**Treaty on the Moratorium of Large-Scale AI Capabilities**

**Research and Development**

**CONTENTS**

1. Preamble…………………………………………………………………………………………3
2. Article I: Definition……………………………………………………………………………....3
3. Article II: Moratorium on Large-Scale AI Capabilities Research and Development……….……4
4. Article III: Criminalization of Artificial General Intelligence and Artificial Superintelligence

Development……………………………………………………….………………...4

1. Article IV: Establishment of an International Oversight Body…………………………………...4
2. Article V: Enforcement and Compliance Mechanisms…………………………………………...5
3. Article VI: Duration and Review…………………………………………………………………5
4. Article VII: Ethical Framework for AI Development…………………………………………….5
5. Article VIII: Assessment and Mitigation of AI Risks…………………………………………….6
6. Article IX: Funding and Investment Guidelines………………………………………………….6
7. Article X: Technical Cooperation and Knowledge Transfer……………………………………...6
8. Article XI: Public Engagement and Awareness…………………………………………………...7
9. Article XII: Exit Strategy and Safeguards…………………………………………………………7
10. Article XIII: Depositary and Notifications………………………………………………………...7
11. Article XIV: Relation to Other Agreements…………………….………………………………...8
12. Article XV: Amendments…………………………………………….…………………………...8
13. Article XVI: Settlement of Disputes……………………………………………………….……...8
14. Article XVII: Signature, Ratification, Acceptance, Accession……………………….…………...9
15. Article XVIII: Entry into Force…………………………………………………………………...9
16. Article XIX: Duration and Withdrawal……………………………………….………….……….9
17. Article XX: Authentic Texts………………………………………………….……...……….…...9
18. Annex A: Technical Specifications………………………………………………………………10
19. Annex B: International Oversight Body…………………………………………………………10
20. Annex C: Review Mechanisms and Evaluation Criteria…………………………………………11
21. Final Provisions………………………………………………………………………………….12

**Preamble**

The participating States, recognizing the potential benefits and risks associated with the rapid advancement of Artificial Intelligence (AI) technologies and the imperative need for responsible development, hereby establish this Treaty on the Moratorium of Large-Scale AI Capabilities Research and Development. The aim of this treaty is to temporarily halt the further progress of AI capabilities on a large scale, while enabling the establishment of international governance and legal frameworks to ensure safe, ethical, and accountable development.

Considering the necessity to address potential risks associated with large-scale AI capabilities development, including but not limited to the refining of powerful AI systems in extensive GPU and TPU clusters,

Recognizing the potential risks of training machine learning models or combinations of models exceeding 500 million parameters,

Mindful of the risks associated with the use of quantum computers in AI-related activities,

Emphasizing the importance of criminalizing the development of Artificial General Intelligence (AGI) or Artificial Superintelligence (ASI) to ensure responsible and controlled progress,

Affirming the need for international cooperation and coordination to effectively manage the development and deployment of AI technologies,

Reaffirming the commitment to ensuring the safety, ethical use, and protection of human rights in the realm of AI,

Desiring to establish an international body to oversee the implementation, enforcement, and ongoing assessment of this treaty,

Hereby agree to the following provisions:

Article I:

*Definition*

For the purposes of this treaty:

1. "AI" refers to systems, algorithms, and technologies that demonstrate intelligent behaviour by analysing their environment and making decisions or taking actions to achieve specific goals by imitating human intelligence.

2. "Large-scale AI capabilities research and development" refers to activities aimed at significantly improving the capabilities, performance, or complexity of AI systems, including but not limited to the refinement of AI models, recreating human intelligence, the utilization of extensive computing resources, and the exploration of advanced technologies.

3. "GPU" stands for Graphics Processing Unit, which is a specialized processor originally designed to accelerate graphics rendering. GPUs can process many pieces of data simultaneously, making them useful for machine learning, video editing, and gaming applications.

4. "TPU" stands for Tensor Processing Unit, which refers to a type of Application-Specific Integrated Circuit (ASIC) developed by Google particularly for accelerating machine learning workloads and lead-time.

5. "Machine Learning (ML) models" refers to algorithms and mathematical models that enable AI systems to learn from data, improve performance, and make predictions or decisions.

6. "Parameters" refers to the adjustable variables or weights within a machine learning model that are optimized during training to improve its performance and quality.

7. "Quantum computers" refers to computational systems that leverage principles from quantum mechanics to perform calculations and offer potential advantages in processing power and problem-solving capabilities.

