10. The shaping of strategic stability by artificial intelligence

CAI CUIHONG*

The world has already begun to enter the artificial intelligence (AI) era. AI and unmanned vehicles have been called the 'second nuclear weapon' with the potential to change the ways in which future wars will be fought.[1] China, Russia and the United States, among other powers, have been competing in AI development. The world is thus embarking upon, or perhaps could be said to have already started, a new cold war, this time driven by AI.

In the light of these developments, this essay considers whether AI will have a similarly profound impact on the strategic stability of the great powers. It begins (in section I) with a review of national AI strategies. It then describes (in section II) how the nuclear strategic stability of the cold war has developed into modern complex strategic stability. The essay then considers the conditions under which AI could have an impact on strategic stability (in section III) and what forms this impact could take (in section IV). It ends by considering (in section V) how AI needs to be included in any framework for maintaining strategic stability.

I. National AI strategies

In recent years the US Government has issued a series of documents on AI strategy.[2] Throughout these documents, the USA emphasizes the use of technological innovation to preserve US military advantage into the future—known as the Third Offset Strategy.[3] Moreover, these documents note that no other technology would have as much of an impact on US military operations as AI and intelligent technologies, whether used in remote sensing, command-and-control networks,

[1] '日媒称日本正加快引入"第二核武器" 紧追中美俄步伐' [Japanese media says that Japan is accelerating the introduction of the 'second nuclear weapon' and closely following the pace of China, the United States and Russia], 参考消息 [Reference News], 28 Jan. 2019.

[2] US National Science and Technology Council (NSTC), Networking and Information Technology Research and Development Subcommittee, *The National Artificial Intelligence Research and Development Strategic Plan* (White House: Washington, DC, Oct. 2016); US Executive Office of the President and National Science and Technology Council (NSTC) Committee on Technology, *Preparing for the Future of the Artificial Intelligence* (White House: Washington, DC, Oct. 2016); US Executive Office of the President, *Artificial Intelligence, Automation and the Economy* (White House: Washington, DC, Dec. 2016); US Department of Defense (DOD), *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity* (DOD: Washington, DC, Feb. 2019); and 'Maintaining American leadership in artificial intelligence', Executive Order no. 13 859, 11 Feb. 2019.

[3] Hagel, C., US Secretary of Defense, Keynote speech, Reagan National Defense Forum, 15 Nov. 2014.

* The views expressed in this essay are those of the author and do not necessarily reflect those of any organization to which she is affiliated. It was translated from Chinese to English by the volume editor, Lora Saalman.

operations or logistical support networks.[4] Reflecting these priorities, in June 2017 the US Government considered limiting China's investment in AI in the USA.[5] In doing so, in accordance with the characteristics and advantages of AI technology, the US military sought to take the lead in proposing a new operational concept of algorithmic warfare with machine learning and deep learning technology as its core.

While China's AI developments started late, they also developed rapidly. China has already become an important force leading global innovation and development of AI. In May 2016 a number of Chinese ministries and agencies established the 'Internet Plus' three-year AI action plan to guide AI technological innovation and industrial development.[6] In July 2017 the State Council issued the New Generation Artificial Intelligence Development Plan, which details medium- and long-term systematic deployment of China's AI development.[7] Use of AI has become a national development strategy and the Chinese Government has been increasing financial and policy support.

According to Russian President Vladimir Putin, AI is 'the future, not only for Russia, but for all humankind' and 'whoever becomes leader in this sphere will become the ruler of the world'.[8] So while the scale of Russia's AI industry and its overall development of AI have lagged behind China and the USA, its achievements in certain areas should not be discounted. The Russian military is currently applying AI to its equipment-renewal programme: a three-dimensional intelligent equipment system has gradually formed, encompassing unmanned ground vehicles, unmanned aerial vehicles (UAVs) and unmanned underwater vehicles (UUVs).[9]

Other technological powers have also joined the AI development race with their own scientific and technological strengths. The Japanese Government has proposed a plan for a super-smart society, the Society 5.0 strategy.[10] The British Government released a report on Artificial Intelligence: Opportunities and Implications for the Future of Decision Making in 2016 and an AI 'sector deal' between

[4] Liao, K. (廖凯), '透视美军抵消战略的变与不变' [The changing and unchanging perspective of the US Third Offset Strategy], 解放军报 [PLA Daily], 5 Sep. 2017, p. 7.

[5] Stewart, P., 'US weighs restricting Chinese investment in artificial intelligence', 14 June 2017, Reuters.

[6] Chinese National Development and Reform Commission, Ministry of Science and Technology, Ministry of Industry and Information Technology, and Central Cyberspace Affairs Office Commission, '"互联网+"人工智能三年行动实施方案' ['Internet Plus' artificial intelligence three-year action plan], 18 May 2016.

[7] Chinese State Council, '新一代人工智能发展规划' [New Generation Artificial Intelligence Development Plan], Order no. 35, 8 July 2017.

[8] '普京大帝谈AI：得人工智能者得天下' [Putin the Great discusses AI: getting AI means getting the world], 搜狐网 [Sohu.com], 4 Sep. 2017, (author translation); and '普京警告：发展AI最成功国家将统治全世界，未来是无人机的战争' [Putin warns: countries that are most successful at developing AI will rule the world, drone wars are the future], 搜狐网 [Sohu.com], 3 Sep. 2017; and '"Whoever leads in AI will rule the world": Putin to Russian children on Knowledge Day', RT, 1 Sep. 2017.

[9] Wang, H. (王慧婕), '发展人工智能已成全球之势' [Developing AI has become a global trend], 人民论坛 [People's Tribune], Jan. 2018, pp. 20–21. See also chapter 7 in this volume.

[10] Japanese Cabinet Office, 'Society 5.0', accessed 26 Apr. 2019.

the government and the British AI sector in 2018.[11] France has also striven to become a European leader in AI, with the government launching the country's national AI strategy in 2017 and publishing a vision of 'AI for Humanity' in 2018.[12] Germany's 'Industry 4.0' strategy includes machine perception, planning, policy and human–machine interaction among the key research directions of its AI development.[13]

II. From nuclear strategic stability to complex strategic stability

Strategic stability is a concept from the cold war era. Its general definition is primarily derived from a 1990 Soviet–US joint statement on non-proliferation and strategic stability.[14] According to this statement, 'strategic stability' may be understood as an equilibrium of strategic forces between the Soviet Union and the USA. In other words, the strategic relationship between the two major powers is such that neither side has the motivation to launch a first nuclear strike.[15] The concept of strategic stability born in the cold war period has two components: crisis stability and arms race stability. Its direct purpose was to use the structure of armaments to eliminate the possibility of a nuclear war between the two superpowers. This theory came to be the main foundation of Soviet and US nuclear strategy, guiding mutually assured destruction (MAD) and having an impact on the development of the two countries' strategic nuclear forces throughout the cold war. Although the concept of strategic stability encountered certain challenges in the post-cold war era, it remains the basis for influencing the balance of international strategic forces.

Since the end of the cold war, the Soviet–US bipolar structure that guided the international security environment has undergone tremendous changes. Many Chinese and foreign scholars quickly concluded that the concept of strategic stability was no longer applicable to the new international situation. However, the concept continues to develop. Strategic stability had been limited to a relationship in which there is a lack of opportunity or motivation to destroy all the nuclear forces of the opponent.[16] Russian experts tend to divide this into a narrow and

[11] Innovate UK, 'Artificial Intelligence 2020 National Strategy', Gov.uk blog, accessed 26 Apr. 2019; British Government, *Industrial Strategy: Artificial Intelligence Sector Deal* (Department for Business, Energy and Industrial Strategy: London, 2018); and Jin, D. (ed.), *Reconstructing Our Orders: Artificial Intelligence and Human Society* (Shanghai University Press/Springer: Shanghai/New York, 2018).

