

Vol.1 No.1(2019)

INTERNATIONAL
SECURITY AND STRATEGY STUDIES
REPORT

Some Thoughts and Analyses
on How AI Will Impact International Relations



清华大学战略与安全研究中心

CENTER FOR
INTERNATIONAL SECURITY AND STRATEGY
TSINGHUA UNIVERSITY

Some Thoughts and Analyses on How AI Will Impact International Relations

Fu Ying

[Abstract] This paper focuses on how artificial intelligence (AI) will impact the international order by influencing the international configuration and norms. In terms of the international configuration, AI may shift the balance of economic and military power among countries, empower the non-state actors in an unprecedented way and intensify international technological competition. With regard to international norms, AI is likely to change the forms and principles of war, thus exerting an impact on the existing international laws and ethics. This paper argues that the security and governance challenges brought by AI require the collective response of the humanity, and that countries, when discussing and exploring future international norms governing AI, may proceed from the vision of building a community with a shared future for all mankind, as well as the concept of common security. Meanwhile, this paper also proposes six principles for AI development, which include well-being, security, sharing, peace, rule of law, and cooperation.

[Key words] AI, International Configuration, International Norms, Common Norms

[About the Author] Fu Ying is Adjunct Professor of Tsinghua University, Honorary Dean of the Institute of International Relations of Tsinghua University and Chairperson of the Center for International Security and Strategy(CISS).

Email: fyding@tsinghua.edu.cn

Launched by the Center for International Strategy and Security of Tsinghua University (CISS) in 2018, the AI and Security Project Team is tasked with exploring the impact of AI technologies on national security and international relations, and the possibility of constructing common norms. Based on a summary of information gathered by the project team, this paper presents a preliminary analysis of how AI will impact international relations and vice-versa.

In 1950, British scientist Alan Mathison Turing introduced the method to determine whether a machine has intelligence, which has later been called the Turing Test.¹ In 1955, American scientist John McCarthy coined the term “artificial intelligence”. In 1956, the first AI conference was held in Dartmouth, New Hampshire, USA, in which AI as an academic discipline was acknowledged by the scientific community. In 1997, an IBM computer called Deep Blue defeated the world’s top chess player Garry Kasparov in a six-game chess match. In the second decade of the 21st century, the research and development of AI technologies have picked up pace. In 2014, Google’s AlphaGo program triumphed in its games against South Korean Go grandmaster Lee Sedol. Over 60 years on AI has been widely applied to more and more sectors. Moreover, in some specialized fields, it is catching up or even surpassed the human brain. As a ubiquitous technology with the potential to change human society, AI has sparked heated discussions in such areas as technology, industry, military affairs, society, and ethics.

Will AI have an impact on international relations and, if so, how? This paper intends to explore these questions. Given that AI per se is characterized by indescribability, complexity and uncertainty and that

¹ In 1950, Turing’s paper, “Computing Machinery and Intelligence”, was published, in which he introduced his concept of what is now known as the Turing test: if a tester cannot tell whether the answers to a set of questions come from a computer or from a human, it can be concluded that the computer is intelligent.

the author is not an AI expert, it should be noted that this paper, based on the history of AI and the trend generally accepted by academia, is mainly aimed at analyzing the impact of AI on international relations and exploring the necessity and possibility of constructing common norms.

Clearly, there have been many inflated expectations on how science and technology would impact modern international relations. For example, in his 1980 book *The Third Wave*, Alvin Toffler predicted that the future world would face prevalent nuclear weapon risks, teeter on the edge of economic and ecological collapses, witness the obsolescence of existing political institutions, thus landing itself in an enormous crisis. Such predictions tend to overestimate the difficulties posed by technology for humanity, but underestimate the will and capability of humankind to overcome such difficulties. Against the backdrop of post-Cold War globalization, multilateralism has gradually become an international consensus, the global nuclear nonproliferation regime has proved effective, global campaigns have been launched against climate change and follow-up cooperation has been further reinforced, and peace movements have been carried out all over the world — all these demonstrate the consensus and sense of responsibility of human beings on safeguarding peace and addressing challenges. Problems brought by technologies can be solved with new technologies, while a rigorous precautionary system based on ethics and laws is also required. In fact, every technological revolution in history has accelerated globalization and enabled the listing of more global issues on the agenda of international politics, thus creating a more transparent and integrated world.

I. What Are We Discussing?

Several questions need to be clarified before touching on the core issue.

The first question is what kind of AI we are discussing? Is it AI in its narrow sense which can simulate one or two specific intelligent behaviors of human beings, such as identifying things, learning, reasoning and judging? Or is it general AI which has consciousness and independent innovative capabilities similar to the human brain? Or is it weak AI which is focused on particularly specific tasks and can only work on voice recognition, image recognition and translation of certain materials, like Google's AlphaGo and iFLYTEK's smart translator? Or is it strong AI, namely human-level intelligence which can think, plan, solve problems, engage in abstract thinking, comprehend complicated ideas, learn rapidly, and learn from experience, like the artificial kid, David, in the movie *A.I. Artificial Intelligence*, or Ava in *Ex Machina*? Or is it super-intelligence in the future, namely "Synthetics" which, having crossed the "singularity"², can compute and think far better than the human brain, and are even "much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills", as described by the Oxford philosopher Nick Bostrom?

