1 Artificial intelligence and international conflict in cyberspace

Exploring three sets of issues

Fabio Cristiano, Dennis Broeders, François Delerue, Frédérick Douzet and Aude Géry

Introduction

Over the last three decades, cyberspace developed into a crucial frontier and issue of international conflict. Disproving the initial fear-mongering expectations of fully-fledged wars occurring in and through cyberspace, this conflict increasingly unfolds 'away from' the traditional categories and thresholds of war and peace.[1] As argued by Lucas Kello, cyberspace is neither truly at war nor at peace but maintains a constant condition of 'unpeace.'[2] International conflict in cyberspace primarily occurs in the so-called grey zone and often pertains to the domain of information, data, and their manipulation, culminating in acts of espionage, sabotage, and subversion. As empirical evidence overwhelmingly shows, confrontation in cyberspace mostly consists of low-impact hacking, espionage, disinformation, and surveillance.[3] In light of this, recent scholarly work interrogates whether we should consider conflict in cyberspace as an 'intelligence competition' rather than through the lenses of traditional warfare.[4] At the same time, this does not mean that we should think of cyberspace as the peaceful, yet ungoverned and ungovernable, oasis envisioned by cyber libertarians in the early days of the internet[5] – quite the opposite. States now conventionally conceive of cyberspace as an issue of national security and increasingly safeguard and promote their national interests through both defensive strategies and offensive operations in cyberspace.[6]

In a context where data and information have become increasingly important, it comes as no surprise that the development and application of artificial intelligence (AI) have gained momentum in the various discourses about international conflict in cyberspace. AI technologies – such as machine learning, natural language processing, quantum computing, neural networks, and deep learning – provide military and intelligence agencies with new operational solutions for predicting and countering threats as well as for conducting offensive operations in cyberspace. Besides automating the production of knowledge about cyber threats, AI can also automate decision-making, which could 'dilute' the role of (human) political agency as an element of international conflict in cyberspace. Concerns at the core of the international

DOI: 10.4324/9781003284093-1

debate about Lethal Autonomous Weapon Systems (LAWS) would then also enter the debate about cyber conflict. Moreover, the operational entanglement of AI technologies in cyberspace further blurs the already contested lines between defence and offence in cyberspace,[7] while also challenging the divide between cyber conflict and information operations.[8] Besides opening up new operational milieus, the adoption of AI-enhanced cyber capabilities also represents an important strategic asset for states, with the ongoing global race towards the adoption of these technologies fully embedded in broader geopolitical conflicts, deterrence, securitisation strategies, and techno-nationalist narratives, such as those about digital sovereignty.[9]

The entanglement of AI technologies with cyber conflict raises several issues primarily related to human-machine interaction, the role of (big) data in society, great powers competition, and regulation. While creating the 'illusion' of scientific and data-driven security, delegating security functions to independent machines might expose networks to a whole variety of new risks emerging because of autonomy and automation.[10] Potential biases in the mechanical processing of data can lead to miscalculations and the creation of a broader 'attack surface' and vulnerability for the systems that AI purports to protect. Similarly, the global race towards the acquisition of these technologies also risks further intensifying and polarising international conflict in cyberspace.[11] For these reasons, AI technologies have also gained interest as a normative issue across ethical and legal debates on responsible (state) behaviour in cyberspace – although the debate about autonomy has not fully crossed over from the military domain 'proper' to that of cyber conflict yet.[12] As this volume shows, specific regulatory frameworks and legislations might be required to capture AI as both a potential asset and threat to national security and to the 'open and secure' cyberspace that some countries seek to uphold.

With the intent of exploring the question 'what is at stake with the use of automation in international conflict in cyberspace through AI?', this volume focuses on three themes, namely: (1) technical and operational, (2) strategic and geopolitical, and (3) normative and legal. These also constitute the three parts in which the chapters of this volume are organised. Scholarly work on the relationship between AI and conflict in cyberspace has been produced along somewhat rigid disciplinary boundaries and an even more rigid sociotechnical divide – wherein technical and social scholarship are seldomly brought into a conversation. This volume addresses these themes through a comprehensive and cross-disciplinary approach. In this sense, the organisation of the volume in three parts should not be considered as an analytical or, even less so, a disciplinary demarcation. The remainder of this introductory chapter outlines, and provides context for, the main debates of each of the three parts of the volume.

Technical and operational considerations

AI has emerged as the defining technology of our times and seems to epitomise the ultimate innovation that everybody wants and about which everybody is 'concerned.' States often have a techno-optimistic view of new technologies and look favourably at the prospect of rationalising and perfecting governance through automation,[13] with AI currently being applied to wide and diverse governance domains and issues. The allure of the concept of 'AI' is perhaps best caught by the fact that many applications in government (and outside of it) would still be more aptly labelled as 'classic' automation rather than AI or the introduction of autonomy in systems. However, developments in AI do start to permeate traditional governance by expanding the range, scale, and complexity of operations that can be meaningfully automated, including those associated with cybersecurity. When compared to other governance branches, the application of AI technologies in cybersecurity represents however less of an innovation. Already in the 1990s, machine learning and neural networks were, for instance, applied to the filtering and classification of spam emails.[14] After all, automation constitutes an inherent feature of internet technology and computation. What is relatively new, and of main interest for this volume, is the internationalisation and 'datification' of conflict in cyberspace, where the potential of AI marks a new operational phase through autonomy.