Article II:

*Moratorium on Large-Scale AI Capabilities Research and Development*

1. All participating States shall impose an immediate moratorium on large-scale AI capabilities research and development, as defined in Article 1.

2. The following are just some of the things that will be included in the moratorium:

a. The shutdown of large GPU and TPU clusters used in the development of sophisticated AI systems.

b. The preclusion of preparing ML models or mixes of models surpassing 500 million boundaries.

c. The preclusion of the utilization of quantum PCs in any computer-based intelligence related exercises.

3. Within their borders, Participating States are required to implement the necessary measures to enforce and monitor compliance with the moratorium.

Exception- Notwithstanding the limitations outlined in Article II, an exception shall be granted to machine learning models that are utilized in sectors critical to preserving human life and safety. Such sectors may include, but are not limited to, healthcare, emergency services, and public safety. The purpose of this exception is to ensure that essential services and activities, which directly contribute to the well-being and protection of individuals, are not unduly hindered by the moratorium.

Article III:

*Criminalization of Artificial General Intelligence and Artificial Superintelligence Development*

1. Participating States shall pass national laws criminalizing the development, creation, or intentional release of Artificial General Intelligence (AGI) or Artificial Superintelligence (ASI).

2. The laws will characterize the extension, punishments, and authorization instruments important to actually stop and address infringement.

3. Participating States shall cooperate in exchanging information and best practices regarding the implementation and enforcement of national laws criminalizing AGI and ASI development.

Article IV:

*Establishment of an International Oversight Body*

1. An International Oversight Body shall be established to oversee the implementation, enforcement, and ongoing assessment of this treaty.

2. The International Oversight Body shall consist of representatives from participating States, selected based on their expertise and commitment to the safe, ethical, and accountable development of AI technologies.

3. The International Oversight Body shall have the authority to:

a. Monitor and evaluate the compliance of participating States with the provisions of this treaty.

b. Facilitate international cooperation and information sharing regarding AI development and deployment.

c. Recommend amendments or revisions to the treaty, as necessary, to adapt to evolving technological, ethical, and security considerations.

Article V:

*Enforcement and Compliance Mechanisms*

1. Participating States shall establish effective mechanisms to enforce and ensure compliance with the provisions of this treaty.

2. The mechanisms shall include appropriate monitoring, reporting, and investigation procedures, as well as penalties and sanctions for non-compliance.

3. Participating States shall cooperate in sharing best practices and expertise to enhance enforcement and compliance mechanisms.

Article VI:

*Duration and Review*

1. This treaty shall remain in force until it is universally agreed among participating States that it is safe, ethical, and appropriate to resume large-scale AI capabilities research and development.

2. Participating States shall periodically review the effectiveness and relevance of this treaty, considering technological advancements, international developments, and the evolving understanding of AI-related risks and opportunities.

Article VII:

*Ethical Framework for AI Development*

1. Participating States shall promote the development and adoption of comprehensive ethical frameworks for AI that align with the principles of transparency, accountability, fairness, and respect for human rights.

2. The ethical frameworks shall address, but not be limited to, the following areas:

a. Data privacy and protection: Ensuring the responsible collection, storage, and use of personal and sensitive data in AI systems.

b. Bias and discrimination: Mitigating and eliminating biases and discriminatory practices in AI algorithms and decision-making processes.

c. Explain ability and interpretability: Promoting the development of AI systems that are explainable and understandable to users and stakeholders.

d. Human oversight and control: Ensuring that AI systems are designed to operate under human supervision and maintain human decision-making authority.

e. Accountability and liability: Establishing clear lines of responsibility and accountability for AI system creators, operators, and users.

f. Socio-economic impact: Assessing and addressing the potential impact of AI on employment, economic disparities, and social dynamics.

g. International cooperation: Encouraging collaboration among participating States to share best practices and address common ethical challenges.

Article VIII:

*Assessment and Mitigation of AI Risks*

1. Participating States shall conduct ongoing assessments of AI risks, with a particular focus on large-scale AI capabilities research and development.

2. The assessments shall include, but not be limited to, the following areas:

a. Safety and security: Evaluating the potential risks and vulnerabilities associated with AI systems, including the prevention of unintended consequences, adversarial attacks, and unauthorized access.

b. Human rights and societal impact: Assessing the impact of AI on fundamental rights, social cohesion, and democratic processes, and implementing necessary safeguards.

c. Dual-use technologies: Addressing concerns related to the military or harmful use of AI technologies and ensuring compliance with international arms control agreements.

d. Environmental impact: Considering the environmental footprint of AI technologies and promoting sustainable practices throughout the AI lifecycle.