The discussion concerning the impact of AI on international relations and configuration should be limited to the existing AI technologies and applications which are based on big data and deep learning and characterized by three major elements: computing power, algorithms, and data. It is impractical to discuss future AI technologies which are yet to

2 In 1990, in his books *The Singularity Is Near: When Humans Transcend Biology and How to Create a Mind: The Secret of Human Thought Revealed*, the US futurist Ray Kurzweil used "singularity" as a metaphor for a point where AI will surpass the intelligence of humanity. After AI crosses the "singularity", all traditions, knowledge, ideas, and common sense people take for granted will not exist. The accelerating progress in technologies will cause a runaway effect wherein artificial intelligence will exceed human intellectual capacity and control and thus radically change civilization.

make breakthroughs in whole brain emulation. Therefore, an analysis of how AI has impacted and may continue to impact international relations should be grounded on the existing and predictable AI technologies and their trends.

The second question is: can AI impact international relations and, ultimately, the international order? At the moment, the answer is yes. Throughout history, technological innovation and dissemination have, for countless times, revolutionized one or several countries and thereby enabled shifts in regional even global configuration. In *The Human Web: A Bird's-Eye View of World History*, William H. McNeill and John R. McNeill drew a vivid picture of how technological revolutions have played a decisive role in warfare, government organization, people's beliefs and ultimately, transfers of power between countries and shifts in regional configuration. The book mentions that circa 1700 BC, the chariot revolution changed the power structures in Mesopotamia, Egypt, India, the Yellow River region of China and so on, resulting in, for example, the Aryans' dominance in northern India and the rise of the Shang dynasty in China. After 1200 BC, the emergence and dissemination of iron smelting technology enabled ordinary infantry equipped with inexpensive iron armor and weapons to overthrow charioteer elites. The combination of cheaper armaments, larger armies, and consolidated bureaucracies made the rise of agricultural empires such as Assyria and Persia possible. In the seventh century BC, a sizable cavalry that excelled at horse archery once again disrupted the military and political balance of Eurasia after which nomadic groups from pastures regained an advantage over farming groups.³

3 William H. McNeill and John R. McNeill, *The Human Web: A Bird's-Eye View of World History*. Translated by Wang, J. X. et al. Beijing: Peking University Press, 2017, p. 72-79.

Nuclear technologies provide another example, since they have changed the political landscape in the modern world and further consolidated the configuration of great powers formed at the end of World War II. Being nuclear states has enabled the US, the former Soviet Union (replaced by Russia after its demise), the UK, France, and China to maintain their permanent seats on the UN Security Council. Moreover, a series of international norms have been created, such as the peaceful use of nuclear energy, nuclear states' obligation to nuclear nonproliferation and non-nuclear states' entitlement to peaceful nuclear technologies. The developments have also given birth to international values that nuclear wars will destroy humanity and that nuclear proliferation is evil and illegal. Additionally, a set of international institutional arrangements have taken shape, such as *Treaty on the Non-Proliferation of Nuclear Weapons*, *Comprehensive Nuclear Test Ban Treaty*, the UN mechanism for nuclear disarmament negotiations, the Nuclear Security Summit and the Southeast Asian Nuclear-Weapon-Free Zone.

As is believed, AI, like nuclear technologies, can be used for both military and civil purposes and will disruptively change the balance of power in the world. In July 2017, the Robert and Renée Belfer Center for Science and International Affairs under the John F. Kennedy School of Government of Harvard University published a report titled “Artificial Intelligence and National Security”, suggesting that future AI might be empowered to serve as a transformative national security technology, on par with four prior cases of military technology – nuclear, aerospace, cyber and biotech.⁴ Therefore, it is reasonable to include AI in the discussions of factors

4 Greg Allen and Taniel Chan, “Artificial Intelligence and National Security,” Paper, Belfer Center for Science and International Affairs, Harvard Kennedy School, July 2017, p. 1.

impacting international relations.

AI can even influence shifts in the international order. Professor Wang Jisi of Peking University holds that the international order consists of two parts: One is the power structure and balance of major countries and state groups; the other is the norms which should be followed in handling inter-state relations.⁵ Professor Yan Xuetong of Tsinghua University stated that the international order is “the state where countries use nonviolent means to address conflicts in accordance with international norms in the international system”, and such a state is composed of international mainstream values, international norms and international institutional arrangements.⁶ He also notes that shifts in the international order are caused by changes in the international configuration, though the latter is not a constituent of the former; and that the establishment of a new international order is essentially a reallocation of international power, which is at the core of international institutional rearrangements.⁷ Both scholars emphasize the international configuration and international norms when defining the international order. It is possible that AI can change the balance of and relationships among global actors, upset the existing international norms and prompt new ones, and thereby shape the international order by changing the international configuration and international norms.

II. How Will AI Impact the International

5 Wang Jisi, *Ultimate Goals of World Politics*. Beijing: CITIC Press, 2018, p.30.

6 Yan Xuetong, “International Order in a Disordered System,” *Quarterly Journal of International Politics*, 2016, 1: pp.13-14.

7 Yan Xuetong, “International Order in a Disordered System,” *Quarterly Journal of International Politics*, 2016, 1: p. 10, 15.