From an operational perspective, AI technologies promise to contribute to one of the core dynamics of international conflict in cyberspace: the identification of vulnerabilities through timely and effective interpretation of data – for either defence or offence. That is, AI has the potential to make conflict in cyberspace more knowable and predictable. When considering aspects of automation and machine autonomy in the context of international conflict in cyberspace, the ability of intelligent machines to make operational choices – at different degrees of independence – points to the question of *who* the actual enactors of international conflict in cyberspace are. As will be further discussed in the third part of this volume, this question is not only analytical or technical. Knowing who enacts conflict in cyberspace also intimately pertains to questions of responsibility.[15] In a context where agency appears to be already diluted through networks, and socio-technical assemblages, exploring the AI-cyber nexus primary means to explore human-mechanic interactions.[16]

The question of autonomy and AI raises an operational interrogative related to the 'place' of humans in relation to the so-called 'loop' of operational decision-making. This dilemma has been foremostly articulated in debates about LAWS where the central question remains whether humans shall be placed in, on, or outside of this loop.[17] In Chapter 2 of this volume, Andrew Dwyer directly addresses this question by analysing the role of deep reinforcement learning (RL) algorithms to question assumptions that AI technologies make conflict in cyberspace more knowable. It argues that, by recognising, performing, and transforming the who, where, and how, of international conflict in cyberspace, AI constitutes more than an epistemic tool for improving operations. In this sense, the chapter also complicates normative considerations about controllable and ethically accountable AI systems and about the place of the human 'in' the loop.

One of the core technical promises of AI for cybersecurity consists in what Tim Stevens defines as a shift 'from known threats to the prolepsis of as-yet-unknown threats and into an anticipatory posture that has received much attention in the critical security literature.'[18] In Chapter 3, Wesley Moy and Kacper Gradon explore the various potential applications of AI in the propagation of disinformation and misinformation, as well as in the context of hybrid and asymmetric warfare. By analysing two methodologies – namely 'Generative Adversarial Networks' and 'Large Language Models' – this chapter explains the relevance of AI for understanding how links are formed, how information is disseminated, and how information can influence opinions and actions in social networks. Taken together, the contributions to the first part of the volume indicate that, while enhancing operational efficiency, AI applications do not necessarily 'make' international conflict more known/predictable and cybersecurity more human-centric. Rather, autonomy and automation further contribute to the problematic understanding of cyberspace as a primarily technical and operational issue or domain.

Strategic and geopolitical considerations

Looking beyond its technical possibilities and operational dilemmas, AI is set to become a constitutional component of economic, political, and military power in the digital age. With the return of great-power competition and the constant contestation and confrontation between states in cyberspace, AI is undergoing a process of securitisation that transforms this dual technology, primarily developed for civilian uses, into a matter of national security and sovereignty.[19] As a result, AI has become fully part of the contested global 'digital arms race,' raising major concerns about the broader risks associated with its use for offensive purposes.[20] This evolution is not surprising. It is in line with the broader securitisation of cyberspace over the past three decades, quickly, but not always correctly, associated with its militarisation in the discourse of states.[21]

With the rise of increasingly sophisticated and targeted state-sponsored cyberattacks since the late 2000s, cyberspace emerged as an imperative of securitisation and a new warfighting domain that required the mobilisation of exceptional means.[22] The representation of cyberspace as predominantly a threat to national security is not self-evident given the complex challenges in this domain, such as those posed by criminal organisations to individual interests that can equally hurt the security of end users and the security and stability of cyberspace itself.[23] Other characterisations that could have prevailed such as economic risk, criminal danger, or threats to individual user privacy have increasingly taken a back seat to national and international security concerns.[24] In the words of internet governance scholar, Milton Mueller 'cybersecurity is eating internet governance' and is pushing out alternative framings.[25] The security frame has progressively extended to all the digital technologies that could be weaponised in the context of digital warfare, including AI, and drives international competition over digital technologies. This competition is both embodied and increasingly shaped by the fierce competition between the United States and China over the production, control, use and governance of digital technologies. Adam Segal argues that during the 1990s and 2000s the integration of the Chinese and American economies was perceived as mutually beneficial, both politically and economically, political decision-makers now consider that the risks outweigh the benefits.[26] And in both state discourses, the issue of security is at the heart of the rivalry. It should be noted that China has launched a massive plan to become the world leader in AI by 2030, with a 150-billion-dollar industry.[27] That is, the talent war is on.