Article IX:

*Funding and Investment Guidelines*

1. Participating States shall develop guidelines and criteria for the responsible funding and investment in AI research and development.

2. The guidelines shall promote the allocation of resources towards AI projects that align with the principles of this treaty, including the promotion of safe, ethical, and accountable AI technologies.

3. Participating States shall encourage private and public entities to adhere to these guidelines and incentivize responsible AI investment practices.

Article X:

*Technical Cooperation and Knowledge Transfer*

1. Participating States shall promote technical cooperation and knowledge transfer among researchers, developers, and stakeholders in the field of AI.

2. The technical cooperation and knowledge transfer shall include, but not be limited to, the following areas:

a. Exchange of best practices: Facilitating the sharing of expertise, methodologies, and lessons learned in AI research, development, and governance.

b. Capacity development: Supporting the training and skill development of individuals and organizations working in AI, particularly in the context of safe and ethical practices.

c. Technology transfer: Encouraging the responsible transfer of AI technologies, while considering potential risks and ensuring compliance with intellectual property rights.

d. Collaboration platforms: Establishing platforms or networks for collaboration, communication, and knowledge dissemination among AI practitioners and stakeholders.

Article XI:

*Public Engagement and Awareness*

1. Participating States shall promote public engagement and awareness initiatives to foster a broader understanding of AI technologies, their benefits, risks, and societal implications.

2. The public engagement and awareness initiatives shall include, but not be limited to, the following activities:

a. Public consultations: Encouraging participatory processes and soliciting public input on AI-related policies, regulations, and ethical frameworks.

b. Education and outreach: Promoting AI literacy and awareness through educational programs, campaigns, and public events.

c. Ethical guidelines for AI use: Developing and disseminating guidelines and resources to assist individuals, organizations, and governments in making responsible and ethical decisions regarding AI deployment.

d. Transparency and accountability: Encouraging AI system developers and operators to communicate openly about the capabilities, limitations, and potential biases of their systems.

Article XII:

*Exit Strategy and Safeguards*

1. Participating States shall establish an exit strategy and safeguards to guide the eventual lifting of the moratorium on large-scale AI capabilities research and development.

2. The exit strategy and safeguards shall include, but not be limited to, the following elements:

a. Risk assessment: Conducting thorough assessments of the progress made in developing international governance and legal frameworks for AI and ensuring that the necessary safeguards are in place to mitigate risks.

b. International consensus: Achieving a broad international consensus among participating States, industry leaders, and relevant stakeholders on the safety, ethics, and responsible development of large-scale AI capabilities.

c. Review mechanisms: Establishing regular review mechanisms to assess the progress of international governance frameworks and the evolving landscape of AI technologies.

d. Gradual resumption: Gradually lifting the moratorium in a phased manner, with clear guidelines, benchmarks, and monitoring mechanisms to ensure the safe and responsible resumption of large-scale AI capabilities research and development.

Article XIII:

*Depositary and Notifications*

1. The depositary of this treaty shall be the Secretary-General. (The depositary can be discussed by participating state during negotiation).

2. The depositary shall be responsible for receiving and circulating notifications, communications, and other relevant documents related to this treaty among the participating States.

3. Each participating State shall promptly notify the depositary of the following:

a. Signature of this treaty.

b. Ratification or acceptance of this treaty.

c. Amendments to this treaty.

d. Withdrawal from this treaty.

e. Any other relevant information or communication pertaining to this treaty.

4. The depositary shall maintain a record of signatures, ratifications, acceptances, amendments, withdrawals, and any other relevant information regarding this treaty.

Article XIV:

*Relation to Other Agreements*

1. This treaty shall be without prejudice to the rights and obligations of participating States under any other international agreements to which they are party.

2. In case of any conflict or inconsistency between the provisions of this treaty and any other international agreement, the provisions of this treaty shall prevail to the extent of such conflict or inconsistency.

Article XV:

*Amendments*

1. Any participating State may propose amendments to this treaty.

2. Proposed amendments shall be submitted to the depositary, who shall promptly circulate them to all participating States.

3. Amendments shall be adopted by consensus or through a defined voting procedure agreed upon by participating States.

4. Amendments adopted shall enter into force for the participating States that have accepted them upon the deposit of their instruments of acceptance with the depositary.

Article XVI:

*Settlement of Disputes*

1. Participating States shall seek to settle any dispute arising from the interpretation or application of this treaty through negotiation or other peaceful means.