Landscape?

First and foremost, AI will affect the balance of power from the economic aspect, and even trigger a new round of rise and fall of major powers.

In *The Rise and Fall of the Great Powers*, Paul Kennedy points out that in the long run, there is an obvious relationship between the economy of every major country and its position as a global power. In June 2017, in *Sizing the Prize*, its 2017 Summer Davos report, PwC predicted that by 2030, AI will make an annual contribution of US\$ 15.7 trillion to the world economy, where China and North America are expected to become the greatest beneficiaries, gaining US\$ 10.7 trillion. In *Notes from the AI Frontier: Modeling the Impact of AI on the World Economy*, released by the McKinsey Global Institute in September 2018, it states that AI will significantly improve global overall productivity. Setting aside competition and transition costs, by 2030, AI will contribute an additional US\$ 13 trillion to global GDP and spur growth by an average of 1.2% per year⁸, matching or even outweighing the revolutionary effects brought about by some other general technologies in history, e.g. the steam engine in the 19th century, industrial manufacturing in the 20th century and information technology in the 21st century. The report also points out that leading countries and regions in the AI industry (mainly developed economies) will witness an economic growth of 20% to 25%, while emerging economies may only enjoy half as much. The “AI divide” may

8 McKinsey Global Institute, *Notes from the AI Frontier: Modeling the Impact of AI on the World Economy*. Discussion Paper, September 2018, p.1. Retrieved February 15, 2019, from <https://www.mckinsey.com/~media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Notes%20from%20the%20frontier%20Modeling%20the%20impact%20of%20AI%20on%20the%20world%20economy/MGI-Notes-from-the-AI-frontier-Modeling-the-impact-of-AI-on-the-world-economy-September-2018.ashx>.

further deepen the “digital divide”.⁹ AI may change the global industry chain. “New Industrialization” represented by industrial robots and intelligent manufacturing will cause the manufacturing to “flow back” to developed economies. As a result, many developing countries will experience “deindustrialization” at an earlier time, or even permanently lose the opportunity of industrialization, and will be locked in the position of resource suppliers. Since AI development and application are capital and technology intensive, it may lead to changes in the employment structure, and highly repetitive jobs with low technology content may gradually disappear.

In addition, in another report published in 2017, based on studies of 46 countries and 800 occupations, McKinsey predicts that by 2030, 800 million people worldwide will be replaced by automatic robots and lose their jobs, and one fifth of the labor force will be affected globally. There will be a mass shift of positions as the one that occurred in the first years of the 20th century, when most people turned from agriculture to industry.¹⁰ In short, labor will be replaced by capital and technology. Meanwhile, the wide application of AI technologies will also increase the need for professional talents in this field.

Research shows that three types of countries are the most likely to benefit from the development of AI technologies. The first are countries with the first-mover advantage, such as the US and China. The second are capital

9 Ibid., p.34.

10 Rich Miller, “Robots Are Coming for Jobs of as Many as 800 Million Worldwide,” Bloomberg, Nov 29, 2017, Retrieved February 16, 2019 from <https://www.bloomberg.com/news/articles/2017-11-29/robots-are-coming-for-jobs-of-as-many-as-800-million-worldwide>.

and technology-intensive countries with relatively limited or declining populations, e.g. Japan, Korea and Singapore, which have the capital and technology for developing AI and will be able to use AI development to make up for their limited, declining or aging populations. The third refers to countries with more scientists, mathematicians and engineers, or those attaching importance to professional education in science, technology, engineering and mathematics (STEM).

Next, AI will change the balance of power between countries in the military field. Advocates for military intelligence believe that AI will revolutionize the form and pattern of war. Mechanized warfare relies on substance such as oil and steel to release energy; while informationized warfare gathers power with a network, which focuses on information and links. Based on current predictions, once warfare enters the era of intelligence, AI will harness the power, where robots and automatic wars will prevail.

It can be foreseen that following intellectualization, the combatants, concepts of fighting and philosophy of winning will all be changed. In conventional wars, despite differences in weaponry, armament and training between two opposing sides, the disadvantaged may still strive with favorable weather and geographical conditions, superior schemes or advanced tactics. For instance, during the wars in Iraq and Afghanistan, the US army was inflicted by IEDs (Improvised Explosive Device). In contrast, in intellectualized wars, the technological advantage in AI obtained by one side can soon become the overwhelming superiority on the battlefield. Therefore, the disadvantaged side will be unable to effectively form the observation-judgement-decision-action cycle, thus falling into a passive position. In its report titled *How Artificial Intelligence Is Transforming the World*, the Brookings Institution puts forward the concept of hyperwar, meaning that a war is a race with time, and the one

that decides and executes the fastest will often prevail. With the aid of AI, the decision-making speed of command and control systems will be much faster than in conventional wars. And with the addition of the automatic weaponry system that are capable of autonomously deciding to launch lethal weapons, the pace of warfare will be greatly accelerated. This is why the new term “hyperwar” is created to describe this mode of war.¹¹ The latter was illustrated in “What Happens When Your Bomb-Defusing Robot Becomes a Weapon” on The Atlantic on April 26, 2018, which explains in detail how military robots can significantly reduce threats posed by roadside bombs.¹²

AI will also trigger revolutionary changes in military equipment. The clustered operation of unmanned automatic lethal weapons may play a leading role in future wars and become a main mode of war-fighting. For example, once the envisaged “swarms” of unmanned aerial vehicles (UAVs), “wolf packs” of unmanned submarines, ground robots and “ant colonies” of unmanned tanks appear in battlefields, the value of “war of attrition” and “huge-crowd strategy” will be reconstructed, making complex and expensive mass weaponry platforms such as aircraft carriers and F35 fighters less advantageous from the comprehensive perspectives of cost and efficiency. It can be imagined that when facing armed UAV clusters with a unit price of a few thousand dollars, F35 fighters, costing hundreds of millions each, will be just like “canons versus mosquitoes”.

