



SELECT COMMITTEE ON THE CCP

DEMOCRATS | RANKING MEMBER RO KHANNA

September 24, 2026

The Honorable Marco Rubio
Secretary
U.S. Department of State
2201 C Street NW
Washington, D.C. 20520

Dear Secretary Rubio:

Yesterday, I convened the first-ever congressional hearing calling for a binding, enforceable, and verifiable treaty between the United States and the People's Republic of China (PRC) on catastrophic AI risks. I came away from this hearing with two key takeaways. **First, there are a number of steps we could pursue today to begin mitigating these risks.¹ This includes establishing technical working groups between our countries.** Critically, we must also undertake a whole-of-society effort to develop verification technologies, including tools that can reliably differentiate inference from training, to ensure any future agreement between the U.S. and China is enforceable. This is particularly important in light of the PRC's track record of not living up to its word. **Second, looking into the future, I came away from this hearing even more convinced that an enforceable and verifiable U.S.-China AI Treaty is not only in the interests of both of our countries, but all of mankind.** In the case of AI, we do not have the luxury of inaction and diplomacy must move at the speed of technology.

Although it appears President Trump unfortunately opted to not commence meaningful discussions with General Secretary Xi Jinping on this topic today, the imperative of pursuing diplomacy on this issue remains. As I stated at the hearing, while America and the PRC have different values, catastrophic AI risks will not differentiate between American citizens or Chinese citizens, or even between democracy and authoritarianism. If humans lose control of AI, or if misaligned superintelligence decides that the interests of mankind differ from its own, the very fate of humanity rests in the balance. No country benefits if a non-state actor or a rogue agentic model collapses the global Internet or unleashes biological havoc.

Accordingly, I write to share my perspective on some of the key provisions that should be discussed in the context of pursuing a treaty or other international agreement between the U.S. and China on artificial intelligence. While this list is not meant to be comprehensive, and it does not cover the initial mechanisms that would help set the stage for seeking a broader pact, it is meant to ensure we begin the process of pursuing long-term consensus. We need to begin exploring ideas that not only foster communication between our respective countries, but also address the underlying risks that make such dialogue essential.

¹ Testimony of Dr. Geoffrey Hinton and Lt. General Jack Shanahan (USAF Ret.) at Shadow Hearing on a U.S.-China AI Agreement before the House Select Committee on China. 119th Cong. (2026).



SELECT COMMITTEE ON THE CCP

DEMOCRATS | RANKING MEMBER RO KHANNA

First, we need a global pause on recursive self-improvement (RSI). To oversimplify, this is when AI itself trains AI. Leading scholars believe it is the most likely path to humans losing control of this technology. Entrusting AI to train itself is analogous to letting the fox guard the henhouse.

Second, we need a global pause on superintelligence. As it has been traditionally defined, this is when AI exceeds human capabilities across all domains. Society simply is not ready for the transformations it could unleash, and no country alone should have the ability to usher in this potentially brave new world.

Third, both the U.S. and China should agree on common protocols for kill switches and require them across all advanced models. Even with a pause on RSI, we cannot guarantee advanced models will not escape human control. If this happens, we need a safety hatch to give our species a shot at reasserting authority over these autonomous systems.

Fourth, every lab in the world should be required to audit their models for alignment and catastrophic misuse in an air-gapped sandbox before releasing them. No country should be allowed to unleash rogue agents on the world. We saw a preview of that future with OpenAI's hack of Hugging Face. Similarly, particularly in light of Chinese open weight models' propensity for forsaking basic safety precautions, we must ensure no model in the world can allow non-state actors to design or unleash biological, chemical, or nuclear havoc. As a part of this fourth pillar, we must establish common sense global bans on using AI to weaponize human biology.

Fifth, we must ensure any U.S.-China agreement on AI is worth the paper it is written on. I cannot overstate the importance of this principle. The PRC is notorious for its failure to implement international commitments, and the stakes are too high in the context of catastrophic AI risks to permit such defiance of international norms. In the words of President Reagan, we must "trust but verify." This means investing in verification technologies and inspection regimes. To complement such tools, we should begin the process of forming and ensuring a leading American role in an International AI Inspection Agency modeled on the International Atomic Energy Agency, or alternatively, a comparable non-governmental organization that could inspire trust from both sides. It would need access to both of our data centers, so it can alert the world if China is using chips to train prohibited models rather than using chips for inference. The time is now to establish technical working groups that can develop the most effective and trustworthy verification mechanisms possible. Our ability to pursue the preceding four pillars depends on supporting the innovation necessary to solve this pressing technical challenge. The same intellectual firepower that is being directed towards model development should be channeled to global guardrails verification.

Although AI is new, the model for how to address its risks on a global scale is not. In 1963, the United States, United Kingdom, and Soviet Union agreed to a Limited Test Ban Treaty to not test nuclear weapons in the atmosphere. In 1996, the world agreed to the Comprehensive Nuclear Test Ban Treaty. We have seismic stations around the world that monitor if anyone violates these agreements. In the late 1980s, President Ronald Reagan and Soviet leader Mikhail Gorbachev signed the Intermediate-Range Nuclear Forces Treaty. It was accompanied by a system of on-site inspections at missile bases that made violating the agreement extremely difficult, and it arguably



SELECT COMMITTEE ON THE CCP

DEMOCRATS | RANKING MEMBER RO KHANNA

saved us from a war in Europe. Just like America and the USSR set aside our vast differences to save the world from nuclear war, the U.S. and China share a common interest in ensuring humans never lose control of AI.

While conversations between the U.S. and China are an important first step, we must also engage in global, multilateral dialogue about AI risks. Washington and Beijing may currently hold a duopoly on advanced models, but in the long term, any durable solution will require consensus among all the world's powers. As Secretary of State, you are uniquely positioned to advance this dialogue into that next phase.

Some issues are bigger than any one country or the geopolitical or domestic political dynamics of any given moment. History will look back at this juncture, the choices made by humanity as we enter this new era of AI, and either celebrate our foresight or condemn our inaction. For the sake of humanity, I urge you to join me in pursuing the U.S.-China and global AI guardrails that will ensure our fate remains our own.

Sincerely,

Ro Khanna
Ranking Member
House Select Committee on the CCP