**Treaty for Moratorium and Responsible Regulation of Artificial Intelligence Development**

The States concluding this Treaty, hereinafter referred to as the “Parties” to the Treaty,

*Bearing in mind* the rapid advancements and potential of AI technologies, and acknowledging the need to ensure their responsible development and use for the benefit of humanity and the global community,

*Recognising* that recent advancements in AI have been progressing at an unprecedented pace, presenting significant challenges in terms of our technological preparedness to comprehensively assess, understand, and regulate its development,

*Acknowledging* the complex and multifaceted nature of AI, which encompasses intricate algorithms, advanced learning capabilities, and potentially far-reaching implications for society, including ethical, economic, and existential considerations,

*Recognising* that the development of powerful AI systems should only proceed when there is a high level of confidence that their effects will be positive and their risks will be effectively manageable,

*Keeping in mind* that various draft pieces of legislation, such as the proposal put forth by the European Union, adopt a risk-based approach in grading AI capabilities and regulating their use,

*Understanding* that the development space of AI encompasses a substantial workforce distributed across the globe,

*Affirming* the significance of promoting the principle of knowledge exchange and collaboration among all Parties to the Treaty, to advance the development of robust regulatory mechanisms for the responsible use AI and harness its potential for maximum societal benifit,

*Desiring* to address the risks and uncertainties associated with the rapid development of AI by implementing a moratorium on its further development until adequate regulatory checks and safeguards have been established, have agreed as follows:

**Article I**

Each of the Parties to this treaty commits to implementing a total moratorium on the development of AI according to other articles of this treaty.

**Article II**

1. Each of the Parties to this treaty agrees to enforce the immediate shutdown of all GPU clusters, in all kinds of establishments without exceptions. The parties will also launch an immediate effort to locate and monitor all GPUs at the domestic level. The goal of this extensive identification effort is to guarantee that the moratorium is effectively enforced.
2. Each of the parties to this treaty agrees grees to implement a prohibition on training of machine learning models. The prohibition on training machine learning models encompasses the development and refinement of models that have the capacity to learn from data and make predictions or decisions based on that learned knowledge.

**Article III**

Each of the parties to this treaty shall commit to enact stringent domestic laws making the development of AI criminal while the moratorium is in effect. The moratorium will be enforced by these rules, which will also serve to deter any unlicensed or unlawful AI development operations. The parties want to assure compliance and stop any possible hazards or unauthorised breakthroughs in AI technology by establishing strong legal frameworks.

**Article IV**

Each of the Parties to this treaty shall recognise the critical importance of transparency in the foundational models of AI to enable thorough investigation and understanding of vulnerabilities and biases.

Each of the Parties will aim to mandate that the developers of AI models shall acknowledge the necessity of collaboration with stakeholders to address potential harms and ensure explainability and interpretability. Furthermore, standardised measurement and benchmarking methodologies, such as HELM, shall be developed to assess the performance, fairness, safety, and robustness of already existing AI models.

Independent audits and assessments by researchers shall be conducted to evaluate the risks associated with deploying AI models, particularly large language models. Regulators shall collaborate with researchers, industry, and auditors to develop guidelines for evaluating high-risk applications and establishing prohibitions if necessary.

**Article V**

Each party to this treaty shall recognise the importance of addressing the employment implications arising from the moratorium on AI development. Each party shall allocate resources for training programs aimed at enhancing skills and knowledge in AI governance, creating research and development projects focused on designing effective frameworks and providing consultation and advisory roles to utilize the expertise of those impacted by the moratorium. By actively engaging individuals affected by the moratorium and offering employment opportunities in the AI governance sector, parties aim to ensure the continued involvement and valuable contributions of these individuals in shaping responsible and comprehensive AI governance frameworks. Regular evaluation and adjustments to these initiatives will be undertaken to optimise support and facilitate a smooth transition for those affected.

**Article VI**

1. Each of the parties will be a member state of the International Organization for Regulation of AI Development (“IORAID”) which is hereby established as the competent organisation responsible for ensuring compliance with the provisions of this treaty. IORAID shall perform the necessary duties and functions to monitor and regulate AI development activities in accordance with the principles outlined herein.
2. IORAID shall serve as the Secretariat for this treaty, providing administrative support and facilitating communication among the parties. It will be in charge of organising meetings, keeping notes, and distributing pertinent data about the application and enforcement of this treaty.
3. Each of the parties to this treaty shall contribute an appropriate share towards the expenses incurred by IORAID in carrying out its duties on the basis of the resources and capabilities of the parties. IORAID shall have the authority to conduct inspections, request reports, and undertake other necessary actions to ensure compliance with the provisions of this treaty.

**Article VII**

Each of the parties to the treaty shall convene meetings at least once every 120 days, or sooner if prescribed by the IORAID to review the status of technological advancements necessary for monitoring the development of AI. By ensuring regular communication and collaboration, the parties aim to stay abreast of technological developments and strengthen their collective efforts in effectively regulating and monitoring AI development during the moratorium period.

**Article VIII**

1. A party may suggest the lifting of the moratorium through the multilateral consultation process if they are confident in their ability to provide thorough safeguards, legal frameworks, and governance systems that adequately address the dangers related to AI development.
2. The party requesting the lifting of the moratorium must offer compelling evidence of the viability, strength, and enforceability of the safeguards they have put in place. A thorough global review process will be used to assess the request to lift the moratorium.
3. A two-thirds majority of the treaty's parties is required to remove the moratorium on AI development.

**Article** **IX**

Nothing in this Treaty shall impede the right of any group of States to enter into regional agreements aimed at ensuring a complete prohibition on the use and development of AI.

IN WITNESS WHEREOF the undersigned, duly authorized, have signed this Treaty DONE in triplicate, at the cities of Tokyo, Canberra and New Delhi, the fifteenth day of July, two thousand and twenty three.