The Parties to this Agreement,

Recognizing the dangers of artificial intelligence, hereinafter referred to as AI, to both the social and economic welfare of the planet as well as the urgent need to stall its breakneck development in order to ensure the general safety of society,

Also recognizing the lack of sufficient oversight expended upon AI and the consequent unpredictability surrounding its behavior,

Gravely concerned with the amplification of pre-existing socioeconomic biases due to obscurities in the decision-making process as well as stringent risks in the areas of ethics, privacy, and security,

Deeply concerned with the development of GPU & TPU farms due to their being tools that proliferate the development of highly advanced, and therefore potentially harmful AI,

Recognizing that the development of artificial general intelligence (AGI) or artificial super intelligence (ASI) in an unrestricted manner could potentially result in rogue AI that pose existential threats to the human civilization,

Recognizing the threat of machine learning models trained on colossal amounts of parameters,

Concerned about the aforementioned tools and methods that cause the proliferation of harmful AI,

Believing that the elimination of tools that contribute to the unrestrained and harmful growth of AI will be executed through a thorough and holistic lens that simultaneously makes the threats of AI less eminent,

Believing that public perception of AI become clearer and not hidden behind the complexity of the topic in order to disseminate information regarding the dangers of AI,

Believing that an international body for oversight be established that focuses on the appropriate extension of measures to control AI, upholding all current measures by focusing on party compliance and resolution of disputes, enforcing such compliance for all pertinent parties, and approving projects in such a manner that development in particularly large-scale AI is stunted,

Affirming that a completely restricted approach to AI development that focuses on present safety and future risk assessment is in the best interest of all parties,

Considering the significance of AI-based research for institutions and a plethora of industries as well as the importance of continued innovation in these fields, it is necessary that a restriction on AI development not implode research, but rather encourage an exploration of the application of current capabilities in numerous contexts,

Convinced that in order to secure a future where AI technology is safely compatible to all spheres of society, all Parties of this Treaty are to engage in the exchange of scientific development and information in the relevant field to their maximum capacity so as to ensure that no stride is counted out, and that AI developments are tightly reviewed across parties to comply with shared safety standards,

Encouraged by the efforts of multiple entities and governments to establish guidelines to restrict AI development in order to streamline more safe applications of technology,

Urging that all parties cooperate in order to achieve a state of safety in the field,

Recalling the commendable efforts of already established frameworks and laws to control AI to achieve a safer space for society and innovation,

Pursuant of a global moratorium on the development of large-scale AI projects until a proper and safe evaluation is reached on the safety and ethicality of such developments,

Have agreed to the following:

**Article I**

1. All Parties to the Treaty undertake to shut down GPU and TPU clusters dedicated to training large-scale AI models. In addition to the shutting down of centralized computing farms, access to resource-heavy decentralized cloud platforms will be restricted for AI-based applications.
   1. All Parties to the Treaty will implement standard protocols such that the use of GPU and TPU clusters will dictate an application to determine eligibility for use. Applications regarding the development of large-scale AI models will be restricted from having access to the computing farm.
2. The manufacturing of Tensor Processing Units (TPUs) and other specialized AI accelerator hardware that bolsters machine learning use will be restricted. Further research and development beyond TPU v5 will be barred.
   1. All Parties to the Treaty will, to their maximum capacity, collaborate in order to limit the total supply of advanced TPUs being offered so as to mitigate the overall risk of large-scale AI models being developed.

**Article II**

1. The development of machine learning models trained on 500 million or more parameters is to be strictly forbidden.
   1. All Parties to the Treaty will mandate reports be made by every major technology company or entity that leverages advanced AI models. These reports are to disclose major machine learning models as well as the number of parameters they are trained on.
   2. Should models exceed to 500 million benchmark, model controllers will be advised to reduce parameter count with techniques such as pruning and knowledge distillation. After an extensive audit, if the parameter count is not reduced after a 1-week (7 day) period from the initial disclosure, then the machine learning model should be shut down indefinitely.
   3. In order to further discourage models that leverage high parameter counts, all Parties to the Treaty will be required to form individual MPC groups (Model Parameter Compliance) that regularly audit models.
   4. Similar to the EU AI Act, model risk will be categorized as “low-risk”, “medium-risk”, or “high-risk”. As an extension of 1.1b, models that demonstrate high risk will be shut down regardless of their parameter count, while models near the parameter limitation that are considered “medium-risk” or lower will be given a 7-day period to reduce parameter count.
   5. As an extension to 1.1c, public awareness to the dangers of models with such large parameter counts is encouraged to be fostered. Public campaigns focused on AI ethics, policy, and regulation will be encouraged heavily.