[12] French Government, '#FranceIA: the national artificial intelligence strategy is underway', 26 Jan. 2017; and Villani, C., *For a Meaningful Artificial Intelligence: Toward a French and European Strategy* (Conseil national du numérique: Paris, Mar. 2018).

[13] Wang (note 9).

[14] Soviet–United States Joint Statement on Future Negotiations on Nuclear and Space Arms and Further Enhancing Strategic Stability, Washington, DC, 1 June 1990.

[15] Wu, T. (吴艇), '从中美战略稳定性看太空武器化问题' [Examining space weaponization via Chinese–US strategic stability], Master's thesis, Fudan University, Apr. 2012, p. 16.

[16] Logan, J., *China's Space Programme: Options for US–China Cooperation*, Congressional Research Service (CRS) Report for Congress RS22777 (US Congress, CRS: Washington, DC, 29 Sep. 2008); and Colby, E. A. and Gerson, M. S. (eds), *Strategic Stability: Contending Interpretations* (US Army War College, Strategic Studies Institute: Carlisle, PA, 2013).

broad sense.[17] In the narrow sense, strategic stability refers to a state in which military strengths and the potentials of strategic forces are roughly equal and neither side seeks to change the military balance to acquire a sustained advantage. In the broad sense, strategic stability refers to the cumulative implementation by two countries or alliances of political, economic, military and other measures that make it impossible for either party to launch a military offensive. In other words, strategic stability may be narrowly characterized as the balance between major powers, in particular the balance of strength and capabilities of strategic weapons. More broadly, it may be defined as a condition in which global actors maintain self- and mutual restraint on a global scale, thereby engendering a relatively stable and balanced strategic situation within the international system.[18]

As noted in a joint statement issued by China and Russia in June 2016, the international community is accustomed to regarding 'strategic stability' as a purely military concept in the field of nuclear weapons. This does not reflect the broad and multifaceted nature of contemporary strategic issues. To achieve the goal of peace and security, strategic stability should be evaluated from a more comprehensive perspective.[19]

Of course, this kind of strategic stability does not mean that disagreements do not occur. However, these differences should not affect the development of overall relations. As such, it could be argued that nuclear strategic stability during the cold war period has developed into the complex strategic stability of today, which is a comprehensive strategic balance in which both the scope and the subject are diversified and intertwined. In transitioning from the narrow to the broad concept of strategic stability, there have been two important changes, as detailed below.

First, the scope of strategic stability has expanded from nuclear power relations via military and security relations to overall strategic relations. The core of maintaining strategic stability is the achievement of mutual deterrence. For this reason, the concept of cross-domain deterrence has begun to replace the concept of nuclear deterrence among decision makers. In recent years, the USA has been committed to creating a system of strategic deterrence that gives it a dominant global role. At the same time, it is also gradually adjusting this system of strategic deterrence at the cognitive and operational levels. At the cognitive level, the USA's greatest threat has transformed from nuclear terrorism to strategic competition and cross-domain threats. At the operational level, the means of cross-domain deterrence have been strengthened across various fields: to reshape the USA's absolute superiority in nuclear deterrence, to establish offensive and defensive conventional deterrence, and to improve its offensive emerging capabilities in

[17] Dvorkin, V., 'Preserving strategic stability amid US–Russian confrontation', Carnegie Moscow Center, Feb. 2019; Berls, R. E. and Ratz, L., *Rising Nuclear Dangers: Assessing the Risk of Nuclear Use in the Euro-Atlantic Region* (Nuclear Threat Initiative: Washington, DC, Oct. 2015); and Margoev, A., *Pursuing Enhanced Strategic Stability through Russia–US Dialogue* (PIR Center: Moscow, May 2019).

[18] Li, Z. (李喆), '"第二核时代"战略稳定性研究' [Study on strategic stability in the 'second nuclear age'], 江南社会学院学报 [Journal of Jiangnan Social University], vol. 17, no. 4 (Apr. 2015), pp. 32–36, p. 32.

[19] 中华人民共和国主席和俄罗斯联邦总统关于加强全球战略稳定的联合声明 [Joint Statement by the President of the People's Republic of China and the President of the Russian Federation on Strengthening Global Strategic Stability], Beijing, 25 June 2016, (author translation).

cyberspace and space.[20] This is being done with the aim of achieving complementary and flexible combinations of advantages among these various deterrents. Furthermore, advanced AI systems can provide deterrence against potential threats, just like the nuclear weapons of the cold war.

Second, the protagonists of strategic stability have expanded from the two major coalitions led by the USA and the USSR to include various global actors. During the cold war, the paramount figures in strategic stability were the two nuclear superpowers, the USA and the USSR, which gave strategic stability certain characteristics. Since the global power game at that time was highly concentrated on the two superpowers, it was difficult for any third-party forces to influence the power balance between the two camps. As a result, strategic stability equated with the dynamics between the two. During the long period that followed the end of the cold war, the focus of global strategic stability also remained the bilateral strategic stability between the two nuclear superpowers, Russia and the USA.

As the world enters the next nuclear era, however, the issue of strategic stability is no longer limited to strategic nuclear confrontation between two militaries. In the global nuclear power system, it is no longer just two nuclear superpowers that can influence and play a decisive role. Furthermore, countries with strategic nuclear power are no longer limited to the five defined as nuclear weapon states by the 1968 Non-Proliferation Treaty (NPT).[21] In fact, many conventional weapons can already replace some of the functions of nuclear weapons.[22] With the deepening of globalization, the nuclear environment is becoming more and more fractured. Within this complex environment, more actors can influence global strategic stability through such high-technology asymmetric means as AI.

III. The feasibility of AI having an impact on strategic stability

The impact of AI on strategic stability is conditional. It is based on three criteria: (*a*) the openness of the strategic stability environment, (*b*) instrumental rationalism in strategic stability thought and (*c*) the expansion of strategic stability factors.

The openness of the strategic stability environment

An important pathway for AI to have an impact on strategic stability among the great powers is the openness of the strategic stability environment. This condition depends on the overall international environment and is reflected in two aspects: changes in the distribution of power and the fragility of strategic stability.

[20] Luo, X. (罗曦), '美国构建全域制胜型战略威慑体系与中美战略稳定性' [US full-domain deterrence and its implications for Sino-US strategic stability], 外交评论 [Foreign Affairs Review], vol. 35, no. 170 (Mar. 2018), pp. 37–62.

[21] Treaty on the Non-Proliferation of Nuclear Weapons (Non-Proliferation Treaty, NPT), opened for signature 1 July 1968, entered into force 5 Mar. 1970.

[22] Li (note 18), p. 32.

Changes in the distribution of power

Openness of hegemony and great power status to incorporate more actors may stem from changes in the distribution of power among states. From the historical rise and fall of great powers, changes have been evident in their strength over time, such as the decline of ancient Rome and the British Empire. If the distribution of power among countries changes, emerging great powers will inevitably challenge the existing hegemonic order. The openness of great power status may be due to the loss of the dominant foundation on which the great powers have relied. For example, the advantages of the sea power era have been gradually surpassed and replaced by the convenience of land transport and air traffic.