The leadership of a few countries in AI capabilities also reveals uncomfortable strategic dependencies for many other countries. It has triggered a debate in the European Union about the risk associated with these dependencies and the need for strategic autonomy to ensure digital sovereignty.[28] But advancing AI technology appears to be a limited policy option to address these issues. In Chapter 4**,** Simona Soare questions the role of AI to advance European strategic autonomy in the field of security and defence. She argues that the adoption of AI is a ‘distraction’ as it introduces additional layers of complexity in the European defence while not contributing significantly to Europe’s strategic autonomy. On the one hand, the integration of AI in the EU decision-making processes and the conduct of operations is challenging because of the EU’s internal functioning in the field of defence. On the other hand, the lack of industrial capabilities and the strategic dependencies towards other powers are real and likely difficult to overcome. In Chapter 5 Arun Mohan Sukumar similarly demonstrates that relying on AI can introduce risks and strategic dependencies, as shown in the case of emerging powers. The chapter examines the role of AI in the development of public services, through examples of the health sector in Brazil, India and Singapore. It shows how, while states are urged to enhance data transparency and to develop digital services for their population, they become exposed to new risks that could set back progress in the digitalisation of states’ mission-critical systems for years. That is, they face a trade-off between furthering digitalisation and accepting more security risks, an instance that speaks to the importance of thinking about the AI-cyber nexus not only in technical/operational terms but also considering broader strategic implications.

Armed forces worldwide have also recognised the strategic relevance of the AI-cyber nexus and have similarly engaged in a profound digital transformation of their operations.[29] On the one hand, this has considerably increased their reliance on digital technologies and data. On the other, it has created new risks and vulnerabilities. Soldiers evolve in a new digital environment that profoundly transforms the way they operate and creates new challenges that are sometimes hard to fully comprehend and govern. In this environment, AI offers promising new capabilities to improve the quality of intelligence, situational awareness, the conditions of training, the ability to operate remotely, the precision and autonomy of weapon systems and, most importantly, the speed and scope of action. As result, the race for AI is thus also a race for military power and superiority and, again, raises strategic problems that are intimately related to operational ones. This representation resonates with a vision deeply ingrained in the US military culture that technology can provide military superiority. In Chapter 6, Jeppe Jacobsen and Tobias Liebetrau argue that this vision goes back a long time before AI and has dominated US military discourse since the Second Offset Strategy of the 1970s. That is, AI represents an operational innovation more than a strategic one.

While providing further evidence to a presumed return of great powers competition, the military superiority approach also feeds the fears inspired by the technology and is a driver for developing offense over defence, to maintain superiority over the enemy. But AI-enabled cyber capabilities might also convey the idea of control that is difficult if not illusory in cyberspace, given the highly dynamic nature of this environment.[30] And it does not take into consideration the vulnerabilities and associated risks that AI technology also brings about. Indeed, with the digital transformation of societies and armed forces, the attack surface keeps increasing. And while AI can considerably improve defence, the emphasis placed on offense could be a source of risk. Jeppe Jacobsen and Tobias Liebetrau demonstrate that the cyber arms race is not just a competition between great powers for AI-enabled cyber capabilities but also a specific arms race between offensive and defensive cyber capabilities, powered by AI. Given the lessons from discussions on how militaries balance offense and defence in cyberspace, they conclude that AI-enhanced cyber offensive capabilities are likely to dominate. And yet AI can backfire in many ways. As our societies grow increasingly dependent on digital technologies, the securitisation of AI technology could have important spill-over effects on the overall level of cyber (in)stability. The ongoing race for data and its exploitation for strategic advantages further blurs the lines between military and civilian operations, with inextricable consequences for the private sector and civil society, raising new legal and normative challenges.

Normative and legal considerations

Stemming directly from the above-mentioned technical/operational and strategic/geopolitical considerations is the necessity of regulating the adoption and use of AI technologies in cyberspace. The development of cyber capabilities, on the one hand, and AI and its possible applications in cyber conflicts on the other, have posed a dilemma to states and other actors: they are interested in these new technologies – notably to enhance their own operational capabilities and strategic posture – but they are at the same time concerned about the potential consequences of these developments for international peace and security. This dilemma lies at the core of the third final part of this volume, which deals with the normative and legal questions raised by AI applications in cyberspace. To understand these, this section also introduces the international processes in which these normative and legal discussions are embedded and become deeply intertwined with states' strategic considerations.

