**TREATY**

**ON** **MORATORIUM OF LARGE-SCALE**

**AI CAPABILITIES RESEARCH AND DEVELOPMENT**

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

Considering the threat that would be visited upon all mankind by the Artificial Intelligence and the consequent need to make every effort to avert the danger of such a security threat and to take measures to safeguard the data security of peoples,

Believing that the further Research and Development of Artificial Intelligence would seriously enhance the danger to privacyand data security,

In conformity with resolutions of the XYZ Assembly calling for the conclusion of an agreement on the moratorium of large-scale Artificial Intelligence capabilities, research and development,

Undertaking to co-operate in facilitating by safe and responsible application of Artificial Intelligence,

Expressing their condemn for research, development and other efforts to further the application, as there are concerns and potential threats associated with Artificial Intelligence such as Job Displacement, Non-Accountability, Privacy and Security, Autonomous control, Socioeconomic Disparities and Dependence on AI,

The necessity for a moratorium on large-scale artificial intelligence capabilities, research and development arises from the recognition of potential risks and the desire to ensure responsible and safe technological advancement,

Large-scale AI capabilities have the potential to significantly impact society, raise ethical dilemmas, and potentially violate human rights. A moratorium allows time for comprehensive ethical frameworks to be established, ensuring AI systems are developed and deployed in a manner that respects fundamental values and protects against discriminatory or harmful outcomes,

Advanced AI systems may possess unforeseen risks and unintended consequences. A moratorium provides an opportunity to conduct thorough research and safety assessments, identifying potential hazards and designing robust safeguards. This proactive approach reduces the likelihood of AI systems causing harm or catastrophic failures,

The complexity of large-scale AI capabilities necessitates effective governance and regulation. A moratorium enables policymakers, experts, and stakeholders to develop appropriate legal and regulatory frameworks. This includes addressing issues such as accountability, transparency, data privacy, and the deployment of autonomous AI systems,

Its development is a global endeavor. A moratorium fosters international collaboration, allowing countries to work together in defining common standards, sharing best practices, and collectively addressing global challenges associated with AI. It promotes cooperation and prevents a fragmented landscape where varying regulations and approaches create potential inequalities and risks,

This technology impacts society at large. A moratorium provides an opportunity for public engagement and democratic deliberation on the future of AI. It allows for education, awareness campaigns, and inclusive discussions involving diverse stakeholders to ensure that AI development aligns with societal values and concerns,

Overall, a moratorium on large-scale AI capabilities research and development serves as a responsible and proactive measure to address potential risks, establish ethical guidelines, promote safety, ensure effective governance, and foster international collaboration and public engagement in shaping the future of AI. It allows for a more considered and inclusive approach to harnessing AI's potential benefits while mitigating its potential harm.

Convinced that, in furtherance of this principle, all Parties to the Treaty are entitled to participate in the fullest possible exchange of scientific information for, and to contribute alone or in co-operation with other States to, condemn further development of the applications of Artificial Intelligence,

Declaring their intention to achieve the objective at the earliest and to undertake effective measures in the direction of suspending further research and development on AI,

Urging the co-operation of all States in the attainment of this objective,

The participating parties, recognizing the rapid advancement and potential risks associated with large-scale artificial intelligence (AI) capabilities, and desiring to ensure the responsible and safe development of AI technologies, hereby establish this Treaty on Moratorium of Large-Scale AI Capabilities Research & Development,

Have agreed as follows:

**ARTICLE I**

**DEFINITIONS**

**I.I** Artificial Intelligence: Artificial intelligence (AI) refers to the field of computer science and technology that aims to create intelligent machines capable of performing tasks that typically require human intelligence. AI systems are designed to perceive, reason, learn, and make decisions, often mimicking human cognitive abilities.

**I.II** Moratorium: Moratorium refers to a temporary suspension or prohibition of large-scale AI capabilities research and development.

**I.III** Large-Scale AI Capabilities: Large-Scale AI Capabilities refers to research and development activities focused on the creation and deployment of highly advanced AI systems with the potential for significant impact on human society, including but not limited to superintelligent AI, autonomous weapons, and AI systems with extensive decision-making capabilities.

**I.IV** Research and Development: Research and development (R&D) of artificial intelligence (AI) refers to the scientific and technological activities aimed at advancing the understanding, capabilities, and applications of AI systems. It encompasses a broad range of activities undertaken to improve AI algorithms, models, frameworks, and methodologies to achieve specific objectives.

**I.V** Data: Data refers to the information that is used to train, validate, and optimize AI models and algorithms. Data is a fundamental component of AI systems and plays a crucial role in enabling machines to learn and make intelligent decisions.

**I.VI** Humans: Humans refer to individuals or the human participants involved in the development, use, and interaction with AI systems. Humans play multiple roles in the AI domain, and their involvement can vary depending on the specific context and application.

**I.VII** Security threat: Security threat refers to any potential risk or vulnerability that can compromise the confidentiality, integrity, or availability of AI systems, their data, or the broader ecosystem in which they operate. These threats can arise from various sources and can have significant consequences if not properly addressed.