The openness of great power status may also be due to the homogenization of technological superiority. Hegemonic powers gain advantage from innovation in fundamental production methods, distinguishing them from other countries.[23] However, this advantage will not last long. As technology spreads and other countries learn from those that have succeeded in the competition to survive in the international community, great powers will increasingly behave the same, the world will soon trend towards homogenization and hegemony will be weakened. For example, despite the efforts of the international community to control nuclear proliferation, the trend is for more states to acquire nuclear weapons. Due to the large temptation of nuclear capabilities, some countries are still eager to try to develop them.

The openness of great power status may also stem from the asymmetric effect of new forces. In the era of AI and cyber means, actors with weak conventional forces may use asymmetric approaches to provoke conflicts. Under the logic of cyberweapon and AI weapon asymmetry, strong powers would prefer defensive strategies, rather than launching attacks. This is because such countries are more dependent on high-technology networks and have higher anticipated losses in conflicts. Even if a weak country and a strong country show the same aggressiveness, an attack launched by a weak country should be more destructive. The inherent logic behind a weak country launching such an attack is to use the asymmetric effect to inflict greater damage.[24]

The fragility of strategic stability

Openness may also stem from the fragility of the strategic stability relationship among great powers. During the cold war, this fragility mainly arose from the balance of terror. Major nuclear powers believed that the use of nuclear force would lead to unacceptable retaliation, so they maintained a relationship of strategic stability primarily by ensuring the ability to use nuclear weapons to engage in counterattack. However, the current balance of nuclear terror has begun to be threatened, particularly following the withdrawal of the USA from the 1972 Anti-

[23] Liu, M. (刘鸣), '美国霸权实力何以能持久延续？' [How can US hegemonic power last forever?], 社会科学 [Journal of Social Sciences], vol. 29, no. 3 (Nov. 2007), pp. 43–53, p. 44.

[24] Liu, Y. (刘杨钺), '网络空间国际冲突与战略稳定性' [International conflict and strategic stability in cyberspace], 外交评论 [Foreign Affairs Review], vol. 33, no. 157 (Apr. 2016), pp. 106–29, p. 114.

Ballistic Missile Treaty (ABM Treaty) in 2002.[25] Unilateralism threatens strategic stability. With the destruction of this nuclear non-proliferation mechanism, the world has fallen into a multilateral nuclear security dilemma. However, mutual vulnerability in the nuclear field is not the only pillar that sustains strategic stability.

Currently, in addition to nuclear factors, strategic stability relations among the great powers are also characterized by interdependence. This encompasses increasing common interests, such as the joint response to international terrorism and the proliferation of nuclear weapons, as well as development of other advanced technologies, such as AI. Common challenges also include failed states, climate change and other threats that can jeopardize economic growth and prosperity.

With deepening economic and political interactions, the great powers find that interdependence on each other and the international system is constantly growing.[26] In this way, strategic stability relations among great powers can be maintained. This is not only because mutual vulnerability means that these states have the ability to cause unbearable damage to each other, but also because they need to achieve more important goals and to confront common challenges and threats. At the same time, while economic and political interdependence are among the cornerstones for the maintenance of strategic stability among great powers, events in the economic and political spheres may also induce instability. On the whole, common interests and interdependence contribute to the strategic stability of great powers, but this stability is fragile.

Instrumental rationalism in strategic stability thought

The second criterion in evaluating the role of AI in the strategic stability of great powers is based on the universal existence of instrumental rationalism in international relations. The realist thinking underlying instrumental rationalism believes in technology and power, typically emphasizing their use to directly achieve its purpose. Strategic stability at the highest level is the stability of will at the political level. However, under the prevailing role of instrumental rationalism, this cannot occur. Instrumental rationalism may create a dilemma, in that attention is often not paid to the effectiveness of the instrument. Instead, it is often dominated by an extreme panic about being overtaken by an adversary, thereby causing strategic instability.[27] The existence of instrumental rationalism in strategic stability thought has greatly enhanced the importance and emphasis

[25] Soviet–US Treaty on the Limitation of Anti-Ballistic Missile Systems (ABM Treaty), signed 26 May 1972, entered into force 3 Oct. 1972, not in force from 13 June 2002, *United Nations Treaty Series*, vol. 944 (1974), pp. 13–17.

[26] Finger, T. (托马斯•芬加) and Fan, J. (樊吉社), '中美关系中的战略稳定问题' [Strategic stability in Chinese–US relations], 外交评论 [Foreign Affairs Review], vol. 31, no. 138 (Jan. 2014), pp. 43–55, p. 44.

[27] Ge, T. (葛腾飞), '工具理性主义的困境与美国冷战决策模式的批判—<保罗•尼采:核时代美国国家安全战略的缔造者>评介' [The dilemma of instrumental rationalism and a critique of the US cold war decision-making model—a review of Paul Nietzsche: the founder of the US National Security Strategy in the nuclear age], 美国研究 [Chinese Journal of American Studies], no. 3, 2018, pp. 135–44, p. 139.

placed by great powers on AI among the most advanced technologies, thus enhancing its role in maintaining strategic stability among countries.

There are three reasons for the proliferation of instrumental rationalism.

The first is cold war mentality. Instrumental rationalism first arose from the fact that this cold war construct has not been overcome. In fact, strategic stability is a legacy of this manner of thinking. As just one example, the National Security Strategy of the US administration of President Donald J. Trump, issued in December 2017, positioned China as a strategic competitor.[28] A cold war mentality has caused great power competition to replace the terrorist threat as a new strategic concern for the USA. Trump believes that the world has entered a new era of competition, such that the military strength, economic strength and political competitiveness of a country are of paramount international importance. In January 2018, the US Department of Defense (DOD), in a summary of the US National Defence Strategy, unabashedly demonstrated that the USA wants to continue to use various means, including AI, to maintain its absolute military superiority and to ready itself for long-term strategic competition among major powers.[29]

Second, instrumental rationalism also stems from the lack of strategic mutual trust among great powers. It could be argued that the current comprehensive strategic stability among major powers must still be based on strategic stability in the traditional military field. While great powers, such as China and the USA, may have good intentions and are working hard to maintain their bilateral relations, ensuring a lack of conflict and confrontation among great powers cannot rely solely on the will and intent of the countries concerned. With the current widespread lack of mutual trust among major powers, their intentions are often difficult to clarify and almost impossible to verify.[30] The relative balance in military power is the key to ensuring that there is no conflict or confrontation. Therefore, instrumental rationalists believe that, even in times of peace, they must maintain stronger military power and strategic strength to ensure that potential attackers can be blocked at any time.

The third is fatalistic realism, from which instrumental rationality also derives. John Mearsheimer sums up the tendency for there to be conflict between a rising power and an established power as a tragedy of great power politics.[31] In Chinese–US relations, fatalistic realism maintains that China's rise will inevitably challenge the dominant position of the USA and will lead to the two countries fighting for hegemony. Belief in unavoidable conflict will inevitably shape each other's cognition and behaviour and poses one of the most serious threats to

[28] White House, *National Security Strategy of the United States of America* (White House: Washington, DC, Dec. 2017).

[29] US Department of Defense (DOD), *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge* (DOD: Washington, DC, Jan. 2018). The full strategy is classified.

