AI-driven automation. They emphasize the importance of fostering education, reskilling, and the creation of new employment opportunities to mitigate potential disruptions and ensure a just and inclusive transition.

Acknowledging the transnational nature of AI development and deployment, the Parties affirm the need for a unified global approach. They underscore the importance of international collaboration, information sharing, and capacity-building to address cross-border challenges, promote best practices, and prevent the emergence of regulatory gaps or uneven standards.

Drawing inspiration from the principles enshrined in international law, including human rights law, humanitarian law, and the principles of responsible innovation, the Parties recognize that AI technologies must adhere to the principles of transparency, explain ability, accountability, and predictability. They emphasize the importance of establishing clear ethical guidelines and standards to guide AI development, ensuring that human values and societal well-being remain at the forefront.

Guided by these principles and aspirations, the Parties come together to draft a treaty that addresses the challenges posed by AI development and deployment, with the aim of fostering responsible and inclusive AI systems that serve the interests of humanity while respecting the principles of human dignity, equity, and justice.

**Article 1: Scope, Principles and Foundations of AI Governance**

1.1 Scope: This treaty applies to all activities related to the development, deployment, and use of artificial intelligence (AI) technologies, including machine learning, deep learning, natural language processing, computer vision, robotics, and autonomous systems. It encompasses both public and private sectors, research institutions, academia, and any other entities engaged in AI-related activities.

1.2 Definitions: 1.2.1 Artificial General Intelligence (AGI): AI systems capable of performing any intellectual task that a human being can do, with a level of proficiency equal to or surpassing humans. 1.2.2 Artificial Superintelligence (ASI): AGI systems surpassing human capabilities across almost all domains, exhibiting intellectual abilities far superior to humans. 1.2.3 Large GPU and TPU Clusters: Computing infrastructures composed of substantial numbers of Graphics Processing Units (GPUs) and Tensor Processing Units (TPUs) used for AI research and development purposes, characterized by their computational power and ability to process large amounts of data. 1.2.4 Parameters: The adjustable weights or variables in a machine learning model that impact its complexity and computational requirements.

1.3 Interpretation: The provisions of this treaty shall be interpreted in line with its objectives, with the aim of adapting to technological advancements while upholding the treaty's fundamental principles.

1.4 Ethical Considerations: AI development, deployment, and use should adhere to ethical standards, including fairness, transparency, accountability, privacy, non-discrimination, human dignity, social benefit, and environmental sustainability.

1.5 International Collaboration: International cooperation, information sharing, and harmonization of regulatory frameworks are crucial in addressing global AI challenges, promoting responsible AI development, and preventing regulatory gaps.

1.6 Inclusivity and Public Engagement: The involvement of diverse stakeholders and public engagement is essential in decision-making processes related to AI development, deployment, and regulation. This ensures that AI systems reflect societal values and promote inclusivity.

1.7 Fundamental Human Rights: AI technologies must respect and promote fundamental human rights, such as privacy, freedom of expression, and non-discrimination, as enshrined in international human rights law.

1.8 Sustainable Development: AI should contribute to sustainable development, economic growth, and the achievement of the United Nations Sustainable Development Goals. It should be environmentally friendly, energy-efficient, and aligned with the principles of sustainability.

1.9 Precautionary Principle: Caution and prudence should guide AI development, particularly in deploying AI systems with significant societal impact or irreversible consequences. Precautionary measures must be taken to prevent or mitigate risks, even in the absence of full scientific certainty.

1.10 Responsible Innovation: AI development and deployment should integrate ethical, legal, and social considerations throughout the entire lifecycle, including research, development, deployment, and end-of-life phases.

1.11 Cultural and Contextual Sensitivity: AI systems should respect and reflect cultural diversity and contextual specificities, taking into account social norms, values, and expectations of the communities they serve.

1.12 International Research Collaboration: International collaboration, open access to research findings, and knowledge exchange are crucial in fostering innovation, advancing AI technologies, and bridging the global AI divide.

1.13 Privacy and Data Protection: AI technologies should respect privacy rights, comply with data protection laws, and incorporate privacy-preserving techniques and robust security measures.

1.14 Transparency and Explain ability: AI systems should be transparent, accountable, and capable of providing clear explanations for their decisions and actions, fostering trust, accountability, and user empowerment.

1.15 Technological Neutrality: AI governance should avoid prescribing specific technologies, allowing for flexibility and adaptation to evolving technological advancements.

1.16 Capacity-Building and Education: Efforts should be made to enhance capacity, digital literacy, and ethical understanding through educational programs, training, and knowledge-sharing initiatives.