**I.VIII** Parties: Parties refer to the states, international organizations, or entities that are signatories to or have ratified the treaty. These parties are the entities that have voluntarily chosen to be bound by the terms and obligations outlined in the treaty.

**ARTICLE II**

**MORATORIUM ON LARGE-SCALE AI CAPABILITIES RESEARCH & DEVELOPMENT**

**II.I** The participating parties agree to implement a moratorium on large-scale AI capabilities research and development for a period of 5 years from the date of this treaty's entry into force.

**II.II** The moratorium shall apply to all participating parties and their affiliated entities involved in research, development, funding, or deployment of large-scale AI capabilities.

**II.III** During the moratorium period, participating parties shall refrain from engaging in large-scale AI capabilities research and development, except for activities related to safety, ethics, and governance of AI technologies, which are encouraged and supported.

**II.IV** The agreement could include a comprehensive ban on the research and development of specific AI capabilities deemed too risky or ethically questionable. This ban could be targeted at areas such as superintelligent AI, autonomous weapons, or AI systems with significant decision-making power.

**II.V** The agreement could impose restrictions on public or private funding for large-scale AI capabilities research and development. This would limit financial support for projects that fall within the scope of the agreement, discouraging further advancements in those areas.

**II.VI** The agreement could include regulations on the transfer of AI-related technologies and knowledge between the signed parties. This would aim to prevent the dissemination of AI capabilities that are subject to the research and development restrictions outlined in the agreement.

**II.VII** The agreement could establish mechanisms for monitoring and oversight to ensure compliance with the research and development restrictions. This may involve regular reporting, inspections, and audits to verify that parties are adhering to the agreed-upon limitations.

**II.VIII** The agreement could encourage and support collaborative research initiatives focusing on AI safety, ethics, and responsible AI development. Parties could work together to explore best practices, develop guidelines, and share knowledge on the responsible use of AI technologies.

**ARTICLE III**

**INTERNATIONAL COLLABORATION AND COOPERATION**

**III.I** The participating parties commit to fostering international collaboration and cooperation in the field of AI safety and ethics during the moratorium period.

**III.II** Participating parties shall share information, best practices, and lessons learned regarding the responsible development and use of AI technologies.

**III.III** Joint initiatives and research projects focused on AI safety, ethics, and risk mitigation shall be encouraged and supported among participating parties.

**ARTICLE IV**

**REVIEW AND ASSESSMENT**

**IV.I** A periodic review and assessment of the moratorium's impact and effectiveness shall be conducted by an independent body established by the participating parties.

**IV.II** The review shall consider advancements in AI technology, potential risks, societal implications, and the progress made in addressing safety, ethical, and governance concerns.

**IV.III** Based on the review's findings, the participating parties may decide to extend, modify, or terminate the moratorium, considering the evolving landscape of AI technology and its implications.

**ARTICLE V**

**DISPUTE RESOLUTION**

**V.I** Any disputes arising from the interpretation or implementation of this treaty shall be resolved amicably through consultations and negotiations among the participating parties.

**V.II** If disputes cannot be resolved through consultations, the participating parties may agree to seek mediation or arbitration by an independent body or organization.

**ARTICLE VI**

**ENTRY INTO FORCE AND RATIFICATION**

**VI.I** This treaty shall enter into force upon ratification by participating parties on 1st August,2023 representing a significant portion of global AI research and development capabilities.

**VI.II** Each participating party shall notify the depository of this treaty regarding their ratification, and the depository shall communicate this information to all participating parties.

**ARTICLE VII**

**WITHDRAWAL**

**VII.I** All participating parties may withdraw from this treaty after providing 1 year of prior notice to all other participating parties.

**VII.II** In the event of withdrawal, participating parties shall provide a detailed justification for their decision and continue to cooperate in AI safety and ethical considerations to minimize potential risks associated with their ongoing AI capabilities research and development.

**VII.III** All Parties shall have the right to withdraw from the Treaty in exercising their national sovereignty if it determines that extraordinary events linked to the subject matter of this Treaty have compromised its country's supreme interests. It shall give three months' notice of such withdrawal to the other Treaty Parties and the XYZ Council. Such a warning must include a description of the extraordinary occurrences that it believes have compromised its supreme interests.

**VII.IV** A conference shall be convened five years following the Treaty's entrance into force to decide whether the Treaty shall continue in force indefinitely or for an extra set period or periods. This decision will be made based on the consent of Treaty Parties.

**ARTICLE VIII**

This Treaty, the English, Russian, French, Spanish and Mandarin texts of which are equally authentic, shall be deposited in the archives of the Depositary Governments. Duly certified copies of this Treaty shall be transmitted by the Depositary Governments to the Governments of the signatory and acceding States.

*IN WITNESS WHEREOF the undersigned, duly authorized, have signed this Treaty.*

*DONE in duplet, at the cities of Queensland and Ferryland, the first day of August, two thousand twenty-three.*