[30] Da, W. (达巍) and Zhang, Z. (张昭曦), '中美关系新阶段中的战略"失语"与战略稳定探索' [Strategic "aphasia" and strategic stability in a new stage of Chinese–US relations], 国际安全研究 [Journal of International Security Studies], no. 5, 2016, pp. 39–59, p. 57.

[31] Mearsheimer, J., *The Tragedy of Great Power Politics* (W. W. Norton & Co.: New York, 2014).

the strategic stability relationship between the two. Many Chinese experts instinctively regard any action taken by the USA that may have a negative impact on China as 'blocking' (遏阻) or 'containment' (围堵). Similarly, US scholars, media and politicians often claim that China's military modernization and activities around the world have a real but unspoken intention to challenge the dominant position of the USA.[32] If both sides believe that conflict is inevitable, the attitudes and policy actions of both countries will be affected. As a result, fatalistic realism may eventually erode all the pillars that maintain the strategic stability of great powers and result in a self-fulfilling prophecy.

The expansion of strategic stability factors

Another criterion for AI to influence strategic stability is through the expansion of strategic stability factors in the new era. Nuclear weapons are no longer the only consideration. To limit strategic stability to the field of strategic nuclear weapons does not guarantee comprehensive and effective security for a country. Nuclear weapons only defend a country's core security interests: ensuring that the country's central territory will not face a large-scale attack from foreign enemies. They will not provide effective support for a country's non-core interests.[33] For any great power, in addition to defending the core interests of the country's central territory, there are many other national interests. To effectively protect these interests requires a greater scope of stability that includes conventional military forces.

Moreover, the factors that influence strategic stability are not limited to the development of strategic military forces: they also cover new threats and instabilities. In other words, strategic stability has become an issue with multiple drivers. Factors such as unilateralism, nuclear proliferation, nuclear terrorism and the development of conventional weapons are evolving as new intervening variables that affect strategic stability.[34] Additionally, AI, cybersecurity, regional conflicts, energy issues, political and diplomatic influence, economic dependence, the level of scientific, technological and economic development, and the extent of participation in international affairs are all considerations for evaluating strategic stability among great powers.

The above-mentioned elements of strategic stability can be divided into three categories: technical factors, behavioural factors and institutional factors.[35] In other words, strategic stability is not only related to a country's deterrence under specific attack and defence patterns, but also to its behaviour and related mechanisms or systems. Technical factors establish the material basis for the comparison of strategic strength among countries. They not only determine the

[32] Finger and Fan (note 26), p. 48.

[33] Bo, E. (波尔特), '战略稳定概念对美国安全战略的影响及启示' [The impact and implications of the concept of strategic stability on US security strategy], 国际论坛 [International Forum], No. 5, 2016, p. 48.

[34] Li, D. (李德顺), '战略稳定性中的相互依赖因素' [The elements of interdependence in strategic stability], Doctoral thesis, Tsinghua University, May 2012, p. 19.

[35] Yu, Q. (俞倩倩), '从战略稳定性看反卫星武器的发展' [A look at the development of ASATs from the perspective of strategic stability], Master's thesis, Fudan University, 2008, pp. 17–18.

size of nuclear weapon forces, but also the level of military technology modernization and conventional forces. They are the fundamental factors in determining strategic stability. Behavioural factors are catalysts, guiding the ability to amplify or reduce material power. Institutional factors are the result of the behavioural interaction of states with one another. They can subtly change the actions of the state, establish a new norm of weapon technology development, and then reconstitute and shift technical and behavioural factors.[36]

As one of the most cutting-edge technologies in the technical factor category, AI plays an important role in all aspects of strategic stability. This is not only because it can affect traditional nuclear relations, conventional force comparison and so on, but also because it is a new variable with an impact on strategic stability. Following the cold war, conventional power advantage clearly shifted to the West. As a result, strategic stability guaranteed by the mutual deterrence of nuclear weapons became, in essence, the last pillar to maintain the balance of international military power. However, AI and cyber means offer an opportunity for a number of countries to garner an advantage. Therefore, as the maturity of AI increases, strategic stability is shaped by the extent of AI factors among technical elements.

IV. The ways in which AI could shape the future path of strategic stability

The core competencies of AI technology driven by deep learning algorithms include cognition, prediction, decision-making and integrated solutions.[37] Cognition refers to the perception and description of the world through the collection and interpretation of information, including such techniques as natural language processing, computer vision and audio processing. Prediction is based on obtaining a wide range of information, analysing different scenarios that may occur through multilayered neural networks, and predicting behaviours and outcomes that may occur in various scenarios in advance. Decision-making is comprised of effective analysis of collected information and completion of predictions regarding specific scenarios, to determine a course of action based on pre-set goals. Once AI is combined with other complementary technologies, it provides an integrated solution for extremely complex activities.

While the fundamental role of AI occurs via these four core competencies, the path of AI's impact on strategic stability can be subdivided into five aspects: (*a*) its empowerment effect on nuclear weapons, (*b*) its enhancement effect on conventional military forces, (*c*) its comprehensive penetrative effect on strategic capabilities, (*d*) its behavioural risk effect that leads to conflict escalation, and (*e*) its psychological anxiety effects.

[36] Li (note 34), p. 19.

[37] Feng, S. (封帅) and Zhou, Y. (周亦奇), '人工智能时代国家战略行为的模式变迁——走向数据与算法的竞争' [The pattern of change in national strategic behaviour in the age of artificial intelligence: towards competition between data and algorithms], 国际展望 [Global Review], no. 4, 2018, pp. 40–41.

Table 10.1. The empowerment effect of AI on nuclear weapons

AI application	Possible result	Impact on strategic stability	
Surveillance, target acquisition and reconnaissance	Higher or perceived higher risk of decapitating strike from an adversary by conventional weapons; higher mutual confidence due to increased transparency	✓ ✓	× ×
Early warning	Possible lower risk of accidental or misinformed launch of nuclear weapons	✓ ✓	×
Air and space defence and ballistic missile defence	Lower confidence in the survivability of second-strike retaliatory capability		×
Nuclear strike capabilities	Possible higher risk of accidental or unauthorized use of nuclear weapons; higher escalation risk		× ×
Command and control	AI as a trusted adviser; possible lower and higher risk, due to hacking or accidental or misinformed launch of nuclear weapons	✓	×
Protection systems for nuclear forces	Attack on nuclear forces or nuclear command and control by conventional weapon systems; higher risk or perceived risk of decapitating strike by an adversary		×

× = negative effect; ✓ = positive effect; AI = artificial intelligence.

Source: Derived from presentations by Nishida Michiru and Petr Topychkanov and subsequent discussion at the East Asia Workshop: The Impact of Machine Learning and Autonomy on Nuclear Risk, Beijing, 6–7 Sep. 2018.

The empowerment effect of AI on nuclear weapons

One of the ways in which AI plays a role in strategic stability is through its empowerment effect on nuclear weapons. Applications of AI that can empower nuclear weapons include in environmental detection, target location, early warning, air and space missile defence systems, nuclear weapon command systems, and protective systems for nuclear storage and transportation equipment.

Nearly all of the resulting scenarios may have an effect on nuclear strategic stability—positive or negative (see table 10.1). Nuclear experts and AI researchers seem to agree that advanced AI may seriously undermine the stability of nuclear strategy and increase the risk of nuclear war.[38] However, not all agree on how and why AI would have an impact. Indeed, AI has a double-edged impact on nuclear strategic stability.