On international cybersecurity, the United Nations General Assembly adopted its first resolution on "*Developments in the field of information and telecommunications in the context of international security*" in December 1998. Since 2004, the United Nations General Assembly has established six successive Groups of Governmental Experts (GGE) on this topic. The first and the fifth GGE failed to adopt a consensus report, reportedly because of disagreement in the discussions on specific branches of international law. The impossibility of the fifth GGE to adopt a consensus report in June 2017 led to disagreement on how to proceed. In 2018, this resulted in the adoption of two concurrent resolutions and the creation of two parallel processes, with largely the same mandate. In addition to the sixth GGE, an Open-Ended Working Group (OEWG) was established. In 2020, a new OEWG was established which will last until 2025 while there is as of, yet no new GGE planned.[31] Moreover, since 2020, some States are advocating for a new process on this topic, a Program of Action (PoA) for advancing responsible state behaviour in cyberspace,[32] which was welcomed in principle in November 2022 by the UN General Assembly. The second, third, fourth and sixth GGE as well as the first OEWG were successful in adopting consensus reports.[33] These reports notably affirmed that international law is applicable to cyberspace and listed specific rules and principles of international law deemed particularly relevant in this context. They also listed 11 norms of responsible behaviour in cyberspace. Taken together these reports constitute a framework of responsible State behaviour in cyberspace, encompassing international law and non-binding norms but also capacity building and confidence-building measures. Interestingly in the context of this book, the development of AI applications has never been mentioned in the GGE or OEWG reports, despite it being discussed in the 2019–2021 rounds of negotiation. While the issue did not make the cut of the 2021 consensus reports it does feature in the so-called Chair's summary of the OEWG process in its section dedicated to 'Threats': "Pursuit of increasing automation and autonomy in ICT operations was put forward as a specific concern, as were actions that could lead to the reduction or disruption of connectivity, unintended escalation or effects that negatively impact third parties."[34] Moreover, both in the context of the UN negotiations and outside of it, states and other actors have started to voice concerns about the role of automation and autonomy in cyber operations.[35]

The discussions on the international security dimensions of AI have been focusing on the development of LAWS. This matter was introduced in 2013 in the agenda of the Meetings of High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons (CCW). After a few informal meetings, these discussions took a similar path as the ones on international cybersecurity, with the establishment a GGE in 2016 which adopted 11 guiding principles on LAWS in 2019.[36] Through these principles, the GGE affirmed the applicability of international law and in particular international humanitarian law as well as a series of ethical and non-binding principles. Surprisingly, Cybersecurity is only briefly mentioned in the sixth principle as one of the "appropriate non-physical safeguards [that] should be considered [w]hen developing or acquiring new weapons systems based on emerging technologies in the area of lethal autonomous weapons systems."[37] There is, however, no mention of autonomous cyber capabilities. Even though the link between cyber security and AI has been made in the context of the OECD[38] and in the UNESCO *Recommendation on the Ethics of Artificial Intelligence*,[39] both of these documents steer clear of national and international security. So until now, 'cyber' and 'AI' seem to be ships passing in the night in the UN's first committee.

This 'absence' is at the heart of the third part of this volume. To navigate this vacuum at the international level. Taddeo, McNeish, Blanchard, and Edgar discuss in Chapter 7 the efforts to define ethical frameworks to guide the use of AI in the defence domain at the domestic level – through the case of the United Kingdom – and propose a possible framework, articulated around five principles: justified and overridable uses; just and transparent systems and processes; human moral responsibility; meaningful human control; and, finally, reliable AI systems. At the core of these ethical considerations are the matter of technological autonomy and the need for some form of human control, involvement, or override: again, where does the human fit in 'the loop'? Going back to the international level, in Chapter 8 Louis Perez navigates the different discussion streams at the UN on Cyber on the one hand and LAWS on the other, before discussing how the current approach to LAWS could also be applied to autonomous cyber operations. Reflecting on the definition of LAWS, this chapter addresses the vital question of whether autonomous cyber capabilities could be considered LAWS and thus be concerned by the discussions on international law and ethics taking place in the framework of the CCW. In Chapter 9, Jack Kenny focuses on a specific principle of international law, the principle of non-intervention, that has been discussed extensively by the GGE and the OEWG. Building on these discussions, as well as on the existing scholarship on the application of this principle in cyberspace, this chapter looks at the specific challenges raised by automation for this principle with a specific focus on its coercion requirement. By going back to one of the operational dilemmas discussed earlier, the chapter elucidates this normative discussion through the analysis of different examples related to the interference in electoral processes using cyber means with a certain degree of autonomy.

The third and last part of the volume shows that the debates at the UN level have a while to go before they will be able to meaningfully address the intersection between AI technology and conflict in cyberspace. There are several reasons for that, which are related to the technical/operational and strategic/geopolitical perspectives outlined earlier. For one thing, most of the richer and top-tier (cyber) military states are often reluctant to forego new military possibilities that may turn out to be game changers.[40] Countries like the United States, Israel and Russia, which are actively developing LAWS are dragging their feet in the GGE negotiations. History does not provide much evidence of weapons being banned before they are used. Also, politically the level of trust between some of the main negotiating parties is at a low point at this moment. The United States are increasingly in an adversarial competition with China – which is one of the main contenders for the 'AI crown' – and since the Russian invasion of Ukraine many states are actively trying to sanction and isolate the Russian Federation. These are not ideal circumstances to discuss restraint as a governance mechanism when it comes to new military and cyber technology. Lastly, there is a mandate mismatch between the two UN processes. The UN GGE on LAWS – as the name indicates – explicitly focuses on a specific technology (AI) in relation to *weapons*. The UN GGE on cybersecurity focuses on *state behaviour* as the focal point for its recommendations and usually aims to be as technology neutral as possible. If a bridge is to be built between these processes it will have to be built on sound reasoning on how technology impacts on, or changes, state behaviour in cyber conflict. Questions like whether 'state control' only exists when there is meaningful human control or also exist when in case of 'system control,' and whether automated and/or autonomous cyber-attacks are or can be (in)discriminate[41] are likely to be at the heart of that. In other words, only by understanding the relationship between AI and conflict in cyberspace as a comprehensive phenomenon, and embedded in broader geopolitical conflicts, can the international community truly move forward with meaningful regulation.