11 Darrell M. West and John R. Allen, How Artificial Intelligence Is Transforming the World. Brookings Institution, April 2018. Retrieved February 16, 2019, from <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/>.

12 Caroline Lester, “What Happens When Your Bomb-Defusing Robot Becomes a Weapon,” The Atlantic, April 2018. Retrieved February 16, 2019, from <https://www.theatlantic.com/technology/archive/2018/04/what-happens-when-your-bomb-defusing-robot-becomes-a-weapon/558758/>.

It is worth explaining that the influence of AI on military affairs is still largely uncertain, including its scope and manner. At the “AI and Security” sub-forum of the 7th World Peace Forum of Tsinghua University in July 2018, experts pointed out that while the future trends of machine learning, industrial robots and material science can be roughly judged, how these technologies will exert a combined influence on future warfare is still far from being accurately estimated. In the first 30 years of the 20th century, European military powers such as Germany, UK, France and Italy all developed tanks, airplanes, and radio communications technology. However, it was only after Germany carried out the “Blitzkrieg” in World War II that the world discovered how new technologies as a whole could bring about such revolutionary changes to warfare. At present, in terms of either the “war of algorithm” or the “swarm” tactics, strategists are still debating the influence of a single technology on military operations. If AI’s military applications cannot be understood as a whole, the counter measures may become a new “Maginot Line”, expensive but futile.

What’s more, according to findings in the research done so far, the influence of AI on the international landscape can be summarized in the following four aspects.

Firstly, it can greatly augment the global wealth. However, the distribution of wealth will become even more unbalanced among countries, with leading countries in AI benefiting more and those lagging behind benefiting less. In this sense, AI will further widen the gap between countries in terms of economic quantity and quality. And the distribution of the additional wealth will be unbalanced among regions in the world as well: whereas the US is prepared to cement its place as the world economic center, East Asia, possibly benefiting most from AI technologies, is expected to gain further momentum. Meanwhile, the wealth distribution among different regions,

sectors and groups within a country may also become more uneven, thus affecting to varying degrees political equilibrium in relevant countries.

Secondly, insofar as the three factors of capital, technology and labor are concerned, AI will be beneficial to countries with comparative edges in the first two factors, but not to those advantaged in labor.

Thirdly, the core elements of the current AI technologies are algorithms, computing power and data. Countries with more computing resources, more advanced algorithms and more data, are more likely to obtain economic advantages. In the future, there may be “algorithm games” and “algorithm hegemony”, and data will become national strategic resources.

Fourthly, due to the fast iteration of AI technology, with the Matthew Effect, countries with first-mover advantage will continuously consolidate their advantages, making it difficult for late-movers to catch up, and leading to “class solidification” in the international society. Since the current development of AI technology is based on progress of computer technologies and is closely related to development of digital economy, countries leading in digital economy and technologies will get the upper hand in the international AI competition.

Last but not least, while the economic and military changes brought about by AI affect the balance of power between countries, they may also mean changes in international trends.

1. Will northern countries gain momentum at the expense of countries in the South, or vice versa? With the subsiding aftermath of the financial crisis, developed countries have been gradually recovering, whereas some BRICS members have successively encountered difficulties, other developing countries are caught up in slower pace, and “the rise of southern countries” has started to lose steam. Setting other factors aside,

the development of AI technology may add to such a momentum. On the one hand, developed economies are set to sustain or even expand their overall edge over developing ones. On the other hand, there can be a further polarization among developing countries in terms of their strengths and interests.

2. Will there be multi-polarization or polarization? Professor Yan Xuetong believes that the international landscape is developing towards bipolarization with China and the US becoming the two poles, rather than multi-polarization.¹³ The development of AI technology has provided some ground for his viewpoint. The US and China are listed in the first echelon by many international AI ranking. As the headstream of AI, the US is obviously a global leader in innovative spirit, basic theories, core algorithms, high-end chips, as well as the number and quality of employees. China is more advanced in face recognition and voice recognition, and has the natural advantages of mass data and application scenarios that are incomparable in other countries. China has the largest number of internet users in the world, constituting the most active group of data generation. The 42nd China Statistical Report on Internet Development, published by China Internet Network Information Center on August 20, 2018, indicates that by June 2018, China had 802 million netizens, generating around 13% of the global data annually.¹⁴ With more graduate students specializing in AI-related sectors, China has more potential in talent supply. According to the Comprehensive Interpretation of AI Industry Development in

13 Yan Xuetong, "Unipolar or Multipolar? A Bipolar World Is More Likely," *China-US Focus*, Vol.6. April 2015, pp.12-15.

14 China Internet Network Information Center: *China Statistical Report on Internet Development*, August 2018, p20. Retrieved February 18, 2019, from http://www.cac.gov.cn/2018-08/20/c_1123296882.htm.