2. If a dispute cannot be resolved through negotiation or other peaceful means, the parties to the dispute may agree to submit it to arbitration or other appropriate mechanisms for settlement.

3. In the absence of an agreement on arbitration, the dispute shall be submitted to the International Court of Justice, if all parties to the dispute are parties to the Statute of the International Court of Justice, or to another court or tribunal of competent jurisdiction, as mutually agreed upon by the parties.

Article XVII:

*Signature, Ratification, Acceptance, Accession*

1. This treaty shall be open for signature by all participating States at New York on 12 July 2023.

2. This treaty shall be subject to ratification, acceptance, or approval by participating States in accordance with their respective constitutional procedures.

3. This treaty shall remain open for accession by any state that is not a participating State upon approval by a majority of the participating States.

4. Instruments of ratification, acceptance, approval, or accession shall be deposited with the depositary.

Article XVIII:

*Entry into Force*

1. This treaty shall enter into force on the 30th day following the deposit of the 25 instruments of ratification, acceptance, approval, or accession with the depositary.

2. For any participating State that deposits its instrument of ratification, acceptance, approval, or accession after the entry into force of this treaty, it shall become a party to this treaty on the [number] day following the deposit of its instrument.

Article XIX:

*Duration and Withdrawal*

1. This treaty shall remain in force indefinitely, unless terminated in accordance with its provisions.

2. Any participating State may withdraw from this treaty by providing written notice to the depositary.

3. The withdrawal shall take effect in 90 days following the receipt of the notice by the depositary, unless a longer period is mutually agreed upon by the participating States.

Article XX:

*Authentic Texts*

1. This treaty, done in English, shall be equally authentic in each of the languages.

2. The original texts of this treaty shall be deposited with the depositary, who shall prepare certified copies in English.

Annex A:

*Technical Specifications*

This annex provides technical specifications and guidelines to support the implementation of the provisions outlined in the main body of the treaty. These specifications aim to ensure consistency and clarity in the application of the treaty's restrictions and requirements.

Section 1: Large GPU and TPU Clusters

1.1 Definition: Large GPU and TPU clusters, as mentioned in Article 2(2)(a) of this treaty, refer to computer systems comprised of a substantial number of Graphics Processing Units (GPUs) or Tensor Processing Units (TPUs) interconnected to perform computationally intensive tasks related to AI research and development.

1.2 Shutdown Process: Participating States shall establish procedures for the orderly shutdown of large GPU and TPU clusters within their jurisdictions, ensuring that the shutdown process adheres to safety protocols and mitigates any potential risks associated with the termination of AI-related activities.

Section 2: Training ML Models Parameters

2.1 Definition: The prohibition mentioned in Article 2(2)(b) of this treaty applies to the training of machine learning (ML) models or combinations of models exceeding 500 million parameters. Parameters refer to the adjustable variables or weights within ML models that are optimized during training.

2.2 Monitoring and Enforcement: Participating States shall establish mechanisms for monitoring and enforcement to ensure compliance with the parameter limitation. This may involve auditing and verifying the parameter count of trained ML models and implementing penalties for violations.

Section 3: Use of Quantum Computers

3.1 Prohibition: As stated in Article 2(2)(c) of this treaty, the use of quantum computers in any AI-related activities is prohibited. Quantum computers refer to computational systems that leverage principles from quantum mechanics to perform calculations and offer potential advantages in processing power and problem-solving capabilities.

3.2 Compliance Measures: Participating States shall take measures to prevent the use of quantum computers for AI-related tasks and ensure appropriate controls and monitoring are in place to detect and address any violations of this prohibition.

Annex B:

*International Oversight Body*

This annex provides guidelines and principles for the establishment and operation of the International Oversight Body mentioned in Article 4 of the treaty.

1. Composition: The International Oversight Body shall consist of representatives from participating States, selected based on their expertise and commitment to the safe, ethical, and accountable development of AI technologies. The composition should be balanced and reflect geographical diversity.

2. Functions and Powers: The International Oversight Body shall have the following functions and powers:

a. Monitor and evaluate the compliance of participating States with the provisions of the treaty.

b. Facilitate international cooperation and information sharing regarding AI development and deployment.

c. Recommend amendments or revisions to the treaty, as necessary, to adapt to evolving technological, ethical, and security considerations.

d. Provide guidance and expertise on technical and ethical matters related to AI development and governance.

3. Reporting and Communication: The International Oversight Body shall prepare regular reports on the implementation and compliance with the treaty, including recommendations for improvements. It shall communicate its findings and recommendations to participating States and relevant international organizations.