Notes

[1] See, for example, Thomas Rid, *Cyber War Will Not Take Place* (London: Hurst and Company, 2013); Erik Gartzke, "The myth of cyberwar: Bringing war in cyberspace back down to earth," *International Security* 38, no. 2 (2013): 41–73; Robert Chesney and Max Smeets, eds., *Deter, Disrupt, or Deceive. Assessing Cyber Conflict as an Intelligence Contest* (Washington, DC: Georgetown University Press, 2023).

[2] Lucas Kello, *The Virtual Weapon and International Order* (New Haven and London: Yale University Press, 2017).

[3] Patryk Pawlak, Eneken Tikk, and Mika Kerttunen, "*Cyber Conflict Uncoded*," EUISS Brief no. 7, European Union Institute for Security Studies, 7 April 2020.

[4] See Richard J. Harknett and Max Smeets, "Cyber campaigns and strategic outcomes," *Journal of Strategic Studies* 45, no. 4 (2022): 534–567; and Chesney and Smeets, *Deter, Disrupt, or Deceive*.

[5] John Perry Barlow, "A Declaration of the Independence of Cyberspace" (8 February 1996).

[6] David J. Betz and Tim Stevens, *Cyberspace and the State: Towards a Strategy for Cyber-power* (Abingdon: Routledge, 2011).

[7] See, for example: Rebecca Slayton, "What is the cyber offense-defense balance? Conceptions, causes, and assessment," *International Security* 41, no. 3 (2017): 72–109.

8 H. Lin and J. Kerr, "On cyber-enabled information warfare and information operations," in *Oxford Handbook of Cybersecurity* (Oxford: Oxford University Press, 2021).

9 Ronald Deibert and Louis W. Pauly, "Mutual entanglement and complex sovereignty in cyberspace," in *Data Politics: Worlds, Subjects, Rights* (London: Routledge, 2019); Christian Ruhl et al., *Cyberspace and Geopolitics: Assessing Global Cybersecurity Norm Processes at a Crossroads* (Washington, DC: Carnegie Endowment for International Peace, 2020); Daniel Deudney, "Turbo change: Accelerating technological disruption, planetary geopolitics, and architectonic metaphors," *International Studies Review* 20, no. 2 (2018); Nicole Perlroth, *This Is How They Tell Me the World Ends: The Cyberweapons Arms Race* (Bloomsbury: Bloomsbury Publishing, 2021).

10 See, for example: Jon R. Lindsay, *Information Technology and Military Power* (Ithaca, NY: Cornell University Press, 2020); and Avi Goldfarb and Jon R. Lindsay, "Prediction and judgment: Why artificial intelligence increases the importance of humans in war," *International Security* 46, no. 3 (2022).

11 Michael C. Horowitz, "Artificial intelligence, international competition, and the balance of power," *Texas National Security Review* 1, no. 3 (May 2018).

12 See Monica Kaminska, Dennis Broeders, and Fabio Cristiano, "Limiting viral spread: Automated cyber operations and the principles of distinction and discrimination in the grey zone," in *13th International Conference on Cyber Conflict: 'Going Viral'* (Tallinn: CCDCOE, 2021).

13 Corien Prins et al, *iGovernment* (Amsterdam: Amsterdam University Press, 2011).

14 Tim Stevens, "Knowledge in the grey zone: AI and cybersecurity," *Digital War* 1, no. 1 (2020); Finn Brunton, *Spam: A shadow history of the Internet* (MIT Press, 2013).

15 This question is also relevant for debates on attribution. On this topic, see Joseph M. Brown and Tanisha M. Fazal, "#SorryNotSorry: Why states neither confirm nor deny responsibility for cyber operations," *European Journal of International Security* 6, no. 4 (2021); for a wider legal perspective, see: R. Liivoja, M. Naagel, and A. Väljataga, *Autonomous Cyber Capabilities Under International Law* (Tallinn, Estonia: CCDCOE, 2019).

16 Noran Shafik Fouad, "The non-anthropocentric informational agents: Codes, software, and the logic of emergence in cybersecurity," *Review of International Studies* 48, no. 4 (2022); Clare Stevens, "Assembling cybersecurity: The politics and materiality of technical malware reports and the case of Stuxnet," *Contemporary Security Policy* 41, no. 1 (2020); Myriam Dunn Cavelty and Andreas Wenger, "Cyber security meets security politics: Complex technology, fragmented politics, and networked science," *Contemporary Security Policy* 41, no. 1 (2020).

17 Benjamin M. Jensen, Christopher Whyte, and Scott Cuomo, "Algorithms at war: The promise, peril, and limits of artificial intelligence," *International Studies Review* 22, no. 3 (2020); Paul Scharre, *Army of None. Autonomous Weapons and the Future of War* (New York: W.W. Norton & Company, 2018); Michael C. Horowitz, "The ethics & morality of robotic warfare: Assessing the debate over autonomous weapons," *Daedalus* 145, no. 4 (2016).