China and US published by the Tencent Research Institute, by June 2017, there were 2,542 AI enterprises worldwide, 1,078 (42%) of which were in the US, with about 78,000 employees; 592 (23%) were in China, with about 39,000 employees; the other 872 AI enterprises were in Sweden, Singapore, Japan, UK, Australia, Israel, India, etc.¹⁵ The AI competition currently happens primarily among giant enterprises, while major internet giants are transferring their resources to the field of AI. Of the internet giants worldwide, the US has five (Apple, Google, Microsoft, Amazon and Facebook), while China has three (Baidu, Alibaba and Tencent). The US and China, which are the top two powers in AI as well as the top two giants in terms of comprehensive national strengths in the world, may leave other countries further behind in the years to come.

3. The power of non-state actors will expand. AI will endow giant enterprises with unprecedented powers. In 2016, Cambridge Analytica Ltd. swayed the US election through Facebook, showing how politically and socially influential these enterprises can be. In the future, many traditional administrative functions can be undertaken by AI enterprises, including statistics and information collection, information release, social welfare payments, public policy discussion, assessment and feedback, as well as campaigns relating to public opinions on military and diplomatic affairs. In an era when information means power, the influence of corporate behaviors on domestic politics and even international politics may continuously increase. Stanford University and Infinite, a US startup company, have jointly developed an autonomous cyber-attack system based on AI processing chips, which can independently study the cyber environment

¹⁵ Tencent Research Institute: *Comprehensive Interpretation of AI Industry Development in China and US*, July 26, 2017, p3. Retrieved February 18, 2019, from http://www.tisi.org/Public/Uploads/file/20170802/20170802172414_51007.pdf.

and generate certain malicious codes to attack or steal information from the designated network. By applying AI to the search for network vulnerabilities, cyber operations will be more efficient, and the means of attack will become more insidious and intelligent. Traditional protection, based on virus databases and behavior recognition, will be unable to cope with flexible AI virus generation systems, where the generation, execution and infection of malicious codes are more obscure, rendering cyber security under a greater challenge.¹⁶

What is more disturbing is that AI may also provide terrorists, hackers and criminals with new means. For instance, terrorists may remotely control UAVs or self-driving vehicles to conduct assassinations and destructive activities, and hackers may apply big-data weapons in more complex and automatic mass attacks targeted at key infrastructures. Frequently mentioned cases include: Russia's military bases in Syria's Khmeimim and Tartus were attacked by militants' UAVs; Venezuelan President Maduro was attacked by UAVs while delivering a speech; and the outburst of the virus WannaCry. The Canadian scholar Amitav Acharya put forward the concept of a "Multiplex World".¹⁷ There is no doubt that AI will add to multiplexity of the world, including the decentralization of political powers, appearance of more diversified actors, and the diversification of politics and ideologies, thus more inclusive and multi-layered governance will be needed.

4.AI will intensify scientific competition. The development and application of AI technologies further enhanced the position of science and technology

16 Academic Plus: "AI Weaponization in US Army," *Sohu Military*, January 7, 2019. Retrieved February 19, 2019, from http://www.sohu.com/a/287109476_100044418.

17 Amitav Acharya: *China and the Crisis of Liberal Internationalism, Global Order* 2018(1), p81.

in international competition. In May 2016, the Committee on Homeland and National Security of the US National Science and Technology Council published the 21st Century Science, Technology, and Innovation Strategy for America's National Security, which points out that science and technology have become the vital part of a country's international competitiveness. Without scientific and technological innovation, there will be no national security. The US has been able to maintain its edge in international competition largely thanks to its innovation power in science and technology. Many Chinese scholars believe that the China-US trade war is essentially one concerning science and technology, and the two countries have begun their power game first in this field. In 2018, the US banned chip sales to two Chinese companies, and even openly prevented its allies from adopting Huawei's 5G technology. On November 19, the Bureau of Industry and Security of the US Department of Commerce listed 14 "representative technology categories" to be discussed for potential control, where the categorization and control of AI and machine learning were the most detailed. It can be foreseen that in the era of AI, international competition relating to science and technology will become more intense, and the result will impact the future international landscape to a large extent.

III. How will AI Impact International Norms?

AI may leave the following impacts on existing international norms.