**Article III**

1. For all Parties to the Treaty, all developments in the domains of artificial general intelligence (AGI) and artificial super intelligence (ASI) are to be halted immediately.
2. All Parties to the Treaty are encouraged to work together in order to create mechanisms that can prohibit complete self-learning to be achieved within AI models so that a singularity event, where the behavior of AI models is so advanced that it becomes unpredictable, may be evaded.
   1. All Parties to the Treaty are encouraged to collaborate specifically on A.I. alignment, which works to align computer system behavior to principle human moral values.

**Article IV**

1. All Parties to the Treaty are to strictly denounce and ban AI-abetted crimes.
2. If AI is used as an accessory to crime, it is to be treated as an innocent agent IF it can be determined that the primary human perpetrator was not influenced by the AI.
3. If the AI committed a crime by itself, an investigation will be launched in order to identify both the actus reus and the mens rea. If both can be convincingly attributed to the AI and not any other external party, then it is to be shut down immediately, and a further investigation into the entity responsible for the AI is to be launched in order to determine for further criminal charges.
4. In order to rectify the grievances of those affected by crimes committed by AI or in which AI has a role, all Parties to the Treaty agree to offer compensations proportional to the degree of implication.

**Article V**

1. All Parties to the Treaty who are found to be in violation of the treaty should be subject to exclusion in discussion regarding the development within the field. Furthermore, said countries may be subject to various diplomatic and trade sanctions according to the discretion of all countries in good standing with the treaty.
2. A compliance board will be established in order to monitor the actions of each country and determine if those actions align with the values of the treaty. Compliance assessments will be completed on a periodic basis in order to determine a country’s standing with the treaty.
   1. Countries found to be lacking in the upholding of the standards of this treaty by the compliance assessment and by other actions taken at the discretion of the compliance board may be subject to the penalties mentioned in 5.1.
3. All Parties to the Treaty are highly encouraged to be in regular communication with each other in order to achieve compliance among all countries.
4. As an extension of 5.2, ethical reviews will become key aspects of compliance assessments. Ethical reviews determine the ethicality of the legislative decisions countries put into place, as well as how well those decisions align with the principles and values for which this treaty is constituted upon.
   1. Countries found to be lacking in the upholding of the standards of this treaty by the ethics review and by other actions taken at the discretion of the compliance board may be subject to the penalties mentioned in 5.1.

**Article VI**

1. In order to address and resolve any disputes among member countries, a third-party interdisciplinary panel composed of technical experts in AI as well as ethical experts will act as arbitral tribunal.
   1. The ruling of said panel is to be binding, and the parties in question waive rights to pursue litigation.
   2. The goal of the panel specifically is to provide guidance and recommendation to relevant parties for effective dispute resolution.
2. Fellow Parties to the Treaty are encouraged to help mediate negotiations between the countries in order to achieve an amicable settlement.

**Article VII**

1. Regular discussion among Parties to the Treaty is encouraged in order to determine the standing and development of the safe and ethical development of AI.
   1. Such discussion is imperative to the functionality of this treaty, for as long as AI development is concluded to not conform to safe and ethical standards, the necessity of this treaty exists.
2. Parties to the Treaty are encouraged to draft and enforce domestic legislation that reflects the core principles of this treaty. Such legislation will allow for a coexistence of global power and national autonomy.
3. Similar to a measure extended by the United Nations Convention Against Corruption (UNAC), protection will be established for people who come forward with violations committed by any Party to the Treaty. Those who demonstrate being in possession of insider knowledge that implicates a party for being in violation of the compliance standards will be incentivized so as to make enforcement procedure more robust.
4. As a part of bolstering enforcement procedure, all countries found to be in violation of this treaty by discreetly investing in the development large-scale AI models will be subject to the various penalties outlined in 5.1.
   1. Any information transmitted between Parties to the Treaty regarding the furthering of the capabilities of large-scale AI is discouraged unless done so in the light of creating preventative measures.

**Article VIII**

1. A globally supervising authority on AI is to be established through the World AI Authority Institute (WAAI)
   1. The goal of WAAI is to supervise developments in AI and to specifically focus on creating standards for AI safety as well as standards for stable innovation
   2. The board of WAAI will be composed of elected members who specialize in a myriad of fields, including but not limited to: AI, Philosophy, and Cybersecurity.
   3. WAAI has the authority of approving of various AI projects, and is specifically tasked with monitoring and restricting development in large-scale and powerful AI.
2. This supervising authority will also be responsible for the adoption of various soft laws.

Soft laws refer to quasi-legal tools that, although are not legally binding, are considered to be etiquette or standard for specific practices. By adopting a ready set of soft laws for AI, development in the future will become safer.

**Article IX**

1. All AI-based projects from official entities, such as corporations, requires approval from a branch of the WAAI, a supervising power in AI development, established by this treaty.
2. Continued development of the project despite failure to obtain approval from the WAAI is subject to punishment as outlined by 5.1.
3. In the case that an application for a project is rejected, that decision may be appealed for reconsideration.
4. If the application for a project is accepted, but the scope or contents of the project change substantially, reapplication is necessary. Failure to reapply under these circumstances is punishable as outlined by 5.1.
5. Parties to the Treaty are unable to interfere and advocate for the acceptance of specific projects.

