From Risk to Regulation: Navigating US-China Trust Gaps and Risk Mismatches in Military AI

Authors

Jean Dong[1], Ahmed Mehdi Inane[2], Riley Simmons-Edler[3,4 *], Ryan Paul Badman[3,4 *]

[1]Harvard Kennedy School, MA, USA, [2]Mila, Quebec, Canada, [3]Harvard Kempner Institute, MA, USA, [4]Harvard Medical School, MA, USA, *equal contributions

Abstract

Lethal Autonomous Weapons (LAWS), defined by the United Nations as systems capable of independently identifying, selecting, and engaging targets without human intervention, represent a transformative yet controversial domain of modern military technology. Since the rise of systems like the Predator drone and earlier autonomous units such as the Israeli Harpy, international calls for regulation have grown. However, most regulatory efforts, including those led by the UN and NGOs, have focused on blanket bans, which global powers often find overly restrictive and impractical. As a result, no significant global agreements on LAWS have been reached.

The proliferation of AI-powered LAWS in conflict zones, combined with advancements in military AI, has exacerbated the urgency for effective regulation. Current proposals have struggled to keep pace with AI advancements, often failing to distinguish between AI-enabled autonomous systems and older, remote-controlled technologies. Diverging perspectives from global superpowers further complicate the regulatory landscape: while the United States has resisted nuanced definitions, China has emphasized regulating only the most advanced AI-enabled LAWS. These fragmented approaches and policy failures reflect the broader breakdown of dialogue, rooted in mismatched categorizations and inadequate assessments of AI's unique challenges relative to traditional warfare.

This paper identifies trust gaps and risk mismatches in military AI between China and the United States as well as military AI applications that pose significant ethical and security risks, advocating for immediate international consensus to ban or limit these high-risk cases. By establishing clear norms and precedents, the global community can build a foundation for more comprehensive regulation of near-future military AI technologies. This approach aims to bridge trust gaps, mitigate the arms race between the U.S. and China, and promote responsible governance of military AI.