First, AI may alter the connotations and forms of war. The US scholar Stephen Pinker points out the downward trend of violence among human beings and holds that "today we are probably living in the most peaceful

time in our species' existence.”¹⁸ Professor Wang Jisi gives five reasons for the notable reduction of wars in today's world: the costs that warring countries will pay might be much higher than the possible benefits; the benefits brought by wars in the past can be gained without resorting to force now; countries have improved their understanding of each other and generally established crisis prevention and control mechanisms; an international arms control regime has been established; peace, as a worldwide belief, has found its way deep into people's heart.¹⁹

However, the development of AI may ease states' and non-state actors' concerns over whether to resolve conflicts by means of arms. It can be inferred at least theoretically that AI is able to cut down the restrictions on wars and lower war thresholds, from the perspectives of decision makers, soldiers and the general public. Unmanned weapons can reduce the number of anticipated casualties of wars to nearly zero. The rapid iteration of AI technology and wide application of new materials will significantly lower the physical cost of wars, making them more economical and even profitable. As a result, to wage a war will tempt decision makers more and make them worry less. The possibility of beyond-visual-range combat operations and the emergence of robot soldiers will allow soldiers to further relax psychologically. In future wars, soldiers may not have to fight on the battlefield since the combat missions can be executed by unmanned automatic weapons led by “General AI” with pre-designed algorithms. And “hyperwars” can considerably shorten the length of wars, thus bringing wars to an end even before public backlash. All the above may change

18 Stephen Pinker, *The Better Angels of Our Nature: Why Violence Has Declined*. Translated by AN Wen. Beijing: CITIC Press, 2015, p1.

19 Wang Jisi. *The Ultimate Goal of World Politics*. Beijing: CITIC Press, 2018, p. 55-59.

people's stances on wars and allow violence to return.

The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism, a publication released by the Institute of Land Warfare of the Association of the United States Army in October 2018, argues that three particular capabilities of AI, i.e. algorithmic content generation, personalized targeting and firehose dissemination, will generate the "Influence Machine", which will be used to produce the exponential benefits of AI-driven influence operations. The publication also holds that "Influence Machine" information operations are more impactful, at a strategic level, than other applications of AI. This is because information operations, with the aid of machine learning, can exploit emotion, bias or values to concentrate on those target groups that are best placed to be affected. Then "customized mental munitions" can be "fired" on those groups at machine gun speed, shaping their wills and cognitions.²⁰

Second, AI may affect global strategic stability. Joseph Nye once argued that the emergence of nuclear weapons has pushed military force to its limit as a means of maintaining security, even contrary to the expectation. It is recognized that there is no winner in a nuclear war and that nuclear weapons may put the whole of mankind in danger of extinction.²¹ However, AI will challenge the classic logic of "mutually assured destruction". In its document titled, "How Might Artificial Intelligence Affect the Risk of Nuclear War?" released in 2018, the Rand Corporation indicates that by 2040, the progress of AI will greatly increase the possibility of targeting and destroying adverse retaliatory nuclear forces, thereby undermining

20 Academic Plus, "A Summary of US AI Weapons," *Sohu Military*. January 7, 2019. Retrieved February 19, 2019, from http://www.sohu.com/a/287109476_100044418.

21 Joseph Nye, *Understanding International Conflicts: An Introduction to Theory and History*, Shanghai: Shanghai People's Publishing House, p. 201-207.

the premise of mutually assured destruction and disrupting the strategic nuclear balance. Even if states have no intention of launching pre-emptive attacks, they will tend to develop this capability so as to use it as a leverage in bargaining with their rivals, which will undoubtedly undermine strategic stability.²²

Third, the characteristics of AI will make it difficult to establish relevant international arms control and non-proliferation mechanisms. During the parallel session with the theme of “AI and Security” at the 2018 World Peace Forum held at Tsinghua University, Gregory C. Allen, co-author of the report *Strategic Competition in an Era of Artificial Intelligence*, said that AI can be used for both military and civil purposes, but its military application, unlike nuclear weapons, is more analogous to electricity. Whereas nuclear technology is regarded as a black-or-white one which means a state is either nuclear or non-nuclear. Every country can use electricity, though how and to what extent the electric power is used may vary greatly. The application of AI for military purposes is more like electricity, which means it is impossible to prohibit any state from using it. There might be an algorithm race between AI powers in the future, just like the arms race between the US and the Soviet Union during the Cold War. The US and the Soviet Union signed a series of agreements on nuclear and missile control after arms control discussions at that time, which established the basic rules. The question is: can future AI powers reach algorithm control agreements in the same spirit? Allen believes that given

22 Edward Geist and Andrew J. Lohn, “How Might Artificial Intelligence Affect the Risk of Nuclear War?” *RAND Corporation*, p.8. Retrieved February 19, 2019, from <https://www.rand.org/pubs/perspectives/PE296.html>.

the current relations between the major powers, it is almost impossible to reach such agreements. Nonetheless, to avoid the risks of disorderly development in the future, the powers must consider seriously the necessity of having a joint exploration in this regard.

Fourth, AI has brought a series of problems for international law. With the application of AI weapons, can the relevant principles of the International Humanitarian Law and the law of war still be held true? For example, the principle of “distinction” which requires one to distinguish between combatants and civilians, the principle of “proportionality”, to prohibit excessive attacks, and “the principle of military necessity”, which stresses that arms can only be used when non-military means fail to work, and the limitations placed on means of war-fighting. Is it necessary to formulate special rules for AI weapons? In the AI-guided wars, how can one distinguish between combatants and non-combatants? Are war robots entitled to humanitarian treatment? Can AI weapons assume the responsibility for the damages they cause? If not, who should be the responsibility bearer, AI weapons manufacturers or users? If AI weapons violate the principle of national sovereignty, will their action lead to state responsibility?

