

Vol.8 No.8(2019)

INTERNATIONAL SECURITY AND STRATEGY STUDIES REPORT

Uncertainties: Why Are We Concerned about the Impact of AI on International Security?



清华大学战略与安全研究中心

CENTER FOR
INTERNATIONAL SECURITY AND STRATEGY
TSINGHUA UNIVERSITY

Uncertainties: Why Are We Concerned about the Impact of AI on International Security?

CHEN Qi (Secretary-General and Professor of Center for International Security and Strategy, Tsinghua University)

ZHU Rongsheng (Post-doc of Center for International Security and Strategy, Tsinghua University)

Scientists generally hold that the development of artificial intelligence (AI) is still at the stage of “weak AI” or “ANI”, which means the technology is only capable of fulfilling a certain task or addressing a specific issue. Nevertheless, some experts assume that, like nuclear weapons, the militarization of AI is going to challenge the paradigms of international strategy. In spite of that, no one has the answer as to when it will evolve into a full-fledged revolution? As we can tell now, some AI enterprises are forced to decline extremely profitable government procurement orders on military equipment due to ethical concerns. Although AI systems are successfully developed, they require another several months or even years to receive different approvals to run the application.¹ As to the effects and approaches to the reform of weapon production, AI is essentially different from nuclear weapons, which demonstrated astonishing destructive power at the onset, and accordingly exerted far-reaching psychological and physical impact on international security. AI is an enabling technology, and only has the potential to trigger a war when combined with other weapon technology. A few pessimistic experts go further to predict that the existing theory is not comprehensive enough to support the development

1 *National Security Commission on Artificial Intelligence Interim Report*, November 11th 2019, <https://drive.google.com/file/d/153OrxnuGEjsUvIxWsFYausIwNeCEkvUb/view>.

of AI, which means AI might expect a “bitter winter” one day in the future. Now that the transformative impact of AI remains uncertain, why is the international community so urgent to find out solutions to AI governance on international security?

While in recent domestic and international meetings, policy makers, scientists, engineers and scholars of politics, among others, feel extremely worried and anxious about the potential impact of AI on international security. Their anxieties are not derived from human vigilance to long-term risks, but from the fact that various uncertainties of AI technology and its applications could have substantial impacts on international security and people’s dynamic perception of these uncertainties. This report primarily discusses the challenges that AI poses to international security, and holds that AI technology not only challenges international security but also promises the potential to stabilize international security. What we should do is to make a prudent and wise choice, not to feel fear and desperate arising from uncertainties. Whether AI will “end” human race or advance the development of human civilization as an “engine” depends on how we use it.

I. Impact of AI on International Security

AI, an increasingly important and fast-growing underlying technology, could enhance a country’s economic competitiveness and trigger a new industrial revolution through empowering many industries.² Previous industrial revolutions have not only changed the international landscape,

² For the development disparities brought about by AI and other enabling digital technologies of the economy, see “The Age of Digital Interdependence,” *Report of the UN Secretary-General’s High-level Panel on Digital Cooperation*, June 2019; “Digital economics How AI and robotics are changing our work and our lives,” *Deutsche Bank Research*, May 14th 2018.

but also led to power shifts and the rise and fall of great powers. Countries which develop and commercialize new technologies ahead of others in the market will secure enormous economic benefits. According to an estimation of the McKinsey Global Institute, netting out competition effects and transition costs, AI could potentially deliver additional economic output of around \$13 trillion by 2030, boosting global GDP by about 1.2 percent a year. If delivered, its impact on the global economy would compare well with that of other general-purpose technologies, including steam engines in the 1800s, industrial manufacturing in the 1900s and IT during the 2000s.³ However, the digital economic growth is not shared fairly. Developed economies with advantageous elements, including technology, capital, data and talents are more likely to secure greater economic benefits from the technology. The widespread application of reliable AI products will foster a more inclusive environment and boost public acceptance of the application of new technology. Economies with underdeveloped digital infrastructure, capital and technological foundation will find it increasingly difficult to catch up with AI leaders. Thus, the digital divide will only be widened.