1.17 International Cooperation for Safety and Security: International cooperation is necessary to prevent and mitigate the risks of malicious uses, cybersecurity threats, and unintended consequences associated with AI technologies.

1.18 Freedom of Research and Innovation: AI governance should strike a balance between responsible oversight and preserving open scientific inquiry, fostering an environment conducive to innovation, experimentation, and scientific progress.

1.19 Universal Access and Benefit Sharing: Efforts should be made to ensure universal access to AI benefits, address the digital divide, and promote fair and equitable benefit sharing, particularly when shared resources are utilized.

1.20 Interoperability and Compatibility: AI systems should be developed with interoperability and compatibility in mind, promoting open standards, data exchange, and collaboration among different AI systems and stakeholders.

1.21 Respect for National Laws and International Obligations: AI governance frameworks should respect national laws, regulations, and international obligations, including human rights law and other relevant international agreements.

1.22 Continuous Evaluation and Adaptation: The effectiveness and relevance of this treaty should be continuously evaluated, and the regulatory framework adapted to address emerging ethical, legal, and social concerns.

1.23 Principle of Non-Discrimination: AI technologies should be developed and deployed without discrimination, ensuring fairness, equity, and inclusivity for all individuals.

1.24 Future-Oriented Approach: AI governance should adopt a forward-looking approach, anticipating future challenges and opportunities through research, foresight studies, and scenario planning.

1.25 Principle of Proportional Regulation: Regulation of AI should be proportional to risks and societal impacts, striking a balance between regulation and innovation while safeguarding public interest, human rights, and societal well-being.

**Article 2: Temporary Moratorium on AI Development**

2.1 Pause on Large GPU and TPU Clusters: Recognizing the need for caution and prudence in AI development, a temporary shutdown is imposed on large GPU and TPU clusters, which are instrumental in refining the most powerful AIs. This pause allows for a thorough assessment of the potential risks, ethical considerations, and regulatory frameworks necessary to ensure responsible AI development. By temporarily suspending the operation of these clusters, the Parties aim to create an opportunity for international cooperation and the establishment of appropriate governance and oversight mechanisms.

2.2 Limit on Model Parameters: In response to the exponential growth in AI model size and complexity, a clear limit is set on the training of machine learning models or combinations thereof. Models with more than 500 million parameters are prohibited, aiming to prevent the uncontrolled proliferation of AI capabilities. This parameter limit promotes manageable and accountable AI research, ensuring that AI systems are comprehensible and transparent to researchers, policymakers, and the public. By imposing this restriction, the Parties seek to uphold principles of transparency and to facilitate the interpretation and auditing of AI systems.

2.3 Prohibition on Quantum Computers: Recognizing the potential risks and uncertainties associated with the use of quantum computers in AI-related activities, a prohibition is placed on their utilization. Until comprehensive regulations are established to ensure the safety, ethics, and legality of quantum-powered AI systems, the Parties commit to refraining from employing quantum computers in AI development and deployment. This prohibition serves to safeguard against unforeseen risks and threats that may arise from the intersection of quantum computing and AI, ensuring that the development and use of quantum-powered AI technologies adhere to stringent safety standards.

2.4 General Moratorium on Large-Scale AI Capabilities Research and Development: In order to strike a balance between scientific progress and responsible AI development, a general moratorium is implemented on large-scale research and development of AI capabilities. This moratorium remains in effect until robust safety protocols, ethical guidelines, and legal frameworks are established. During this period, the Parties prioritize research on AI safety, explain ability, interpretability, and ethical considerations. By placing a temporary hold on large-scale AI capabilities research and development, the Parties demonstrate their commitment to ensuring that AI technologies are developed responsibly, with due regard for potential risks and societal implications. This moratorium also allows for the exploration of mechanisms to prevent the concentration of power, ensure fairness, and address any potential biases or discriminatory impacts of AI systems.

2.5 Duration and Review: The temporary measures outlined in this article shall remain in effect until it is universally agreed, through consensus among the Parties, that it is safe, ethical, and in the best interests of humanity to resume large-scale AI capabilities research and development. The Parties commit to conducting regular assessments and reviews to evaluate progress in addressing risks, establishing responsible AI governance frameworks, and determining when it is appropriate to lift the temporary moratoriums. These reviews ensure that the treaty remains adaptive and responsive to the evolving landscape of AI technologies and the associated challenges.