4. Decision-Making Process: The International Oversight Body shall adopt its own rules and procedures for decision-making, which shall ensure transparency, inclusivity, and the fair representation of participating States.

Annex C:

*Review Mechanisms and Evaluation Criteria*

This annex outlines the review mechanisms and evaluation criteria to assess the effectiveness, relevance, and safety of the treaty's provisions, as mentioned in Article VI.

1. Periodic Reviews: Participating States shall conduct periodic reviews of the treaty's implementation and effectiveness, taking into account technological advancements, international developments, and evolving understanding of AI-related risks and opportunities. The reviews shall be conducted at intervals of 1 years.

2. Evaluation Criteria: The reviews shall assess the following aspects:

a. Progress in developing international governance and legal frameworks for AI.

b. Advancements in AI technologies and their impact on society, economy, and security.

c. Effectiveness of the treaty's provisions in mitigating risks associated with large-scale AI capabilities research and development.

d. Compliance with ethical principles, human rights, and responsible AI practices.

e. Input from relevant stakeholders, including academia, industry, civil society, and affected communities.

3. Reporting and Recommendations: Participating States shall submit reports on the outcomes of the reviews to the International Oversight Body, including recommendations for potential revisions or adjustments to the treaty based on the evaluation.

4. International Cooperation: Participating States shall engage in international cooperation and knowledge sharing to enhance the effectiveness of the review mechanisms. This may include sharing best practices, conducting joint assessments, and exchanging expertise on AI development and governance.

*Final Provisions*

1. This treaty shall be open for signature by all United Nations Member States and any other states invited by the participating States.
2. This treaty shall enter into force upon ratification or acceptance by a minimum of 25 participating States.
3. This treaty, including its annexes, shall form an integral part thereof and shall be equally binding.
4. The treaty and its provisions shall be implemented in accordance with the respective domestic laws and regulations of participating States.
5. Amendments to this treaty shall be communicated to the depositary and shall enter into force in accordance with the procedures outlined in Article XVI.
6. This treaty may be terminated or modified through mutual agreement by participating States, subject to the provisions of international law.
7. The depositary shall register this treaty, including its annexes, with the Secretariat of the United Nations and other relevant international organizations.

IN WITNESS WHEREOF, the undersigned, being duly authorized by their respective governments, have signed this Treaty on the Moratorium of Large-Scale AI Capabilities Research and Development.

DONE at New York on 19 July 2023, in 10 original copies, each in the English, all texts being equally authentic.

[Signature block]

Notes-

1. Regarding Article II

During my research work, I have discovered that the effectiveness or performance of a machine learning model cannot be solely determined by the number of parameters it possesses. The appropriate number of parameters is influenced by several factors, including the complexity of the problem, the size of the dataset, and the computational resources available.

Merely imposing a limitation of 500 million parameters for all machine learning models may not be sufficient. Therefore, it is crucial to include an exception clause in this section, granting permission to important sectors to utilize AI machine learning models that exceed the specified limits.

For instance, in the medical sector, machine learning models are extensively employed in applications such as medical imaging using technologies like MRI, CT scans, and X-rays. These models possess the capability to analyse vast amounts of electronic health records to extract valuable insights. Imposing strict limitations on parameters without exceptions would significantly impede progress in this sector.

Hence, it is imperative to consider the unique requirements of each sector and incorporate flexibility in the parameter limitations to ensure continued advancements and innovations.

1. In article XII, all the important checkpoint are placed, which at this point of time seems necessary requirements for the exit strategy, but it can have other specific requirements too. For example- The Level of awareness among the people, the capability of AI at the time of discussion about the exit strategy.
2. The place, numbers, duration, and other relevant aspects mentioned in this document are fictional and have been created by me for demonstration purposes. In reality, these factors would be determined by the participating states during the negotiation process.
3. I sincerely apologize in advance for any mistakes or errors that may be found in this document. Despite my best efforts to ensure accuracy and comprehensiveness, there is always a possibility of unintended errors or omissions. I appreciate your understanding and kindly request your diligent review and verification of the content for grammatical correctness and appropriateness.
4. Just for clarity, it is important to note that while the content of the treaty is contained within the eight pages of the articles, an additional page is allocated for the cover page, one page for the table of contents, two pages for the annexes, and one page for the notes. These additional pages, including the cover page, table of contents, annexes, and notes, are not counted within the specified page limit for the Draft. Therefore, the overall document would consist of the specified number of pages for the articles, along with the additional pages for the aforementioned sections, which are exempt from the page limit.