### **Preamble**

The Parties to this Treaty,

Recognising the rapid advancement of Artificial Intelligence (AI) technologies and their potential to bring about profound changes in all aspects of human life,

Acknowledging the potential risks associated with the unregulated development and deployment of AI, particularly in the form of Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI),

Understanding the need for a coordinated international response to address these challenges and to ensure the safe and ethical development and use of AI,

Believing that a temporary pause in the further development of AI, pending the establishment of appropriate international governance and legal structures, is a necessary and prudent step,

Inspired by the principles and spirit of the article "Pausing AI Developments Isn’t Enough. We Need to Shut it All Down,"

Hereby agree to the following provisions in the interest of global safety, security, and the common good.

### **Article 1: Definitions**

For the purposes of this Treaty:

"Artificial Intelligence (AI)" refers to the development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

"Artificial General Intelligence (AGI)" refers to a type of artificial intelligence that has the ability to understand, learn, and apply its intelligence to any intellectual task that a human being can.

"Artificial Superintelligence (ASI)" refers to a form of artificial intelligence that surpasses human intelligence across most economically valuable or practically relevant fields, potentially leading to an intelligence explosion.

"Large-scale AI capabilities" refers to the development and deployment of AI systems that have the potential to significantly impact society or the environment, either directly or indirectly.

"Quantum computers" refers to a new type of computing device that uses the principles of quantum mechanics to process information, potentially enabling faster and more complex computations than traditional computing devices.

### **Article 2: Prohibition of Large-Scale AI Capabilities**

The Parties agree to cease the operation of large GPU and TPU clusters, defined as those consisting of more than 1000 processing units, used for the refinement of powerful AI systems. This includes any computational infrastructure primarily used for the training, testing, or deployment of AI models. This measure is taken to prevent the creation of AI systems that could potentially exceed our ability to control or understand them due to their computational power.

The Parties agree to prohibit the training of Machine Learning (ML) models, or combinations of models, with more than 500 million parameters. This limit is set based on current understanding of AI model complexity and is intended to prevent the creation of excessively powerful and potentially uncontrollable AI systems. The limit is subject to review and adjustment by the International AI Safety Body as our understanding of AI systems evolves.

### **Article 3: Prohibition of Quantum Computing in AI Activities**

The Parties agree to prohibit the use of quantum computers in any AI-related activities. This includes, but is not limited to, the development, training, testing, and deployment of AI models. This prohibition is based on the potential for quantum computing to significantly accelerate AI capabilities beyond our ability to control or understand them. The potential risks posed by the intersection of quantum computing and AI warrant a cautious approach until a robust framework for safe use is established.

### **Article 4: Moratorium on AI Capabilities Research and Development**

The Parties agree to a general moratorium on large-scale AI capabilities research and development. This pause will remain in effect until it is universally agreed that it is safe and ethical to resume such activities. "Large-scale AI capabilities research and development" refers to projects that aim to significantly advance the state of AI beyond current capabilities, particularly in ways that could pose risks to safety, security, or societal stability. The moratorium is intended to provide time for the development of robust safety and ethical guidelines and does not preclude research aimed at improving the safety and ethical use of AI.

### **Article 5: Criminalisation of AGI and ASI Development**

The Parties agree to pass national laws criminalising the development of any form of AGI or ASI. This measure is intended to prevent the emergence of AI systems that could pose a threat to human safety or societal stability. The specifics of these laws, including definitions, penalties, and enforcement mechanisms, should be determined in accordance with each Party's legal and regulatory framework, with guidance from the International AI Safety Body.

### Article 6: Establishment of an International AI Safety Body

The Parties agree to establish an International AI Safety Body (IASB) to oversee the implementation and enforcement of this Treaty. The IASB will be an independent, international organisation with a mandate to promote the safe and ethical development and use of AI.

The IASB will be composed of representatives from each of the Parties, as well as experts in the fields of AI, law, ethics, and international relations. Representatives will be chosen through a transparent and democratic process, with the aim of ensuring broad representation and expertise.

The IASB will have the power to monitor compliance with the Treaty, investigate alleged violations, and recommend sanctions for non-compliance. It will also provide guidance on the safe and ethical development and use of AI, and will facilitate international cooperation and knowledge sharing in this field.

The IASB will be funded by contributions from the Parties, supplemented by donations from other sources. It will operate transparently, with regular reporting to the Parties and the public.

### Article 7: Enforcement Mechanisms and Penalties

The IASB will be responsible for monitoring compliance with this Treaty. This will include regular inspections of AI development facilities and audits of AI research activities. The IASB will develop detailed protocols for these inspections and audits, with the aim of ensuring thoroughness, fairness, and respect for national sovereignty.

In the event of a suspected violation of this Treaty, the IASB will conduct a thorough investigation. If a violation is confirmed, the IASB will recommend appropriate sanctions. These sanctions may include fines, suspension of AI research activities, and in severe cases, referral to the International Court of Justice. The process for determining sanctions will be transparent and will include opportunities for the accused Party to present a defence.

### Article 8: Duration and Termination

This Treaty shall remain in effect indefinitely, until it is universally agreed by the Parties that it is safe and ethical to resume large-scale AI capabilities research and development.

Any Party may propose the termination of this Treaty. Such a proposal must be approved by a two-thirds majority of the Parties.

Finally, we'll include the final provisions, such as how the treaty can be amended and the process for ratification.

### Article 9: Amendments and Ratification

Any Party may propose amendments to this Treaty. Proposed amendments must be approved by a two-thirds majority of the Parties.

This Treaty shall be subject to ratification by the Parties in accordance with their respective constitutional processes. The instruments of ratification shall be deposited with the Secretary-General of the United Nations.

This Treaty shall enter into force upon the deposit of instruments of ratification by two-thirds of the Parties.