There will be winners and losers in the AI race. Countries failing to develop AI technology and achieve commercialization could potentially suffer economic losses. Likewise, Countries that underinvest in AI R&D for military applications will put their national security at greater risk, and thereby diminish their own geopolitical influence.⁴ Many countries have released AI strategies since 2016, either positioning themselves as AI leaders or expressing strong intentions to play a key role amidst the trends. We can

3 “Notes From The AI Frontier Modeling The Impact of AI On The World Economy,” *McKinsey Global Institute*, September 2018.

4 Daniel Castro, Michael McLaughlin, Eline Chivot, “Who Is Winning the AI Race: China, the EU or the United States?,” *Center for Data Innovation*, August 2019.

tell from this that technological competition is anything but neutral. With the development of telecommunications technology, well-situated leading countries want to make sure that their power and influence are reinforced in the global digital transition or ensure that they do not fall behind. For this reason, the competition of securing technological superiority is set to cause a “tragedy of great power politics”.⁵

This assumption seems to be supported by the China-US trade war and competition for technology leadership. Chances are that policy makers overestimate the power and influence generated by technological breakthroughs but underestimate the potential role of global cooperation in boosting international security. Realistically, major powers dare not maintain peace of mind in the competition, which might change their fate, because they are not sure whether they could beat their rivals or where they will eventually land in the new revolution of a technological order, thereby triggering the China-US great power competition, which is splitting the world into two opposite systems. “The world (is) splitting in two, with the two largest economies (China and the US) on earth creating two separate and competing worlds, each with their own dominant currency, trade and financial rules, their own internet and artificial intelligence capacities, and their own zero sum geopolitical and military strategies.”⁶

The international order appears to head for confrontation, but it is not the only thing that disturbs the world. The ongoing technological advances are shaking the foundation of strategic deterrence, and the strategic

5 John Mearsheimer, *The tragedy of great power politics*, W.W.Norton & Company, 2001.

6 António Guterres, “Address to the 74th Session of the UN General Assembly,” *United Nations Secretary General Speech*, September 24th 2019, <https://www.un.org/sg/en/content/sg/speeches/2019-09-24/address-74th-general-assembly>.

stability and expectations of major powers. Nuclear-weapon states (NWS) are clearly going to deploy a powerful second-strike capability to ensure that their adversaries do not resort to the first strike at the risk of being destroyed. On that account, identifying the level of retaliatory credibility has become an important starting point in the nuclear strategic rivalry between major powers.⁷ Yet, instead of thoroughly destroying the foundation of strategic competition among major countries, all AI needs to do is to undermine the level of retaliatory credibility.

AI is now capable of helping detect missile silos from a massive dataset.⁸ The interest of the United States (US) in developing the capability to track and target mobile missile launchers makes Russia and China fear that its powerful reconnaissance capabilities could mature into a threat to their more sophisticated retaliatory forces.⁹ If unmanned weapons feature sufficient stealth and greater levels of autonomy, a nation will benefit from strike options with lower risks but higher combat effectiveness,¹⁰ which gives the attacker greater strategic advantages. This type of tactics cannot guarantee that a country could totally avoid a second strike, but the possibility per se can be extremely dangerous for nation. Nuclear deterrence pushes decision makers to make a decision in a very limited

7 Keir A Lieber and Daryl G Press, “The new era of counterforce: Technological change and the future of nuclear deterrence,” *International Security*, Vol. 41, No. 4, 2017, p. 9.

8 Richard A Marcum, et al, “Rapid broad area search and detection of Chinese surface-to-air missile sites using deep convolutional neural networks,” *Journal of Applied Remote Sensing*, Vol. 11, No. 4, Nov. 13th 2017.

9 Edward Geist, Andrew J. Lohn, “How Might Artificial Intelligence Affect the Risk of Nuclear War?” *Rand Corporation*, 2018.