**Article 3: Governance and Oversight**

3.1 Establishment of an International AI Governance Body: To ensure effective regulation and coordination of AI development and deployment, an international body, hereinafter referred to as the "AI Governance Body," shall be established. The AI Governance Body shall consist of representatives from participating nations and relevant stakeholders, including experts in AI, policymakers, industry representatives, and civil society organizations. The AI Governance Body shall serve as a central hub for discussions, collaboration, and decision-making on matters pertaining to AI governance and oversight.

3.2 Mandate and Responsibilities of the AI Governance Body: 3.2.1 The AI Governance Body shall be responsible for formulating and recommending policies, guidelines, and standards related to AI development, deployment, and use. It shall promote the adoption of ethical principles, best practices, and regulatory frameworks that align with the objectives of this treaty. 3.2.2 The AI Governance Body shall facilitate international cooperation, information sharing, and capacity-building initiatives to support the development of responsible and accountable AI systems. It shall encourage the exchange of knowledge, experiences, and expertise among participating nations and relevant stakeholders. 3.2.3 The AI Governance Body shall monitor and assess the global AI landscape, including technological advancements, emerging risks, and societal implications. It shall conduct regular evaluations to identify gaps, challenges, and areas requiring further attention, proposing necessary measures to address them. 3.2.4 The AI Governance Body shall foster transparency and public engagement by providing mechanisms for public input, consultation, and feedback in decision-making processes related to AI governance. It shall promote inclusivity, ensuring that the voices of diverse stakeholders, including marginalized communities and affected individuals, are heard and considered. 3.2.5 The AI Governance Body shall collaborate with existing international organizations, networks, and initiatives related to AI governance, fostering synergy and avoiding duplication of efforts. It shall engage in dialogue with relevant stakeholders, including academia, industry, and civil society, to leverage their expertise and insights. 3.2.6 The AI Governance Body shall establish mechanisms for monitoring compliance with the provisions of this treaty, ensuring adherence to ethical principles, legal frameworks, and responsible AI practices. It shall recommend enforcement mechanisms, including sanctions or penalties, for non-compliance with treaty obligations.

3.3 Decision-Making and Governance Processes: 3.3.1 The AI Governance Body shall adopt transparent and inclusive decision-making processes, guided by principles of accountability, fairness, and consensus-building. It shall ensure that decisions are based on scientific evidence, ethical considerations, and the best interests of humanity. 3.3.2 Participating nations shall have equal representation in the AI Governance Body, with decisions made through a consensus-based approach. The AI Governance Body shall strive to accommodate the specific needs, priorities, and circumstances of each participating nation, while upholding the overarching objectives and principles of this treaty. 3.3.3 The AI Governance Body shall establish expert working groups, task forces, or committees to address specific AI-related issues, such as data governance, algorithmic transparency, and AI ethics. These groups shall provide recommendations and expertise to inform the decision-making process of the AI Governance Body. 3.3.4 The AI Governance Body shall convene regular meetings, conferences, or assemblies to facilitate dialogue, knowledge sharing, and the review of progress in AI governance. It shall publish reports, guidelines, and recommendations to disseminate best practices and promote transparency in its activities.

3.4 Cooperation with National Authorities: 3.4.1 Participating nations shall cooperate with the AI Governance Body by providing relevant information, data, and expertise to support its work. They shall align their national policies, regulations, and standards with the recommendations and guidelines issued by the AI Governance Body. 3.4.2 Participating nations shall establish national AI governance authorities or bodies to ensure effective implementation of the treaty provisions at the national level. These national authorities shall collaborate with the AI Governance Body, sharing experiences, best practices, and lessons learned to strengthen global AI governance.

3.5 Funding and Resources: 3.5.1 Participating nations shall commit to providing adequate funding and resources to support the functioning of the AI Governance Body and its associated activities. They shall allocate resources for research, capacity-building, and technical assistance programs aimed at promoting responsible AI development and governance. 3.5.2 The AI Governance Body may seek voluntary contributions from public and private entities, including international organizations, philanthropic foundations, and industry stakeholders, to enhance its financial sustainability and operational capabilities.

3.6 Review and Evaluation: 3.6.1 The effectiveness, relevance, and performance of the AI Governance Body shall be subject to periodic review and evaluation. Participating nations shall assess the body's achievements, impact, and alignment with the objectives of this treaty, making necessary adjustments and improvements to enhance its effectiveness. 3.6.2 Participating nations shall engage in ongoing dialogue and collaboration to ensure that the AI Governance Body remains adaptive and responsive to the evolving landscape of AI technologies, societal needs, and emerging challenges.