18 Stevens, "Knowledge in the grey zone," 166.

19 Lene Hansen and Helen Nissenbaum, "Digital disaster, cyber security, and the Copenhagen School," *International Studies Quarterly* 53, no. 4 (2009).

20 Heather M. Roff, "The frame problem: The AI "arms race" isn't one," *Bulletin of the Atomic Scientists* 75, no. 3 (2019).

21 Sergei Boeke and Dennis Broeders, "The demilitarisation of cyber conflict," *Survival* 60 no. 6 (2018).

22 Max Smeets, *No Shortcuts: Why States Struggle to Develop a Military Cyber-Force* (London: Hurst Publishers, 2022).

23 Nazli Choucri and David D. Clark, *International Relations in the Cyber Age: The Co-Evolution Dilemma* (Cambridge, MA: MIT Press, 2019).

24 Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust, and Fear Between Nations* (Oxford: Oxford University Press, 2016).

25 M. Mueller, "Is cybersecurity eating internet governance? Causes and consequences of alternative framings," *Digital Policy, Regulation and Governance* 19, no. 6 (2017).

26 Adam Segal, "Une guerre froide fluide: Les Etats-Unis, la Chine et la guerre technologique," *Hérodote* 2022, no. 184–185 (2022); see also, Kai Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston, MA: Houghton Mifflin, 2018).

27 Graham Webster et al., "China's plan to 'lead' in AI: Purpose, prospects, and problems," *New America Foundation* (1 August 2017).

28 Dennis Broeders, ed., *Digital Sovereignty: From Narrative to Policy?* (EU Cyber Direct, 2022); T. Christakis, *'European Digital Sovereignty': Successfully Navigating Between the 'Brussels Effect' and Europe's Quest for Strategic Autonomy* (Multidisciplinary Institute on Artificial Intelligence/Grenoble Alpes Data Institute, 2020); Benjamin Farrand and Helena Carrapico, "Digital sovereignty and taking back control: From regulatory capitalism to regulatory mercantilism in EU cybersecurity," *European Security* 31, no. 3 (2022).

29 See, for example: Lindsay, *Information Technology*; James Johnson, "The AI-cyber nexus: Implications for military escalation, deterrence and strategic stability," *Journal of Cyber Policy* 4, no. 3 (2019); Joe Burton and Simona R. Soare, "Understanding the strategic implications of the weaponization of artificial intelligence," in *11th International Conference on Cyber Conflict (CyCon): Silent Battle*, ed. T. Minárik, S. Alatalu, S. Biondi, M. Signoretti, I. Tolga and G. Visky (Tallinn: CCDCOE, 2019).

30 Martin C. Libicki, "Cyberspace is not a warfighting domain," *Journal of Law and Policy for the Information Society* 8, no. 2 (2012).

31 For an overview, see: Dennis Broeders, "The (im)possibilities of addressing election interference and the public core of the internet in the UN GGE and OEWG: A mid-process assessment," *Journal of Cyber Policy* 6, no. 3 (2021): 277–279.

32 Aude Géry and François Delerue, "A new UN path to cyber stability," *Directions* (6 October 2020); Valentin Weber, "How to strengthen the program of action for advancing responsible state behavior in cyberspace," *Just Security* (10 February 2022).

33 UNGA, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN Doc A/65/201 (30 July 2010); UNGA, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN Doc A/68/98 (24 June 2013); UNGA, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN Doc A/70/174 (22 July 2015); UNGA, *Report of the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security*, UN Doc A/76/135 (14 July 2021); UNGA, *Report of the Open-ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN Doc A/75/816 (18 March 2021).

34 Chair of the OEWG, *Chair's Summary of the OEWG on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN Doc AC.290/2021/CRP.3 (10 March 2021).

35 See Kaminska, Broeders, and Cristiano, "*Limiting Viral Spread*," 62–64.

36 CCW, *Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System*, annexed (Annex III)

to the *Final Report* of the meeting of the high contracting parties to the convention on prohibitions or restrictions on the use of certain conventional weapons which may be deemed to be excessively injurious or to have indiscriminate effects, 19 December 2019, CCW/MSP/2019/9, Annex III, 10.

[37] Ibid., principle (f).

[38] OECD, *Recommendation of the Council on Artificial Intelligence* (Paris: OECD, 2019).

[39] UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (UNESCO, 2021).

[40] See, for example: John Arquilla, *Bitskrieg. The New Challenge of Cyberwarfare* (London: Polity, 2021).

[41] Kaminska, Broeders and Cristiano, "*Limiting Viral Spread*."

Bibliography

Arquilla, John. *Bitskrieg. The New Challenge of Cyberwarfare*. London: Polity, 2021.

Barlow, John Perry. *A Declaration of the Independence of Cyberspace*. 8 February 1996. http://www.eff.org/~barlow/Declaration-Final.html.

Betz, David J., and Tim Stevens *Cyberspace and the State: Towards a Strategy for Cyber-Power*. Abingdon: Routledge, 2011.