10 Michael Mayer, “The New Killer Drones: Understanding the Strategic Implications of Next-Generation Unmanned Combat Aerial Vehicles,” *International Affairs*, Vol. 91, No. 4, July 2015, pp. 765-780.

amount of time, and thereby levies heavier pressure on them regarding the launch of the first strike. Or a country could seek to make up for its inadequate deterrence by developing more dangerous weapons. The arms race thus incurred will force it to make deployments of unsafe AI systems, thus aggravating strategic instability. There is another alternative available to the defensive side that is to attack the probing devices of its adversary, and generate an adversarial network or adopt strategic deception to guard against the other party's potential retaliatory capabilities. These efforts will more or less ensure its own nuclear retaliatory facilities are covert and safe. This aggravation of security dilemma may, however, intensify the complexity and misjudgment of strategic deployments.¹¹ Eventually, the country will either have to launch the attack in advance or lose the war altogether.

Apart from the threat to strategic stability, AI may also threaten the balance of power and exacerbate conflicts among major powers. In the era of algorithm warfare, data collection and algorithm training lead to "substantial" progress in military might.¹² The "substantial" progress will strongly influence the battlefield. Countries with technological advantages will be able to deploy more sophisticated weapons or develop new combat concepts, whereas those with backward technology will suffer from relative vulnerabilities due to the lack of countermeasures. From the security perspective of realism, a major power may seek comparative advantages or fear that its own security might be threatened by the military advantages

11 Jurgens Altmann and Frank Sauer, "Autonomous Weapons and Strategic Stability," *Survival*, Vol. 59, No. 5, 2017, pp. 121-127; Vincent Boulanin, "The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk", Vol. I "Euro-Atlantic Perspectives," *Stockholm International Peace Research Institute*, 2019.

12 James Johnson, "Artificial intelligence & future warfare: implications for international security," *Defense & Security Analysis*, Vol. 35, No. 2, 2019, p. 157.

of other countries, or seek alternatives to offset the advantages of its adversary. It would not only start a new arms race, but also increase mutual strategic distrust among different nations, and give rise to unexpected international conflicts. Imaging, a deep fake video showing US soldiers killed by Russia's poison gas has the potential to provoke both powers into nuclear deterrence.¹³ Along the spiral trajectory of an arms race, more and more advanced autonomous weapons will make warfare outpace human reaction, or countries will be scrambling to deploy unsafe AI weapons. Either way, decision makers will face overwhelming mental pressure, and their rational judgments of strategic response will be distorted.

In addition, the proliferation of autonomous weapons will lower the threshold of war and increase the risks of war between countries. Military actions always come with high risks of casualties, so state leaders have to think twice before launching military operations, taking people's sensitivity to casualties into account. However, the process of strategic decision-making will be changed by the expectation that autonomous weapons could reduce the casualty rate. Decision makers can persuade people that small expense can generate larger returns of war. Major countries will free themselves from domestic limitations on the use of force, which creates a more favorable environment for them to exert military influence other countries. However, the proliferation of autonomous weapons actually benefit medium countries built on a robust foundation of technological development the most, rather than major countries. It is because these middle powers can overcome their disadvantages of resources and population with the help of technology, and then change the allocation of conventional international force to enhance their position in the

¹³ Mark Fitzpatrick, "Artificial Intelligence and Nuclear Command and Control," *Survival*, Vol.61, No.3, 2019, pp.81-92.

international system's balance of power. Nowadays, lethal autonomous weapons systems(LAWS) are not effectively regulated globally, so people are concerned about how changes in the balance of power between major and medium countries will affect the international security system. Moreover, the vulnerability of AI technology increases the possibility of these emergencies.