Top of Form

**Article 4: Ethical Guidelines and Human Rights**

4.1 Ethical Guidelines: Clear and comprehensive ethical guidelines shall be developed and promoted to guide the design, implementation, and use of AI technologies. These guidelines shall prioritize principles such as fairness, transparency, accountability, privacy protection, non-discrimination, human dignity, and social benefit.

4.2 Human-Centered AI Design: AI technologies shall be designed to augment human capabilities, enhance well-being, and promote human flourishing. Human-centered design approaches shall be adopted, ensuring user empowerment, inclusivity, and avoiding undue concentration of power.

4.3 Respect for Human Rights: Fundamental human rights, including privacy, freedom of expression, and non-discrimination, shall be protected and upheld in AI development and deployment. AI systems shall be designed to prevent violations of human rights and safeguard against discrimination.

4.4 Privacy Protection: AI technologies shall comply with data protection laws and regulations. Privacy-preserving techniques, secure data handling practices, and robust security measures shall be implemented to protect personal information and privacy rights.

4.5 Transparency and Explain ability: AI systems shall be transparent, accountable, and capable of providing clear explanations for their decisions and actions. Efforts shall be made to enhance understanding of AI system functioning, limitations, and potential biases through research and standards.

4.6 Non-Discrimination and Fairness: AI technologies shall be developed and deployed in a manner that prevents discrimination and promotes fairness. Measures shall be taken to prevent the creation or perpetuation of biases in AI systems and ensure fairness, equity, and inclusivity.

4.7 Human Oversight and Accountability: Mechanisms shall be established to ensure human oversight and accountability in AI development and deployment. Human intervention and control shall be maintained in critical decision-making processes to prevent undue reliance on AI systems.

4.8 International Cooperation on Ethical Guidelines: Collaboration shall be fostered to develop international ethical guidelines for AI technologies. Best practices and experiences shall be shared to promote convergence and harmonization of ethical standards globally.

4.9 Regular Ethical Impact Assessments: Regular ethical impact assessments of AI systems shall be conducted to evaluate risks, benefits, and societal implications. These assessments shall inform the design and deployment of AI technologies and facilitate proactive risk mitigation.

4.10 Education and Awareness: Initiatives shall be undertaken to enhance understanding of ethical considerations in AI. Education programs and awareness campaigns shall be promoted to foster digital literacy, ethical awareness, and responsible use of AI technologies.

**Article 5: Enforcement and Compliance**

5.1 Compliance with Treaty Obligations: 5.1.1 The Parties shall undertake all necessary measures to ensure compliance with the provisions of this treaty within their respective jurisdictions. They shall align their national laws, regulations, and policies with the objectives and principles of this treaty. 5.1.2 Participating nations shall establish effective mechanisms and enforcement measures to monitor and enforce compliance with treaty obligations. These mechanisms may include regulatory bodies, inspection procedures, reporting requirements, and penalties for non-compliance.

5.2 International Cooperation on Enforcement: 5.2.1 The Parties shall engage in international cooperation to promote effective enforcement of this treaty. They shall share information, best practices, and experiences on enforcement mechanisms, facilitate mutual assistance, and provide technical support to enhance enforcement capabilities. 5.2.2 Participating nations shall cooperate in investigations, exchange information, and collaborate in cross-border enforcement actions when AI-related activities violate treaty obligations or national laws.

5.3 Reporting and Monitoring: 5.3.1 Participating nations shall submit regular reports on their implementation of this treaty, including the measures taken to enforce its provisions and address emerging challenges. These reports shall facilitate transparency, accountability, and mutual learning among the Parties. 5.3.2 The AI Governance Body, as established in Article 3, shall be responsible for monitoring compliance with treaty obligations. It shall review the reports submitted by participating nations, conduct assessments, and provide recommendations for improvements in enforcement and compliance mechanisms.

5.4 Dispute Settlement: 5.4.1 The Parties shall strive to resolve any disputes arising from the interpretation or application of this treaty through peaceful means and in accordance with international law. They shall engage in dialogue, negotiation, and mediation to reach amicable solutions. 5.4.2 In the event that disputes cannot be resolved through negotiation or mediation, the Parties may resort to established dispute settlement procedures, including arbitration or other peaceful means as agreed upon by the disputing Parties.