Fifth, AI may impact the democratization of international relations. The development of AI is likely to consolidate the international power structure and deepen small and weak countries’ dependence on major powers in science and technology, economy, and security. Once the “winner-takes-all” principle of technological and commercial competition in the AI arena is applied to international relations, it will inevitably affect the principle of sovereign equality between countries, whether they are large or small,

strong or weak. When the competition among major powers intensifies, will alliances once again be adopted as an unavoidable choice for weak countries, though the alliance itself represents inequality between the leader and its allies?

Sixth, AI has brought new challenges for global governance. The development of AI is of great significance for the world to solve the three major predicaments (aging, digitization, and climate change).²³ The graver challenge might be how to treat “human value in the digital age”²⁴. Some countries and cities have begun to try out universal personal income (UPI).²⁵ However, this is an inevitable issue that requires global wisdom and efforts.

We may take data as an example in discussing the relevant global governance. The international community in the recent history of mankind has established generally recognized rules and institutions on cross-border

23 Kathryn Reilly, “UK’s Ageing Crisis Has Become the ‘New’ Climate Change: Can Digital Innovation Save the Care System from Collapsing?” *MedTech Engine*. Retrieved February 20, 2019, from <https://medtechengine.com/article/uks-ageing-crisis-has-become-the-new-climate-change-can-digital-innovation-save-the-care-system-from-collapsing/>; McKinsey Global Institute, “Digitization, AI, and the future of work: Imperatives for Europe,” *Briefing Note*, September 2017, p.1. Retrieved February 20, 2019, from

<https://www.mckinsey.com/~media/McKinsey/Featured%20Insights/Europe/Ten%20imperatives%20for%20Europe%20in%20the%20age%20of%20AI%20and%20automation/Digitization-AI-and-the-future-of-work.ashx>.

24 Pricewaterhouse Coopers, “Human Value in the Digital Age,” December 2018. Retrieved February 20, 2019, from <https://www.pwc.nl/nl/assets/documents/pwc-human-value-in-the-digital-age.pdf>.

25 Annie Nova, “More Americans Now Support A Universal Basic Income,” *CNBC*, Feb 26, 2018. Retrieved February 20, 2019, from <https://www.cnn.com/2018/02/26/roughly-half-of-americans-now-support-universal-basic-income.html>; Peter S. Goodman, “Finland Has Second Thoughts About Giving Free Money to Jobless People,” *New York Times*, Apr 24, 2018. Retrieved February 20, 2019, from <https://www.nytimes.com/2017/07/20/opinion/finland-universal-basic-income.html>.

movement of natural persons, of transportation vehicles such as ships and aircraft, and of capital and commodities, etc. Today, since data has become a resource of increasing importance, several questions appear. What rules and institutional arrangements should be observed in cross-border movement of data? What are the rights and obligations of the stakeholders such as the producer, user, exporter and receiver of the data? What is the relationship between local storage of data and cross-border transfer of data? What is the relationship between the state and the data enterprises? All these will be difficult problems for future global governance. Japan, host of G20 Summit 2019, has already proposed listing global data governance on the G20 agenda.

IV. How Should We Choose?

Today, even those researchers who are at the forefront of AI technology have to admit that there is still a long way to go before robots with intelligence as high as humans are produced. The security threat that AI may impose, for the most part, actually stems from some humans' usage of AI to threaten other humans. Therefore, how humans will exercise self-restraint is the most pressing moral challenge in the progress of AI technology.

On October 14, 2015, an organization called "Intelligence Squared" invited several experts on international issues to debate on the topic, "China and the US are Long-term Enemies" in New York. John Joseph Mearsheimer, professor of political science at the University of Chicago, unsurprisingly talked about "the tragedy of the great power politics". As an opponent, Kevin Rudd, former Prime Minister of Australia, argued that the challenge of diplomacy is to prevent war, and he believed we can. Rudd's words also make sense for AI, in that AI technology can do both good and evil.

Although how to judge and choose a stance regarding good or evil is a long-standing philosophical topic on which a consensus is hard to reach, yet mankind, which boasts rationality and modern civilization, ought to decide to do good on the key issue that concerns the fundamental existence of human beings.

In June 2018, when I was invited to visit iFLYTEK Beijing, there was a discussion about whether there will be a competition between human translators and machines in the future. One executive of iFLYTEK argued that the purpose of developing AI translation technology is not to replace human, but to make translation more accurate and easier. His words helped me to understand the situation better. After all, as Chairman Mao Zedong said, “It is the people that decide the outcome of a war, not one or two new types of weapons”. The end of developing AI technology is to make it serve humanity itself. Whether to do good or evil with AI depends on the people who control the technology.

At present, the discussion on the international order is quite heated. One question that the academia of China and the United States are most concerned about and other countries follow with interest, is this: will the US and China launch a “new Cold War”, thus leading to the re-emergence of parallel orders as what was like during the Cold War? When talking with some Americans and Europeans who came to Beijing for the China Development Forum in March 2018, I found them to be pessimistic about the prospect of China-US relations, as they believed that China’s rapid growth was a structural threat to the US. When I asked if there would be any other choices except confrontation between the two powers, Martin Wolf, the associate editor of the Financial Times, replied, “You have no choice. You can’t stop growing and they can’t stop worrying.” China and the US will not unite with each other unless the mankind is confronted with

a common enemy, an invasion from Mars, for example.