II. Exploration of Safeguarding International Security Governance

Observers who strongly believe in technological advances are actually extremely pessimistic about human affairs, and regard AI as “the end of enlightenment”¹⁴, “the cause of the Third World War”¹⁵, “the end of the human race”¹⁶, etc. However, those observers who hold totally opposite opinions are much more optimistic about our ability to prevent disasters. As for now, AI is a continuously developing enabling technology, and it is hard to give it a perfect definition. The transformative effects brought about by AI have already spread to the whole world. Since it is impossible to prohibit it completely, and inconsistent with international concerns and shared future for mankind to let it run its course without check, here comes the question: how to conduct AI international governance? For decision makers, it is not only necessary for them to focus on uncertainties of international security brought about by AI, but also to continuously enrich

14 Henry Kissinger, “How the Enlightenment Ends,” *The Atlantic*, June 2018, <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>.

15 Elon Musk on AI: we are summoning the demon, October 30th 2017, http://tech.ifeng.com/a/20171030/44736042_0.shtml.

16 Stephen Hawking warns AI could spell the end of the human race, April 8th 2017, http://www.xinhuanet.com/tech/2017-04/28/c_1120889914.htm.

their knowledge about AI, which will change the options of existing and future governance. Therefore, what we are facing now is not a doomed tragedy, but a significant choice that could boost human development.

Under the UN framework, several mechanisms have had international discussions on how to limit the development of autonomous weapons. Among them, the UN Convention on Certain Conventional Weapons (CCW or CCWC) has held three informal meetings of experts and three formal meetings of governmental experts since 2014. Although certain breakthroughs have been made in the establishment of international mechanisms and regulations, parties involved hold hugely different opinions about a feasible definition of lethal autonomous weapons systems, and argue against the ambiguous objects in all definitions. No matter what definition is under discussion now, it now seems to be contrary to the existing rules of weapons. According to the statistics and analysis of 154 weapon systems by Stockholm International Peace Research Institute (SIPRI), only 49 of them are able to be applied to wars under human supervision but not human interference. Those weapon systems are used to defend our own facilities, for example, protecting warships or bases, or handling oncoming missiles and so on.¹⁷ Strategic dividends associated with LAWS influence the security interests of all countries, and impose more difficulties and challenges to subsequent arms control actions.

Even so, it does not mean that we are backed into a corner or we are bound to have endless and meaningless debates. At the first informal meeting of experts in 2014, parties involved were clearly divided. Some representatives thought the development and use of LAWS should be

¹⁷ Vincent Boulanin, and Maaïke Verbruggen, “Mapping the development of autonomy in weapon systems,” *SIPRI Report*, November 14th 2018, p.26.

completely prohibited, because they could be easily proliferate and extremely dangerous, and more importantly, they will impose huge threats to mankind. Conversely, others supported the development of LAWS, because they will be “smart” enough in future to understand the ethics of human war and even provide more humane options of war to people. These two different opinions agree that AI will be much stronger, but come up with opposite solutions. Till now, their projections have not materialized yet, but debates that seem to be irreconcilable have gradually produced consensus. Parties concerned agree that the arms control of LAWS should not stand in the way of the development of civil technologies and economic growth, and progress still should be made in other issues even with ambiguous definitions.¹⁸ This indicates that the cognition of decision makers is not unchangeable. Relevant negotiations of arms control are not to reach some binding international conventions at once, but to formulate the bottom line of LAWS and then combine multiple means of arms control to establish “soft law” progressively. This requires self-restriction in each country and non-binding agreements among countries, for instance, “codes of conduct” in line with the existing international law and norms.

With regard to the maintenance of global strategic stability, it is not a matter of life and death for nations yet, and AI can play a significant role in safeguarding strategic stability. The increased accuracy of intelligence gathering and analysis may result in more credible deterrence, assurance and reassurance. Ideally, reassurance to the opponent can be enhanced with more comprehensive intelligence and analysis. This process will promote

18 Report of the 2017 Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS), groups of governmental experts from contracting parties of Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, December 22th 2017.

a virtuous circle and ultimately lower the risks of war significantly. When nuclear-weapon states hardly trust each other or know each other's intention, more effective reconnaissance will provide decision makers with more reliable information.¹⁹ Even though AI is able to locate the missile silo of a country, the country can enhance the resilience of its silo against the first strike and thus conserve strategic power for the second attack which fosters strategic deterrence. In other words, from the perspective of technologies, strategic instability associated with AI is not unsolvable. In addition, AI can also supervise and review nuclear facilities in nuclear disarmament and denuclearization.