The use of AI in two scenarios—in tracking missiles and as a decision aid on the use of nuclear weapons—illustrates the role that AI may play in nuclear warfare from both sides.[39] If AI is applied in tracking missiles, it will greatly improve the accuracy of monitoring potential enemy attacks. This increased transparency may enhance the strategic mutual trust between two parties, thereby reducing

[38] Geist, E. and Lohn, A. J., *How Might Artificial Intelligence Affect the Risk of Nuclear War?* (RAND Corporation: Santa Monica, CA, 2018). This report is based on a series of 3 workshops in May–June 2017.

[39] Geist and Lohn (note 38).

the possibility of nuclear war and improving strategic stability. However, in the event of a crisis, using or providing AI-enabled intelligence, surveillance and reconnaissance (ISR) may also increase tensions and the possibility of unexpected escalation of the conflict. Moreover, if the AI missile-tracking function is flawed or hacked, the probability of triggering a nuclear war will greatly increase, thereby reducing strategic stability. Accordingly, AI decision-making support has a dual impact on nuclear war.

With the use of AI, the number of factors that have an impact in the fragile MAD-based nuclear balance will significantly increase. AI-enabled autonomy and sensor integration are of strategic importance since they can enhance ISR, automatic target recognition (ATR) and terminal guidance capabilities, which may seriously weaken nuclear force survivability. This would thereby shake a country's sense of security and undermine crisis stability. This has a greater impact on China and Russia, since they primarily rely on mobile intercontinental ballistic missiles (ICBMs) for deterrence. Of course, the ability to develop ATR, sensor integration and signal processing remains extremely difficult. According to one report, in an increasingly multipolar strategic environment, AI is likely to lead to the breakdown of the balance of nuclear weapons and the failure of current means of nuclear deterrence by 2040.[40]

Furthermore, involvement of AI technology will also introduce new variables into the stability of the global system of nuclear deterrence. In an era of weak AI, only a combination of AI technology and nuclear weapons can form an effective deterrent system. When AI technology is involved in all aspects of a nuclear deterrent structure, the original system of stability will change. As a data tool, AI provides countries with new offensive capabilities and has a direct impact on the reliability of nuclear weapon use. In a big data environment, however, there are also a number of subjective factors, such as the unpredictability of national will and strategic intent. When subjected to deep learning algorithms, intent may be clarified. These shifts could result in an imbalance in the MAD-based system of nuclear deterrence. The party with a command of AI technology will have the ability to clearly assess the possibility and destructiveness of the other party's nuclear counterattack, thus having more flexible strategic options, while the side with relatively backward technologies will possess less credible retaliatory capabilities.[41] The gap between military powers will again expand and countries' military strategic aims will accordingly readjust. In other words, the traditional international security system will become unstable.

The enhancement effect of AI on conventional military forces

The second path for AI to have an impact on strategic stability is through its upgrading of conventional military forces. While nuclear weapons were the most important pillar of strategic stability during the cold war, they were not an

[40] Geist and Lohn (note 38).

[41] Feng, S. (封帅), '人工智能时代的国际关系: 走向变革且不平等的世界' [International relations in the AI age: towards a world of change and inequality], 外交评论 [Foreign Affairs Review], no. 1, 2018, pp. 140–41.

Table 10.2. The enhancement effect of AI on conventional military forces

AI application	Possible result	Impact on strategic stability	
Target country and battlefield situational awareness	AI has the ability to collect battlefield information more comprehensively and efficiently. The use of natural language processing systems can more efficiently collect and process audio signals. Machine vision can enhance the ability of automatic weapon systems to identify and analyse battlefield conditions. This allows for increased transparency, strategic mutual trust, reduced motivation to launch war. However, false information may also increase risk perceptions.	✓ ✓	×
Military command human–machine cooperative decision-making	An intelligent command system with functions of reasoning, analysis, prediction, decision-making, etc., can greatly improve the accuracy and effectiveness of military command activities. AI can quickly process battlefield information and has the rapid response capability that humans lack. AI offers multithread processing capability, can simultaneously handle multiple military operations and can propose complex strategies that are beyond the capabilities of human thought.	✓	× ×
Assisting human activity	This includes portable electronic equipment and auxiliary power units to ensure that the soldiers get help in a variety of possible emergencies. This could strengthen existing power distribution among states and at the same time reduce the fear of activities of war.	✓	× ×
Collaborative operations (advanced manned or unmanned combat teaming)	This consists of using AI systems to coordinate actions, optimize operational strategies, and flexibly adjust to battlefield conditions and operational objectives to maximize battlefield advantage. At the same time, it will increase asymmetry.	✓	× ×
Network empowerment and autonomous high-speed weapons for cyberattacks and electronic warfare	This covers everything from real-time identification of defects and vulnerabilities by computer systems that completely lack human intervention, to the ability to quickly and automatically complete software repair and system defence in billions of lines of code, to creation of a hacker robot with both offensive and defensive capabilities. Because of non-lethality, use may increase. Developed and intermediate countries may be the biggest beneficiaries of empowerment with autonomous weapons.	✓	× ×
Lethal autonomous weapon systems	This features self-discovery of targets, self-determination and implementation of attacks. It is relatively controllable among rational state actors, but it is uncontrollable in the case of non-state actors such as terrorist organizations.		× ×

× = negative effect; ✓ = positive effect; AI = artificial intelligence.

operational option in the great power competition. This was due to the balance of nuclear terror and the consequences of mutual destruction. With the expansion of the concept of strategic stability, conventional military forces have become an important consideration. The world's military has transformed from an era of mechanization to one of information. Algorithm-based AI is an important promoter of this military revolution. It is expected to give birth to new combat styles and to change the mechanism of winning wars. In doing so, it has become an important means to change the rules of the game in warfare and to shape subversive military capabilities (see table 10.2).[42]

AI can also play a broad role in non-nuclear forces. For example, the proliferation of autonomous weapons is not limited to such traditional fields as UAVs, but rather may be fully rolled out in a variety of military fields. One scholar has warned that 'If autonomous weapons are developed and deployed, they will eventually find a home in every domain—air, space, sea, land, and cyber'.[43] Unlike previous technological changes, AI technology in the military field has led to changes in all aspects: from military weapons to strategic design and from global military power balance to military ethics, all will inevitably be affected.

In terms of environmental situational awareness on the battlefield, AI has the ability to collect more comprehensive battlefield information. For example, the use of machine vision can enhance the ability of the automatic weapon system to identify and analyse battlefield conditions. Moreover, the natural language processing system can efficiently collect and process audio signals. For the strategic environment of competitors in peacetime, AI is also able to employ big data for statistical analysis to sense changes in strategic posture in a timely manner. In terms of military command, an intelligent command system with functions of reasoning, analysis, prediction and decision-making, among other capabilities, can greatly improve the accuracy and effectiveness of military command formulation. Combat commanders are thereby able to grasp battlefield information and to gain more precise tactical advice.

In practice, before a conflict begins, the AI system would be able to provide a more comprehensive set of battlefield information, simulate the deployment and combat capabilities of both sides, complete a relatively accurate format of the battlefield from deductive simulations and quantify all potential outcomes from a range of probabilities derived from various military strategies. In line with this quantitative probability, an effective operational plan of force distribution and strategic deployment could be selected and carried out. This is because AI has two advantages that humans are unable to match. First, AI systems can exceed human capacity in quickly processing battlefield information and engaging in rapid response. Second, AI systems have multithreading processing capabilities that can undertake multiple military operations simultaneously and propose

[42] Long, K. (龙坤) and Zhu, Q. (朱启超), '"算法战争"的概念、特点与影响' [The concept, features and impact of 'algorithmic warfare'], 国防科技 [National Defense Science & Technology], vol. 38, no. 6 (2017), p. 39.