5.5 Sanctions and Remedies: 5.5.1 Participating nations shall establish sanctions and remedies for non-compliance with treaty obligations. These may include administrative penalties, fines, injunctions, or other appropriate measures to ensure adherence to ethical guidelines, responsible AI practices, and regulatory frameworks. 5.5.2 Sanctions and remedies shall be proportionate and effective in deterring non-compliance, promoting responsible AI development, and addressing any harm or violations resulting from non-compliant AI activities.

5.6 Capacity Building and Technical Assistance: 5.6.1 The Parties shall provide capacity-building support and technical assistance to facilitate compliance with treaty obligations, particularly for developing nations or nations with limited resources. Capacity building initiatives may include training programs, knowledge-sharing platforms, and access to expertise and resources. 5.6.2 The AI Governance Body shall coordinate and facilitate capacity-building efforts, collaborating with relevant stakeholders to ensure the effective transfer of knowledge, skills, and technologies to support compliance with the treaty.

5.7 Review and Evaluation: 5.7.1 The enforcement and compliance mechanisms established under this article shall be subject to regular review and evaluation. The Parties shall assess their effectiveness, efficiency, and fairness to make necessary improvements and adaptations over time. 5.7.2 The AI Governance Body shall play a central role in reviewing and evaluating the enforcement and compliance mechanisms, making recommendations for enhancements, and facilitating information sharing and mutual learning among participating nations.

**Article 6: Duration and Review**

6.1 Duration of the Treaty: This treaty shall remain in force indefinitely until it is universally agreed, through consensus among the Parties, that it is safe, ethical, and in the best interests of humanity to resume large-scale AI capabilities research and development.

6.2 Regular Reviews: The Parties shall conduct regular reviews of the treaty's effectiveness, relevance, and performance. These reviews shall evaluate progress in addressing risks, achieving the treaty's goals, and adapting to emerging ethical, legal, and social concerns related to AI technologies.

6.3 Adaptation and Amendments: Based on the reviews and evolving AI landscape, the Parties may propose amendments or modifications to the treaty. Any amendments or modifications shall be made through consensus among the Parties and shall be aimed at improving the treaty's efficacy, responsiveness, and alignment with technological advancements.

6.4 Emerging Technologies: The Parties shall monitor and assess the development and impact of emerging AI technologies and other related fields. They shall consider the need for specific regulations, guidelines, or provisions to address the unique challenges and opportunities presented by such technologies.

6.5 International Collaboration on Research and Development: During the treaty's duration, the Parties shall foster international collaboration on research and development of AI technologies. This collaboration shall focus on sharing knowledge, expertise, and resources to promote responsible AI innovation, while respecting the temporary moratoriums and ethical guidelines established under this treaty.

6.6 Universal Agreement for Resumption: The resumption of large-scale AI capabilities research and development shall require universal agreement among the Parties. This agreement shall be based on scientific evidence, consensus on ethical principles, and confidence in the establishment of robust governance and oversight mechanisms for safe and responsible AI development.

6.7 Withdrawal from the Treaty: A Party may withdraw from this treaty by providing written notice to the AI Governance Body. The withdrawal shall take effect one year after the receipt of the notice, and the withdrawing Party shall continue to fulfill its obligations until the withdrawal becomes effective.

**Conclusion:**

In conclusion, this treaty provides a comprehensive framework for addressing the challenges associated with AI development, deployment, and use. It emphasizes the importance of ethical considerations, human rights, governance, and enforcement mechanisms to ensure responsible and accountable AI technologies. The treaty establishes temporary moratoriums on specific AI activities, sets parameter limits, and prohibits the use of certain technologies until comprehensive regulations and oversight structures are in place. It also establishes an international body to oversee the treaty, promote collaboration, and monitor compliance.

The treaty's effectiveness will rely on the Parties' commitment to aligning national laws, fostering international cooperation, and sharing knowledge and resources. Regular reviews and evaluations will enable the adaptation of the treaty to emerging ethical, legal, and social concerns. Additionally, the treaty emphasizes the importance of capacity building, education, and awareness initiatives to promote responsible AI development and foster a global understanding of ethical considerations.

By implementing this treaty, the Parties aim to strike a balance between scientific progress and the responsible development of AI technologies. They recognize the need to ensure transparency, fairness, privacy protection, and the prevention of discrimination in AI systems. Furthermore, the treaty emphasizes the respect for fundamental human rights and the principle of human-centered design to ensure AI technologies benefit humanity.

Overall, this treaty provides a clear and comprehensive roadmap for the responsible development and use of AI technologies. It promotes collaboration, transparency, and accountability, with the ultimate goal of ensuring that AI advancements align with ethical principles, respect human rights, and contribute to the well-being of society as a whole.