The "Great Fracture" is not the only choice for both of China and United States. We should be optimistic that AI competition is not a zero-sum game, but a tool to promote the welfare of human beings and global cooperation. Although the Trump administration's foreign policy to China shows signs of "decoupling", scientific research cooperation between these two countries instead witnesses more progress. According to the statistics of AI papers from Clarivate Analytic, from 2013 to 2017, international papers co-authored by scholars in China and the US boasted the highest growth rate with up to over 4,000 pieces, and they were the most active partner of each other in the past five years.²⁰ Studies based on the statistics of publications in the field of science and engineering show that papers co-

19 Edward Geist, Andrew J. Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?" *Rand Corporation*, 2018; Vincent Boulanin, "The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk", Volume I; "Euro-Atlantic Perspectives," *Stockholm International Peace Research Institute*, 2019.

20 Competitive Analysis of Proliferating Countries/Regions on AI Papers, Clarivate Analytics, December, 2018, p13; cited in Fu Ying, A Preliminary Analysis of AI's Influence on International Relations, *Quarterly Journal of International Politics*, No.1, 2019, p.16.

authored by scholars in China and the US grew by 55.7% during the period of 2014-2018, and conclude that US research article publications would have declined considerably without co-authorship with China.²¹ Scientific and technological advances and their exchanges in the world benefit all countries, and technological breakthroughs in a country may be conducive to the social development in another. For example, the Beijing-based Microsoft Research Asia (MSRA) once supported four Chinese young scholars to publish a paper about deep residual learning, which has become an important reference in recent years. After that, one of the four authors joined Facebook AI Research in the US, and the other three devoted themselves to AI startups in China.²² Their great success has undoubtedly benefited Chinese and American enterprises and AI R&D as a whole.

In respect of international security affairs, uncertainties are normal, whereas certainties are rare. As AI technology advances, the definition of LAWS will not be limited to technological considerations anymore, but focus on the autonomous level. For instance, the intelligent level of weapons will become the evaluation criteria of nature and limitations. What policy makers need to worry is not the loss of power resulted from uncertainties, but the long-term risks brought about by the crystallization of the cognition. They should not look to address the dilemma of international security through a robust and secure AI system enabled by the principle of “technology governing technology”. Safe military systems will lower the possibility of emergencies between countries at the technological level, but it means that a nation gets advantages that only it has, which

21 Jenny Lee and John Haupt, “Winners and losers in US-China scientific research collaborations,” *Higher Education*, November 7th, 2019.

22 Matt Sheehan, “Who Benefits From American AI Research in China?” October 21th 2019, <https://macropolo.org/china-ai-research-resnet/>.

will inevitably exacerbate the arms race and even lead to disastrous consequences by weakening strategic stability. Meanwhile, countries should not deploy unsafe AI systems in order to win the arms race. Policy makers should neither underestimate impacts of different strategic cultures on new technology, nor overestimate security impacts brought about by AI. Currently, AI technology remains immature, and there is not enough evidence to prove that it is able to replace existing weapon systems. Policy makers should recognize that AI security governance determined by uncertainties is a dynamic process, and any plans based on the best or the worst results which could potentially emerge in the future may be hard to achieve. Against the ever-changing international landscape, policy makers should seek international cooperation in developing technology and changing humanity.

Executive editor: Xu Xinyun

Center for International Security and Strategy of Tsinghua University

Address: Room 217, Mingzhai, Tsinghua University

Tel: +010-62771388

E-mail: ciss_thu@163.com