Research on Artificial Intelligence and International Security 2026

SUMMARY AND KEY OUTCOMES China–Europe Track II Dialogue on the Impact of Artificial Intelligence on International Security

April 2026

清华大学战略与安全研究中心
CENTER FOR INTERNATIONAL SECURITY AND STRATEGY
TSINGHUA UNIVERSITY

Centre for Humanitarian Dialogue

SUMMARY AND KEY OUTCOMES
China–Europe Track II Dialogue on the Impact of Artificial Intelligence on International Security

I. Overview

From 2024 to 2025, the Center for International Security and Strategy (CISS) at Tsinghua University and the Centre for Humanitarian Dialogue co-hosted a series of four dialogues between Chinese and European experts on the potential risks arising from the development of artificial intelligence technologies and their impact on international security. The dialogues sustained in-depth discussions and academic exchanges with a wide range of experts and scholars, including representatives from think tanks, international organizations, technology companies, and academic institutions.

The dialogue series aimed to:

- Deepen understanding of the potential security risks associated with AI applications, including risks inherent in AI-enabled weapons and decision-support systems (AI-DSS) particularly in relation to risks of unintended escalation towards or within conflicts;
- Identify possible confidence-building measures (CBMs) to reduce the risk of escalation associated with the use of AI-enabled military systems;
- Explore avenues for collaboration between China and Europe on AI governance.

The dialogue series began with discussions on categorizing AI-related international security risks and identifying shared concerns and gradually expanded to scenario-based case analyses. It has ultimately led to reflections and concrete proposals for confidence-building measures and governance frameworks. While participating experts acknowledged differences in priorities and perspectives on certain issues between China and Europe, there was consistent agreement on two core points: the urgency of addressing AI-related international security risks; and the importance of continued dialogue as a foundation for enhancing mutual understanding and exploring global AI governance approaches.

II. Key Issues

1. Shared Understanding of Security Risks related to AI

Participants identified a common set of risks associated with the use of AI, including:

- **Technical vulnerabilities:** Algorithmic bias, lack of transparency, vulnerability to adversarial attacks and data poisoning, all of which may undermine the reliability and safety of military AI systems.
- **Human factors:** Automation bias and overreliance on AI-generated recommendations, as well as the impact of AI on decision-making speed and decision-makers' judgment, potentially leading to misjudgment or errors.
- **Operational and Contextual risks:** Misinterpretation or malfunctioning of AI systems in real-world scenarios may lead to misjudgments, and disinformation and deepfake technologies could intensify information and cognitive warfare, further exacerbating regional and international tensions.
- **Strategic instability:** AI may challenge traditional frameworks of strategic stability, lower the threshold for the use of force, and deepen mistrust among states.

2. Scenario-Based Discussions

The dialogues featured detailed scenario analyses, including:

- **Scenario A (Maritime domain):** Focused on the risk of misidentification of civilian vessels during the use of military AI decision-support systems (AI-DSS). Key concerns included deficiencies in training data, biases, challenges of automated decision-making, and the absence of standardized communication and coordination protocols across agencies and countries in the case of incidents. Such failures could contribute to unnecessary escalation of maritime incidents and potentially escalate tensions.
- **Scenario B (Land/conflict zones):** Examined the ethical and operational risks of deploying AI-enabled weapon systems in conflict environments. Discussions emphasized the need to clarify accountability in the case of unintended consequences —whether related to AI system/algorithm, operators, and/or decision-making oversight—and to distinguish the roles of technical design standards and operational policies in mitigating risks.

3. Confidence-Building Measures (CBMs)

A central focus of the dialogues was exploring practical CBMs and discussing the full lifecycle management of AI to mitigate risks associated with AI applications and enhancing mutual trust and international cooperation between China, European states, and the broader community. Participants emphasized that the development of

AI-related CBMs should follow a phased and incremental approach:

- Prioritize addressing core technical risks, such as algorithmic bias and system reliability, while improving mutual understanding to reduce misperceptions stemming from differences in technical approaches or information asymmetry.
- Promote international cooperation, strengthen risk identification and prevention capabilities, and enhance information exchange and sharing to enable timely communication and response to incidents, including technical vulnerabilities and major security risks.