Boeke, Sergei, and Dennis Broeders. "The demilitarisation of cyber conflict." *Survival* 60, no. 6 (2018): 73–90. https://doi.org/10.1080/00396338.2018.1542804.

Broeders, Dennis. "The (im)possibilities of addressing election interference and the public core of the internet in the UN GGE and OEWG: A mid-process assessment." *Journal of Cyber Policy* 6, no. 3 (2021): 277–297. https://doi.org/10.1080/23738871.2021.1916976.

Broeders, Dennis, ed. *Digital Sovereignty: From Narrative to Policy?* EU Cyber Direct, 2022.

Brown, Joseph M., and Tanisha M. Fazal. "# SorryNotSorry: Why states neither confirm nor deny responsibility for cyber operations." *European Journal of International Security* 6, no. 4 (2021): 401–417.

Brunton, Finn. *Spam: A Shadow History of the Internet*. MIT Press, 2013.

Buchanan, Ben. *The Cybersecurity Dilemma: Hacking, Trust, and Fear Between Nations*. Oxford: Oxford University Press, 2016.

Burton, Joe, and Simona R. Soare. "Understanding the strategic implications of the weaponization of artificial intelligence." In *2019 11th International Conference on Cyber Conflict (CyCon): Silent Battle*, edited by T. Minárik, S. Alatalu, S. Biondi, M. Signoretti, I. Tolga and G. Visky, 1–17. Tallinn: CCDCOE, 2019. https://doi.org/10.23919/CYCON.2019.8756866.

CCW. *Guiding Principles Affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System*. CCW/MSP/2019/9, 19 December 2019.

Chair of the OEWG. *Chair's Summary of the OEWG on Developments in the Field of Information and Telecommunications in the Context of International Security*. UN Doc AC.290/2021/CRP.3, 10 March 2021.

Chesney, Robert, and Max Smeets, eds. *Deter, Disrupt, or Deceive. Assessing Cyber Conflict as an Intelligence Contest*. Washington, DC: Georgetown University Press, 2023.

Choucri, Nazli, and David D. Clark. *International Relations in the Cyber Age: The Co-Evolution Dilemma*. Cambridge, MA: MIT Press, 2019.

Christakis, T. *'European Digital Sovereignty': Successfully Navigating between the 'Brussels Effect' and Europe's Quest for Strategic Autonomy*. Multidisciplinary Institute on Artificial Intelligence/Grenoble Alpes Data Institute, 2020.

Deibert, Ronald, and Louis W. Pauly. "Mutual entanglement and complex sovereignty in cyberspace." In *Data Politics: Worlds, Subjects, Rights*, edited by Didier Bigo, Engin Isin and Evelyn Ruppert, 81–88. London: Routledge, 2019.

Deudney, Daniel. "Turbo change: Accelerating technological disruption, planetary geopolitics, and architectonic metaphors." *International Studies Review* 20, no. 2 (2018): 223–231.

Dunn Cavelty, Myriam, and Andreas Wenger. "Cyber security meets security politics: Complex technology, fragmented politics, and networked science." *Contemporary Security Policy* 41, no. 1 (2020): 5–32.

Farrand, Benjamin, and Helena Carrapico. "Digital sovereignty and taking back control: From regulatory capitalism to regulatory mercantilism in EU cybersecurity." *European Security* 31, no. 3 (2022): 435–453. https://doi.org/10.1080/09662839.2022.2102896.

Fouad, Noran Shafik. "The non-anthropocentric informational agents: Codes, software, and the logic of emergence in cybersecurity." *Review of International Studies* 48, no. 4 (2022): 766–785.

Gartzke, Erik. "The myth of cyberwar: Bringing war in cyberspace back down to earth." *International Security* 38, no. 2 (2013): 41–73.

Géry, Aude, and François Delerue. "A new UN path to cyber stability." *Directions*(6 October 2020). https://directionsblog.eu/a-new-un-path-to-cyber-stability/.

Goldfarb, Avi, and Jon R. Lindsay. "Prediction and judgment: Why artificial intelligence increases the importance of humans in war." *International Security* 46, no. 3 (2022): 7–50.

Hansen, Lene, and Helen Nissenbaum. "Digital disaster, cyber security, and the Copenhagen school." *International Studies Quarterly* 53, no. 4 (2009): 1155–1175.

Harknett, Richard J., and Max Smeets. "Cyber campaigns and strategic outcomes." *Journal of Strategic Studies* 45, no. 4 (2022): 534–567.

Horowitz, Michael C. "The ethics & morality of robotic warfare: Assessing the debate over autonomous weapons." *Daedalus* 145, no. 4 (2016): 25–36. https://doi.org/10.1162/DAED_a_00409.

Horowitz, Michael C. "Artificial intelligence, international competition, and the balance of power." *Texas National Security Review* 1, no. 3 (May 2018): 36–57.

Jensen, Benjamin M., Christopher Whyte, and Scott Cuomo. "Algorithms at war: the promise, peril, and limits of artificial intelligence." *International Studies Review* 22, no. 3 (2020): 526–550.