[43] Roff, H., 'To ban or regulate autonomous weapons—a US response', *Bulletin of the Atomic Scientists*, vol. 72, no. 2 (Mar. 2016), pp. 122–24; and Roff, H., 'Banning and regulating autonomous weapons', *Bulletin of the Atomic Scientists*, 24 Nov. 2015.

complex strategies that human thought patterns are unable to grasp.[44] AI can also help humans with complementary actions, such as portable electronic equipment and auxiliary power units, to help in a variety of possible emergencies.

Humans can coordinate operations with AI systems and optimize warfare tactics, while flexibly adjusting to battlefield conditions and combat objectives to maximize battlefield advantage. Automated technology allows a weapon system to achieve greater flexibility and self-determination to solve problems. An intelligent weapon system not only achieves a substantial separation between human and weapon, but also completely transforms the war activity into a task of the weapon system. This brings the casualty rate among combatants to near zero and maximizes the efficiency of weapon use and coordination among various weapon systems. More importantly, the use of intelligent weapons makes the traditional combat laws, such as killing enemy combatants, lose their practical significance.[45] At the same time, human–machine collaboration can also accomplish a good deal of the work that cannot be done by humans alone. The USA and Europe have made breakthroughs on a number of key technologies such as UAV synergistic flight, unmanned vessel bee colony combat, unmanned submersible network detection and manned or unmanned combat aircraft formation flight tests.

Beyond these capabilities, network empowerment and autonomous high-speed weapons for cyberattack and electronic warfare are areas in which AI is particularly promising. Cyberweapons must operate outside communication range and respond rapidly. As a result, attacks initiated and controlled by AI systems have great potential. Further, the non-lethal nature of cyberweapons may increase their use. This being said, development of autonomous cyberweapons may differ from traditional weapons in that requirements at the technical level are higher. In this case, it could be argued that a technologically developed medium-sized country may be the largest beneficiary of autonomous weapons and may rewrite conventional power distribution, thereby injecting more uncertainty and instability into the international system.[46]

Additionally, countries are also vigorously developing lethal autonomous weapon systems (LAWS) that can independently identify targets, make independent judgments and carry out attacks. These types of system have the ability to engage in automatic attack and may engage in inhumane killing. For a national actor with a rational decision-making model, such systems are relatively controllable. What truly affects strategic stability and the international system is the use of LAWS by non-state actors, such as irrational terrorist organizations. This is because the rapid pursuit of new advanced technologies has not only enabled great powers to develop and deploy new weapon systems for a revolution in military affairs, it has also provided new possibilities for the proliferation of weapons of mass destruction and LAWS.

[44] Feng (note 41), p. 140.

[45] Feng (note 41), p. 139.

[46] Work, R. O. and Brimley, S. *20YY: Preparing for War in the Robotic Age* (Center for a New American Security: Washington, DC, Jan. 2014), p. 33.

The enhancement effect of AI on conventional military forces also results in another two forms of change in strategic stability among great powers. Due to this upgrade, technologically advanced countries may encounter lower risks, combined with more effective attack tools, such that they are able to pose a serious challenge to their opponent's strategic deterrence. Thus, for countries that have historically had the ability to fend off an attack, the introduction of greater mobility, concealment and autonomy capabilities with the next generation of equipment may make their retaliation-based deterrence strategy ineffective.[47] The impact of AI technology will thereby aggravate the imbalance of conventional military power confrontation. Armed forces lacking AI technology will find it increasingly difficult to compensate for their disadvantages on the battlefield through tactics and strategies. Conventional confrontation will no longer be a rational strategic option and they will have to resort to asymmetric warfare.[48] At the same time, the development of new unmanned weapons may also change the traditional casualty counts of conflicts, thus increasing the rate of use of these weapons. These trends are undoubtedly not helpful for great power strategic stability. However, the strategic mutual trust generated by AI-enabled mutual battlefield situational awareness and attack capabilities will also increase, which will be beneficial to a certain extent.

The comprehensive penetrative effect of AI on strategic capabilities

The third way for AI to have an impact on strategic stability is through its full penetrative effect on strategic capabilities. From the vantage point of international politics, the most important value of AI lies in a potential shift in allocation of strategic capacity among countries.[49] Competition in science and technology is an important part of strategic jockeying among great powers and competition in the field of AI is a core element. Therefore, the speed and impact of promoting the application of AI in various fields will not only profoundly affect future victory in war, but also the strategic competitiveness of great powers (see table 10.3). In a broad sense, the strategic competitiveness of great powers is ultimately the foundation of strategic stability in peacetime.

The comprehensive penetrative effect of AI on strategic capabilities is mainly due to its high penetrative advantage. AI has become an irresistible technological trend and is entering all aspects of life and all social fields. From the perspective of technological development, the new generation of AI not only represents a new direction in science and technology but also has an extremely important impact on the path of research and development (R&D) tools, costs and even the paradigm of how R&D is conducted in other scientific fields. From an economic vantage point, a new generation of AI will reconstruct all aspects of economic activities,

[47] Liu, Y. (刘杨钺), '全球安全治理视域下的自主武器军备控制' [Arms control of autonomous weapons under global security governance], 国际安全研究 [Journal of International Security Studies], no. 2, 2018, pp. 49–71, p. 64.

[48] Feng (note 41), p. 140.

[49] Liu (note 47), p. 50.

Table 10.3. The comprehensive penetrative effect of AI on strategic stability

AI application	Possible result	Effect on strategic stability	Impact on strategic stability	
Economy	Leads to major changes in economic structure, promotion and upgrade of industrial transformation, and achievement of a new leap in productivity	Winner-takes-all (×) AI technology catch-up cycle shortened (✓)	✓	× ×
Society	Greatly improves the level of targeted public services and comprehensively improves the quality of people's lives		✓	× ×
Politics	Increases political governance of the country and enhances freedom of speech		✓	× ×
Security	Increases the maintenance of national security measures and enhances national competitiveness		✓	× ×

× = negative effect; ✓ = positive effect; AI = artificial intelligence.

such as production, distribution, exchange and consumption. It will also form new macro- and micro-level intelligent demands and promote the advancement of new technologies, new products and new industries. These major structural changes will promote industrial transformation and upgrade, achieving a new leap in productivity.

Within social development, a new generation of AI will bring new opportunities for social construction. The extensive application of AI in education, medical care, elderly care, environmental protection, urban operation and judicial services will greatly improve the level of targeted public service to comprehensively improve the quality of people's lives. In terms of global competition, AI has become a new focus of international competition. Major developed countries regard the development of AI as a major strategy to enhance national competitiveness and safeguard national security. Whoever takes the lead in achieving breakthroughs in the field of AI will dominate future development.

Further, AI is a strategic technology that affects a country's developmental destiny and is related to the comprehensive strength of the country. In order to seize the initiative within this technological competition, countries have made plans for national AI strategies. At present, the world's major technological powers—China, Russia and the USA among others—all attach great importance to AI development. A 2018 report that systematically examined the possible impact of AI on national security from an economic, information, and military perspective recommends that the USA pay special attention to controlling the potential catastrophic risk of AI being used by hostile countries or through unanticipated incidents.[50] A parallel

[50] Horowitz, M. C. et al., *Artificial Intelligence and International Security* (Center for a New American Security: Washington, DC, July 2018).

report suggests that the USA should introduce an overarching national AI strategy as soon as possible.[51] This would be to guarantee that the USA is the global leader in top-level design, overall planning and key investment in AI technology, to allow it to win strategic competition in the AI field against China, India, Russia and South Korea, among others.