III. Key Challenges

Despite broad consensus on the need to address risks posed by AI and the importance of continued dialogue as foundations for constructing global AI governance frameworks, several challenges remain:

- **Transparency:** AI systems often function as "black boxes," lacking explainability in their decision-making processes. This undermines public trust and complicates regulation. Differences across countries in data access, algorithm disclosure, and regulatory standards further exacerbate information asymmetries. There is a need to advance research on explainability, promote information exchange, and narrow the gaps created by differing regulatory standards.
- **Decision-making speed:** While AI significantly enhances decision-making efficiency, high-speed automated processes may increase the risk of misjudgment or escalation. It is therefore necessary to appropriately slow down AI-assisted decision-making and ensure sufficient space for human judgment and intervention, especially in critical stages, to mitigate risks of misperception and escalation. In addition, time constraints may limit careful human evaluation in urgent or high-pressure situations, making it essential to balance efficiency with safety while maintaining appropriate human control and final decision authority.
- **Role of the private sector:** As private companies play an increasingly prominent role in AI development and application, traditional state-centric governance models face structural challenges. While corporate leadership accelerates technological innovation, it may also weaken direct state control over critical technologies and increase the complexity of cross-border governance and regulatory coordination. There is a need to clarify the division of responsibilities between governments and enterprises and to promote a multi-stakeholder collaborative governance framework.

IV. Recommendations for Future engagement

1. Sustain dialogue to enhance mutual understanding

Given existing understandings and remaining differences, it is recommended to continue a regularized expert dialogue mechanism. Periodic, high-level expert exchanges can deepen discussions on AI-related international security risks and governance pathways, reduce information asymmetry, mitigate misunderstandings, and gradually build consensus as a foundation for further cooperation.

2. Enhance policy relevance and global impact

As two major forces for global peace and stability, China and Europe play a vital role in international security governance. It is recommended that discussions and outcomes are closely aligned with policy needs, and that representatives from defense, foreign policy, and governance institutions be invited as observers. This would facilitate broader discussions on how different regions perceive AI-related international security risks, strengthen shared understanding, and improve the policy translation and international impact of the dialogue.

Introduction to the Artificial Intelligence Project at the Center for International Security and Strategy (CISS)

The Center for International Security and Strategy (CISS) of Tsinghua University was established on 7 November 2018. As a university-based think tank, it focuses on research within the fields of international strategy and security. Since 2019, CISS has focused on the cutting-edge developments in artificial intelligence technology and international security governance issues, establishing a dedicated expert group for its Artificial Intelligence Project to advance research on AI and international security.

Concurrently, through ongoing collaborations with the Brookings Institution in the United States and the Centre for Humanitarian Dialogue in Switzerland, CISS maintains track-two dialogues on AI and international security between China and the United States, as well as between China and Europe. CISS also continuously expands project cooperation on global AI governance with international organizations and think tanks such as the United Nations Institute for Disarmament Research (UNIDIR) and the International Committee of the Red Cross.

These efforts have yielded a series of significant research findings and policy reports, contributing to building international consensus for advancing cooperation in the field of AI and international security. Furthermore, CISS actively undertakes research projects commissioned by national ministries including the Ministry of Foreign Affairs, Ministry of Science and Technology, and Ministry of Finance, while deepening Area Studies on AI governance to contribute to policy formulation and international cooperation.

Introduction to the Henry Dunant Centre for Humanitarian Dialogue (HD)

The Henry Dunant Centre for Humanitarian Dialogue (HD), founded in 1999, is dedicated to preventing, mitigating, and resolving conflicts through dialogue. Headquartered in Geneva, Switzerland, HD is an international organization registered under Swiss federal law and an official partner of the United Nations. HD was awarded the Carnegie Wateler Peace Prize for its outstanding contribution to conflict mediation and the promotion of peace. Previous recipients of this award include the UNHCR, UNICEF, former United Nations Secretary-General Dag Hammarskjöld, the first Secretary General of the League of Nations Eric Drummond, and the former U.S. Secretary of State Henry Kissinger, among others.

HD works in a rigorous, pragmatic, and understated manner in accordance with the principles of humanity, impartiality, and independence, focusing on innovation and always working in a results-oriented manner. With more than 300 staff across Asia, Africa, Latin America, the Middle East, and Eurasia, HD has mediated in more than 80% of the world's violent conflicts. HD has worked in China since 2008 to support communication between China and relevant countries and organizations. HD's work aims to promote frank dialogue, exchange of views, mutual trust, and solutions towards peaceful resolution and effective prevention of conflicts and disputes.

清华大学战略与安全研究中心
CENTER FOR
INTERNATIONAL SECURITY AND STRATEGY
TSINGHUA UNIVERSITY

Centre for
Humanitarian
Dialogue

ciss@tsinghua.edu.cn
86-10-62771388
428A Mingli Building, Tsinghua University,
Haidian District, Beijing, China (100084)
http://ciss.tsinghua.edu.cn

WeChat Website Contact Us