**Article X**

1. WAAI branches will be encouraged to form at provincial levels in order to gain access to more information and obtain a better understanding of the development of AI and AI safety standards.
2. Communication among these branches will not be strictly hierarchical so that more interconnectivity is present, and more communication among branches is fostered so that AI safety can be more thoroughly evaluated.

**Article XI**

1. Funds will be allocated and dispersed proportionally to the Parties to the Treaty so as to improve public comprehension of AI and to improve awareness of the dangers it poses.
2. These funds will be spent on whatever the party sees fit to spend it on as long as it directly improves public understanding of the ethical and safety threats that AI poses. Examples for funding may include but is not limited to the initiation of community classes, marketing campaigns, and more.

**Article XII**

1. All Parties to the Treaty may propose amendments to this treaty at any time. The addition of said amendment is contingent on a majority vote from all parties involved.
2. Any Party to the Treaty may withdraw from the treaty. Withdrawal must be notified to the compliance board at least 3 months in advance, and compliance to the treaty is still expected for 1 year after withdrawal.
3. Every year a conference will be held in-person in Canberra, Australia for the discussion of the progression of AI safety standards. The moratorium outlined by this treaty is to be upheld for at least 5 years, after of which a grand conference will be held in order to determine the state of AI safety and the continued relevancy of the treaty.
   1. After 5 years, if it is determined that this treaty is still necessary (by a majority vote from Parties to the Treaty) in order to prevent the proliferation of harmful AI applications, then a decision regarding the upholding of this treaty is to occur every one year after that at the Canberra conference.
   2. After 5 years, if it is determined that this treaty is no longer necessary (by a majority vote from Parties to the Treaty) at the grand conference in Canberra, then this treaty is no longer in effect.

Works Cited:

1. Ayling, Jacqui, and Adriane Chapman. “Putting AI Ethics to Work: Are the Tools Fit for Purpose?” *AI and Ethics*, vol. 2, no. 3, 12 Sept. 2021, pp. 405–429, https://doi.org/10.1007/s43681-021-00084-x.
2. Coeckelbergh, Mark. *Ai Ethics*. The MIT Press, 2020.
3. Darrell M. West, John R. Allen, et al. “Soft Law as a Complement to AI Regulation.” *Brookings*, 9 Mar. 2022, www.brookings.edu/articles/soft-law-as-a-complement-to-ai-regulation/.
4. Fortner, Robert S. *Ethics in the Digital Domain*. Rowman & Littlefield, 2021.
5. “Home.” *The Artificial Intelligence Act*, 3 Apr. 2023, artificialintelligenceact.eu/.
6. Jobin, Anna, et al. “The Global Landscape of AI Ethics Guidelines.” *Nature Machine Intelligence*, vol. 1, no. 9, 2 Sept. 2019, pp. 389–399, https://doi.org/10.1038/s42256-019-0088-2.
7. LeaderGPU. “GPU Farm.” *LeaderGPU*, 25 Dec. 2019, www.leadergpu.com/gpu\_farm.
8. Marr, Bernard. “The 15 Biggest Risks of Artificial Intelligence.” *Forbes*, 5 June 2023, www.forbes.com/sites/bernardmarr/2023/06/02/the-15-biggest-risks-of-artificial-intelligence/?sh=49b275c52706.
9. Morley, Jessica, Anat Elhalal, et al. “Ethics as a Service: A Pragmatic Operationalization of Ai Ethics.” *Minds and Machines*, vol. 31, no. 2, June 2021, pp. 239–256, https://doi.org/10.1007/s11023-021-09563-w.
10. Morley, Jessica, Libby Kinsey, et al. “Operationalizing Ai Ethics: Barriers, Enablers and next Steps.” *AI &amp; SOCIETY*, vol. 38, no. 1, 15 Nov. 2021, pp. 411–423, https://doi.org/10.1007/s00146-021-01308-8.
11. Nyuytiymbiy, Kizito. “Parameters and Hyperparameters in Machine Learning and Deep Learning.” *Medium*, 28 Mar. 2022, towardsdatascience.com/parameters-and-hyperparameters-aa609601a9ac.
12. “Supported Reference Models | Cloud TPU |  Google Cloud.” *Google*, cloud.google.com/tpu/docs/tutorials/supported-models. Accessed 16 July 2023.
13. ÓhÉigeartaigh, Seán S., et al. “Overcoming Barriers to Cross-Cultural Cooperation in AI Ethics and Governance.” *Philosophy &amp; Technology*, vol. 33, no. 4, 15 May 2020, pp. 571–593, https://doi.org/10.1007/s13347-020-00402-x.