Johnson, James. "The AI-cyber nexus: Implications for military escalation, deterrence and strategic stability." *Journal of Cyber Policy* 4, no. 3 (2019): 442–460. https://doi.org/10.1080/23738871.2019.1701693.

Kaminska, Monica, Dennis Broeders, and Fabio Cristiano. "Limiting viral spread: Automated cyber operations and the principles of distinction and discrimination in the grey zone." In *13th International Conference on Cyber Conflict: 'Going Viral,'* edited by T. Jančárková, L. Lindström, G. Visky and P. Zotz, 59–72. Tallinn: CCDCOE, 2021.

Kello, Lucas. *The Virtual Weapon and International Order.* New Haven, CT and London: Yale University Press, 2017.

Lee, Kai Fu. *AI Superpowers: China, Silicon Valley, and the New World Order.* Boston, MA: Houghton Mifflin, 2018.

Libicki, Martin C. "Cyberspace is not a warfighting domain." *Journal of Law and Policy for the Information Society* 8, no. 2 (2012): 321–336.

Liivoja, R., M. Naagel, and A. Väljataga. *Autonomous Cyber Capabilities under International Law.* Tallinn, Estonia: CCDCOE, 2019.

Lin, H., and J. Kerr. "On cyber-enabled information warfare and information operations." In *Oxford Handbook of Cybersecurity*, edited by P. Cornish, 251–272. Oxford: Oxford University Press, 2021.

Lindsay, Jon R. *Information Technology and Military Power.* Ithaca, NY: Cornell University Press, 2020.

Mueller, M. "Is cybersecurity eating internet governance? Causes and consequences of alternative framings." *Digital Policy, Regulation and Governance* 19, no. 6 (2017): 415–428. https://doi.org/10.1108/DPRG-05-2017-0025.

OECD. *Recommendation of the Council on Artificial Intelligence.* Paris: OECD, 2019.

Pawlak, Patryk, Eneken Tikk, and Mika Kerttunen. *Cyber Conflict Uncoded.* EUISS Brief no. 7, European Union Institute for Security Studies, 7 April 2020. https://www.iss.europa.eu/content/cyber-conflict-uncoded.

Perlroth, Nicole. *This is How They Tell Me the World Ends: The Cyberweapons Arms Race*. Bloomsbury: Bloomsbury Publishing, 2021.

Prins, Corien, Dennis Broeders, Henk Griffioen, Anne-Greet Keizer, and Esther Keymolen. *iGovernment.* Amsterdam: Amsterdam University Press, 2011.

Rid, Thomas. *Cyber War Will Not Take Place*. London: Hurst and Company, 2013.

Roff, Heather M. "The frame problem: The AI "arms race" isn't one." *Bulletin of the Atomic Scientists* 75, no. 3 (2019): 95–98.

Ruhl, Christian, Duncan Hollis, Wyatt Hoffman, and Tim Maurer. *Cyberspace and Geopolitics: Assessing Global Cybersecurity Norm Processes at a Crossroads*. Washington, DC: Carnegie Endowment for International Peace, 2020.

Scharre, Paul. *Army of None. Autonomous Weapons and the Future of War.* New York: W.W. Norton & Company, 2018.

Segal, Adam. "Une guerre froide fluide: Les Etats-Unis, la Chine et la guerre technologique." *Hérodote* 2022, no. 184–185 (2022): 271–284.

Slayton, Rebecca. "What is the cyber offense-defense balance? Conceptions, causes, and assessment." *International Security* 41, no. 3 (2017): 72–109. https://doi.org/10.1162/ISEC_a_00267.

Smeets, Max. *No Shortcuts: Why States Struggle to Develop a Military Cyber-Force.* London: Hurst Publishers, 2022.

Stevens, Clare. "Assembling cybersecurity: The politics and materiality of technical malware reports and the case of Stuxnet." *Contemporary Security Policy* 41, no. 1 (2020): 129–152.

Stevens, Tim. "Knowledge in the grey zone: AI and cybersecurity." *Digital War* 1, no. 1 (2020): 164–170.

UNESCO. *Recommendation on the Ethics of Artificial Intelligence.* UNESCO, 2021. https://unesdoc.unesco.org/ark:/48223/pf0000379920.

UNGA. *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security.* UN Doc A/65/201, 30 July 2010.

UNGA. *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security.* UN Doc A/68/98, 24 June 2013.

UNGA. *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security.* UN Doc A/70/174, 22 July 2015.

UNGA. *Report of the Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security.* UN Doc A/75/816, 18 March 2021.

UNGA. *Report of the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security.* UN Doc A/76/135, 14 July 2021.

Weber, Valentin. "How to strengthen the program of action for advancing responsible state behavior in cyberspace." *Just Security*, 10 February 2022. https://www.justsecurity.org/80137/how-to-strengthen-the-programme-of-action-for-advancing-responsible-state-behavior-in-cyberspace/.

Webster, Graham, Rogier Creemers, Paul Triolo, and Elsa Kania. "China's plan to 'lead' in AI: Purpose, prospects, and problems." *New America Foundation*, 1 August 2017. https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/.