The strengths of cutting-edge technologies, enabled by AI and big data, may contribute to the formation of a new strategic balance. On the one hand, this is because the fourth industrial revolution centred on AI may lead to a winner-takes-all situation among countries. The comprehensive penetrative effect of AI on strategic capabilities is not conducive to the strategic balance of major powers. This is an important reason why countries hope to seize the opportunity in this unstable state. On the other hand, the shortening of the AI technology cycle of catching up is favourable for the strategic balance among major powers. In the previous industrial revolutions, the time advantage of the leading country over those working to catch up was large. For example, when the United Kingdom launched the First Opium War in 1839, China was still an agricultural society. The technology gap between the two countries could have been measured in decades, if not centuries. However, in the era of intelligent revolution, developed countries have realized such achievements as smartphones, driverless cars and cashless payment. As a result, developing countries have the chance to make similar progress within a year or two of these advances. As a result, the time differential is becoming smaller and smaller. This is also conducive to the formation of a multipolar world and the improvement of strategic stability.

The behavioural risk effect of AI that leads to conflict escalation

The fourth pathway for AI's impact on strategic stability is through its shaping of behavioural risk that could contribute to conflict escalation. It can do this in three ways (see table 10.4).

First, AI may blur the boundaries between conventional and nuclear warfare, thereby causing conflict escalation. As Paul Bracken of Yale University, USA, points out, the continued improvement of technologies such as AI has the potential to weaken the strategy of minimum nuclear deterrence and to blur the boundaries between conventional and nuclear warfare.[52] AI technology can help achieve new breakthroughs in tracking, targeting and anti-submarine warfare or make it easier for high-precision conventional ammunition to destroy reinforced ICBM silos.[53] This ability to destabilize is particularly significant because policy-makers are more likely to threaten to use conventional weapons than to conduct any form of nuclear attack. In a crisis, the threat of conventional weapon use can put tremendous pressure on the opponent. Doing so may force the country to yield

[51] Horowitz, M. C. et al., *Strategic Competition in an Era of Artificial Intelligence* (Center for a New American Security: Washington, DC, July 2018).

[52] Bracken, P., 'The intersection of cyber and nuclear war', Strategy Bridge, 17 Jan. 2017.

[53] Holmes, J., 'Sea changes: the future of nuclear deterrence', *Bulletin of the Atomic Scientists*, vol. 72, no. 4 (July 2016), pp. 228–33.

Table 10.4. The behavioural risk effect of AI that leads to conflict escalation

AI application	Behavioural risk of AI application	Main impact	Impact on strategic stability
Blur the boundaries between conventional and nuclear war	The opponent believes that it is necessary to use nuclear weapons before being disarmed or to counter an attack that fails to engage in successful decapitation.	Causes conflict escalation	× ×
Increase armed behaviour options	AI applications such as autonomous weapons do not necessarily involve human casualties and can alleviate the pressure of domestic public opinion that a country may face when launching and participating in foreign military operations.		× ×
The intention behind using AI to perform tasks may be misunderstood	This may be interpreted as a serious provocation against a target country's security interests, leading to more stringent response measures. Hacking may lead to misjudgement or an escalatory response.		× ×

× = negative effect; ✓ = positive effect; AI = artificial intelligence.

but could also trigger a nuclear war. The reasons for conflict escalation are that the opponent believes that it is necessary to use nuclear weapons (*a*) before being disarmed, (*b*) to counter a partially successful attack, or (*c*) in the event of a crisis that leads to accidental use.

Second, AI may increase the options for armed behaviour and cause conflict escalation. For national actors, one of the advantages of AI applications such as autonomous weapons is that they do not necessarily involve human casualties (on the attacking side). They can alleviate the pressure of domestic public opinion that decision makers may face when launching and participating in foreign military operations, while increasing the tools available for performing tasks. In particular, AI applications such as autonomous weapons can reduce the potential cost of certain postures and activities that may be necessary but that could lead to an excessive deterioration of the situation.[54] At the same time, for the problems that could be solved through diplomatic negotiation and other means, the risk of conflict is increased.

Third, the intention behind the use of AI may be misunderstood, increasing the risk of conflict escalation. With the development of technology and the evolution of the global situation, national actors may increasingly use AI weapons. But how well they send unambiguous signals to demonstrate their intent is a challenge when performing these tasks. Instead, these activities may be interpreted as a serious provocation to security interests, leading to a more stringent response from the target country. This could result in unnecessary conflict escalation.

[54] Liu (note 47), p. 67.

Moreover, autonomous weapons are highly dependent on perception and exchange of information about the external environment. As a result, the likelihood of accidents and human-induced malicious interventions increases. For example, if a drone is subjected to a hack or other form of electromagnetic interference while performing a reconnaissance mission and this results in abnormal behaviour such as a crash, an impact or an explosion, the target may misjudge or make an escalatory response.[55]

The psychological anxiety effect of AI

The fifth way in which AI can affect strategic stability is through its psychological anxiety effect. This can lead to strategic mutual suspicion and arms racing and thus affect strategic stability (see table 10.5).

First, there is a concern among countries that their level of AI technology will be surpassed. Technology is the foundation of various strategic capabilities in nuclear, space and conventional forces. It is therefore generally believed that a new generation of AI will become an important strategic deterrent. As described above, AI may overturn the foundation of the nuclear deterrence strategy by 2040.[56] Just like the cold war of the 1940s and 1950s, each side has a reason to fear that its opponents could gain a technical advantage. In late 2017, President Putin hinted that AI may be the way in which Russia rebalances US power in defence.[57] Russian state media subsequently reported that AI is the key to Russia's defeat of the USA.[58]

Second, there are concerns among states that the AI-related rules system will be pre-emptively formulated by the major powers. Elements of strategic stability include technical and behavioural factors, as well as institutional ones. The rules system of AI technology and applications can rebuild technical and behavioural factors. At present, however, AI research is still in its infancy. As a result, the international norms at the relevant technical and behavioural level are still in essence unwritten. Historical development shows that the successful pioneers of technological development are often the makers of the rules and regulations. Generally, latecomers can only passively accept rules and regulations. Even if it is possible for them to formulate new rules, this is difficult. Therefore, major countries have stepped up their R&D related to AI, hoping to take the lead in this rule-making round of competition.

The third concern of states is how their loss of great power status could have an impact on their voice in international diplomacy. AI will become another status symbol of great power. Without occupying the commanding heights of AI, it will be difficult to have a prominent stake in the future international arena.

[55] Liu (note 47), p. 65.

[56] Geist and Lohn (note 38).

[57] President of Russia, 'Расширенное заседание коллегии Министерства обороны' [Extended meeting of the board of the Ministry of Defence], 22 Dec. 2017.

[58] '新的冷战？专家警告说，人工智能是全球军备竞赛的"首选武器"' [New cold war? Experts warn that artificial intelligence is the 'preferred weapon' of the global arms race], 网易号 [NetEase], 31 Jan. 2018.