If we still hold a zero-sum game mindset and pursue absolute security, AI will, without a doubt, like atomic bombs and satellites in the 1940s and 1950s, become a new focus in major powers' competition and a catalyst for dividing the world into two or more parallel orders. However, if we look at the problem from the perspective of building a community with a shared future for mankind and proceed from the pursuit of common security, it is easy to realize that the security and governance challenges brought by AI technology are problems faced by all humankind. In this way, we will be inclined to discuss the norms acceptable to all stakeholders in the spirit of equal consultation. The question is: will AI become a challenge like “an invasion from Mars” that can unite China, the US, Russia and other countries in the world?

Technology experts are speaking up about the issue. In a parallel session titled “China-US Relations in Cyberspace” of the fifth World Internet Conference held in Wuzhen, China in 2018, Dave Faber, a former expert from the US Federal Communications Commission, called for the formulation of international guidelines on AI as soon as possible to prevent a repeat of the tragedy and its follow-up consequences caused by the absence of international consensus and self-discipline when nuclear-weapon technology appeared.

In fact, China and the US are conducting in-depth exchanges and cooperation in science and technology research. Statistics from Clarivate Analytics on the number of papers on AI show that from 2013 to 2017, within the field of “Computer Science, Artificial Intelligence” in the Web of Science (WoS) database, the Chinese mainland ranks the first among the 167 countries and regions that participate in the global assessment, with 59,573 articles, accounting for 25.02 % of the total. The United States

ranks the second, accounting for 13.66 %, with 32,527 articles. Among others, the number of articles co-authored by Chinese and Americans grows the fastest. For example, in the past five years, scientists from the US co-authored more papers with their colleagues from China than from any other country, with a total of 4,307 papers. Similarly, China is the largest partner for the US in terms of co-authoring AI papers, far surpassing other countries.²⁶

International discussions and studies on related issues have been conducted in an increasingly in-depth manner. The Group of Governmental Experts (GGE) on Lethal Autonomous Weapons Systems (LAWS) of UN Disarmament Commission has taken into account the possible impacts of LAWS on international security, such as the arms race in LAWS technology, the wider technology gap between developed and developing countries and possibly, a lower war threshold. GGE also suggests that states should reach a common understanding that they must take the legal responsibility for their LAWS; a legally binding document should be formulated after the protocol on the Convention on Certain Conventional Weapons, to preventively ban LAWS.

President Xi Jinping made clear China's stance on AI in his congratulatory letter sent on the occasion of the opening of the World Artificial Intelligence Conference in September 2018. He said that a new generation of AI is flourishing all over the world, giving fresh impetus to economic and social development and profoundly changing people's life and work. All countries should deepen cooperation in addressing new issues brought by AI in law, security, employment, ethics and government governance in a bid to seize this development opportunity. And China is willing to work

²⁶ Clarivate Analytics, "Analysis on the Competitiveness of Countries Productive in AI Technology Papers," December 2018, p. 13.

with other countries in the field of AI to promote development, maintain security and share the outcome.²⁷

Based on the findings in the research by the “Artificial Intelligence and Security” Project Team of Tsinghua University, we propose six principles for AI development:

First, the principle of well-being. The development of AI should serve the common well-being and interests of humankind. Its design and application must follow the basic ethics of human society and respect human dignity and rights;

Second, the principle of security. AI shall not harm human beings. AI systems must be secure, applicable and controllable. Personal privacy should be protected and data breach and abuse prevented. AI algorithms must be traceable and transparent and there should be no algorithm discrimination;

Third, the principle of sharing. The economic prosperity created by AI should be shared by all mankind. A reasonable mechanism should be built to enable more people to benefit from the development of AI technology, and to prevent digital divide;

Fourth, the principle of peace. AI technology must be used for peaceful purposes. Efforts must be made to improve its transparency and establish confidence building measures. Peaceful use of AI should be encouraged whereas LAWS arms race should be prohibited;

Fifth, the principle of the rule of law. The application of AI technology

²⁷ Xinhua News Agency, “Xi’s Congratulatory Letter to World Artificial Intelligence Conference 2018,” Retrieved February 22, 2019, from http://www.xinhuanet.com/politics/leaders/2018-09/17/c_1123441849.htm.

should conform to the purposes of the UN Charter and the basic principles of modern international law such as sovereign equality of states, peaceful settlement of disputes, prohibition of the use of force and non-interference in internal affairs of other states;

Sixth, the principle of cooperation. Countries around the world should promote the exchanges of AI technologies and talents, in an effort to promote and regulate upgrading of AI technologies within an open environment. These principles can serve as the basis for the discussion and formulation of international rules on AI.

The above principles are still abstract and not specific enough. It would need greater effort and input by interested members and the stakeholders in the international community to explore further and find the common denominator. It is worth noting that a large number of young experts engaged in high-end scientific research have called for voluntary self-discipline for the common interests of mankind and that many entrepreneurs have also stated that they will pay attention to ethics and moral factors in their AI technological research and application and stay away from any endeavor that is detrimental to human well-being.

I hope that international discussions on these issues can be held in an increasingly in-depth way and I look forward to more contributions in this regard by Chinese think tanks and scientific and technological groups in cooperation with their counterparts from other countries in the world.

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