Table 10.5. The psychological anxiety effect of AI

Psychological anxiety effect on strategic stability	Main impact	Impact on strategic stability
Concern that AI technology will be surpassed	Psychological anxiety leads to a blind pursuit of strategic advantage, rather than strategic stability Strategic mutual doubt can be caused by psychological anxiety	× ×
Concern over pre-emptive AI-related rulemaking		× ×
Concern over the loss of great power status and its impact on a country's international diplomatic voice		× ×

× = negative effect; AI = artificial intelligence.

Nuclear weapons were once the most important symbol of great power status. Today, the strategic capabilities of AI not only illustrate military power, but also demonstrate the level of a country's technological and industrial development. Having AI strategic capabilities will greatly enhance a country's voice within the international diplomatic struggle.

Of course, in addition to the concerns affecting strategic stability, great powers have also emphasized other concerns about AI and national security in important documents, such as national security strategies and science and technology development strategies. Among these are fear of losing control of AI technology. In essence, AI is easy to obtain through technical means, extremely difficult to control and has a low threshold for abuse. It can easily fall into the hands of extremist individuals, criminal gangs or even terrorist organizations, thus posing a major threat to political security and social stability. Another example is the fear of major security risks in AI applications. The application of AI technology has many uncertainties. As such, without predictive, early-warning and preventive capabilities, systematic and catastrophic risks in what could be called the 'AI era' are inevitable.

The impact of psychological anxiety caused by AI can be divided into two categories.

First, the anxiety brought on by the blind pursuit of strategic advantage is a destructive factor when it comes to strategic stability. Because of the instrumental rationality of strategic stability thinking, the strategic goal of a great power is often not strategic stability, but rather the pursuit of strategic advantage. Yet strategic stability is worth pursuing instead of strategic advantage. Strategic stability is a state in which great powers can pursue strategic advantage. According to a 2017 report, phenomena similar to the development of nuclear weapons by the USA and the USSR after World War II are taking place.[59] Countries may agree to propose a digital Geneva Convention that limits AI weapons, but this does not prevent independent nationalist groups, militias, criminal organizations, terrorists

[59] Allen, G. and Chan, T., *Artificial Intelligence and National Security* (Harvard Kennedy School, Belfer Center for Science and International Affairs: Cambridge, MA, July 2017).

and other countries from developing AI and carrying out AI attacks. Moreover, a country can withdraw from any treaty. So, it is almost certain that one party will turn AI into a weapon, even if this is just based on a desire to engage in self-defence. Between strategic advantage and strategic stability, the blind pursuit of AI-related strategic advantages is a potential hazard for the maintenance of strategic stability, because technology is viewed as an important factor in changing the balance of offence and defence. According to the theory of offence and defence, when the balance between the two shifts to make offence dominant, the weapon system with higher mobility and self-protection will enhance the attack advantage and increase the possibility that a pre-emptive attack will be launched.

Second, strategic mutual doubt caused by psychological anxiety is also a destructive factor of strategic stability. From the point of view of AI, no one can accurately predict what kind of conditions will be produced by unmanned vehicles and intelligent warfare. Lowered warfare thresholds, expanded arsenal scales and uncertain technological evolution paths make these AI-related arms races a new source of strategic mutual distrust among states.[60] Incomplete mastery of AI will only increase uncertainty about the ability of a particular opponent's AI to engage in attack and defence. According to some nuclear experts and AI researchers, China and Russia seem to believe that the USA is trying to use AI to threaten the viability of their strategic nuclear forces and trigger mutual suspicion.[61] As a result, disastrous consequences can occur in a crisis. Strategic mutual distrust has also led to a lack of information sharing among major powers in the field of AI. This exposes decision makers to the risk of potentially unwise judgments and reduces strategic stability.

V. Conclusions

Strategic stability did not end along with the cold war. On the contrary, the concept of strategic stability broadened following the conclusion of the cold war. Nuclear strategic stability during the cold war has developed into today's complex strategic stability. Its categories have expanded from nuclear power relations via military and security relations to overall strategic relations. Its protagonists have grown beyond the United States and the Soviet Union to include various global actors.

There is a feasible basis for AI as a 'second nuclear weapon' to have an impact on strategic stability. This is based on the openness of the strategic stability environment, which includes hegemony, great power status and the fragility of great power strategic stability relations. It is also based on instrumental rationalism derived from cold war thinking, fatalistic realism and low strategic trust among great powers. Most importantly, the numerical growth of these factors suggests that AI has great potential for strategic stability. Among the three elements of

[60] Liu (note 47), p. 63.
[61] Geist and Lohn (note 38).

strategic stability—technical factors, behavioural factors and institutional factors—technical factors establish the material basis for the comparison of strategic strength among countries. Technical factors not only determine the level of nuclear forces, but also the ability to engage in military modernization and the level of conventional armed forces. These are fundamental elements in determining strategic stability.

There are five pathways for AI to have an impact on strategic stability: its empowerment effect on nuclear weapons, its enhancement effect on conventional military forces, its comprehensive penetrative effect on strategic capabilities, the behavioural risk effects that lead to conflict escalation and the psychological anxiety effect. Although some factors can enhance strategic stability, the impact of AI may be negative in most cases, such as its blurring of the boundaries between conventional and nuclear wars, increasing the choices of armed behaviour and resulting in misunderstanding of intent when employed. The escalation of conflict, the psychological pursuit of strategic advantage instead of strategic stability, and strategic mutual distrust among countries are destructive factors that have an impact on strategic stability.

AI applications have great potential and may have a significant impact on strategic stability. However, many of the limitations of these applications also merit the attention of strategists. Among AI's many characteristics are its military and civilian use, easy proliferation and data dependence. It will bring significant challenges to existing laws, security and ethics. In terms of security, AI systems are inherently fragile and unpredictable. As such, system accidents and enemy cyberattacks can be catastrophic. Malicious actors may use these vulnerabilities to infiltrate nuclear weapon systems, while the injured state may be unaware. The 2018 US Nuclear Posture Review specifically addresses the impact of cyberthreats on nuclear command, control and communications (NC3) systems.[62]

In addition, the development of AI weapons represented by LAWS and arms racing may endanger human peace, stability and even survival. On the legal front, the rapid development of AI and militarization trends have seriously affected the core principles of distinction, proportionality and humanity in the existing international law of armed conflict. On the ethical side, the rise of machines brought about by AI has brought enormous challenges to traditional human–machine relations. Whether, what and how human moral standards should be embedded in increasingly intelligent machines needs to be studied in depth.

Given the potential impact of AI on strategic stability, it is necessary to design a framework for maintaining strategic stability in the AI era as soon as possible. Regarding technical factors, countries can cooperate on researching the vulnerability of AI systems, while maximizing the role of AI. On behavioural factors, major countries should not only establish a communication channel for crisis management but also consider a response plan for machine learning, judgement and execution. Nuclear attacks cannot be withdrawn, so the real dilemma is in how to prevent nuclear crises and how to mitigate the transformation of

[62] US Department of Defense (DOD), *Nuclear Posture Review* (DOD: Washington, DC, Feb. 2018).

traditional behaviour into nuclear crises. Once a potential nuclear crisis has occurred, it must be prevented from further escalation. In terms of institutional factors, major countries need to jointly build AI-related mechanisms to prevent the illegal proliferation and malicious use of AI technology, rationally regulate the military application of AI, and prevent excessive dependence on AI. Most importantly, countries should build strategic mutual trust in the era of AI on the basis of all these factors, thereby promoting strategic stability and advancing the process of world peace and development.