INTERNATIONAL LAW DOES NOT COMPUTE: ARTIFICIAL INTELLIGENCE AND THE DEVELOPMENT, DISPLACEMENT OR DESTRUCTION OF THE GLOBAL LEGAL ORDER

MATTHIJS M MAAS*

Within the coming decade, the deployment of artificial intelligence ('AI') appears likely to have a disruptive impact on global affairs. What will such 'globally disruptive' AI imply for the form, function and viability of international law? I briefly sketch the long history of technological innovation driving, shaping and destroying international law. Drawing on scholarship on the relation between new technologies and international law, I argue that new technology changes legal situations both directly, by creating new entities or enabling new behaviour, and indirectly, by shifting incentives or values. I argue that development of increasingly more disruptive AI may produce three types of global legal impacts. The first is 'legal development' (patching); the second is 'legal displacement' (substitution); the third is 'legal destruction' (erosion). I discuss the potential impact of AI in all three modalities, and the implications for international relations. I argue that many of the challenges raised by AI could in principle be accommodated in the international law system through legal development, and that while AI may aid in compliance enforcement, the prospects for legal displacement — a shift towards an 'automated international law' — look slim. However, I also conclude that technical and political features of the technology will in practice render AI destructive to key areas of international law: the legal gaps it creates will be hard to patch, and the strategic capabilities it offers chip away at the rationales for powerful states to engage fully in, or comply with, international law regimes. This suggests some risk of obsolescence of distinct international law regimes.

CONTENTS

* PhD Fellow, Centre for International Law, Conflict and Crisis, University of Copenhagen; Research Affiliate, Centre for the Governance of AI, Future of Humanity Institute, University of Oxford. ORCID ID: 0000-0002-6170-9393. Email: matthijs.maas@jur.ku.dk. Versions of this paper were previously presented at the 2018 Copenhagen Inaugural Workshop on AI and Legal Disruption, at the 8th Annual Cambridge International Law Conference 2019, as well as in a December 2018 interview on the *Algocracy and Transhumanism Podcast*. I would like to thank the participants of these workshops, particularly Hin-Yan Liu, Léonard Van Rompaey, Luisa Scarcella and John Danaher, for their comments on early versions of this argument, and the Center for the Governance of AI for support during a research stay in Fall 2018. I would also like to thank Haydn Belfield, Linnéa Nordlander, Martina Kunz and two anonymous referees for extensive and valuable feedback. Finally, I thank Lyria Bennett Moses, Roger Brownsword and Thomas Burri for their comments and encouragement. Any remaining errors are all my own. No conflict of interest is identified.

I INTRODUCTION: 'GLOBALLY DISRUPTIVE' AI AND INTERNATIONAL LAW

In recent years, the field of artificial intelligence ('AI') has advanced at an unprecedented pace, with remarkable and rapid gains in a wide range of performance metrics.[1] In a technical sense, AI is 'that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment'.[2] However, in a practical sense, AI can be understood as a general-purpose technology, for automating and improving the accuracy, speed and/or scale of machine decision-making, pattern-recognition and prediction in complex or large environments, with the aim of substituting for, or improving upon, human performance in specific tasks.

As a general-purpose technology for domain-general decision tasks, there are few things that AI might not help us with. The technology holds large promise for improving lives and addressing global challenges, such as the achievement of the United Nations 2030 Sustainable Development Goals.[3] Indeed, the eventual historical impact of AI has drawn comparisons to past generally-enabling 'technologies', such as electricity, fire or the internal combustion engine.[4]

At the same time, however, many experts also anticipate the deployment of profoundly disruptive AI capabilities — systems which will, directly or indirectly, spark conflict, enable oppression or inflame tensions, generating diverse and far-reaching governance challenges for the world.[5] These challenges include, but are not limited to, large labour displacements and inequality,

---

[1] Most famously, the breakthroughs with the game-playing artificial intelligence ('AI') systems AlphaGo, AlphaZero and, most recently, AlphaStar: see generally David Silver et al, 'Mastering the Game of Go with Deep Neural Networks and Tree Search' (2016) 529(7587) *Nature* 484; David Silver et al, 'Mastering the Game of Go without Human Knowledge' (2017) 550(7676) *Nature* 354; The AlphaStar Team, 'AlphaStar: Mastering the Real-Time Strategy Game StarCraft II', *DeepMind* (Blog Post, 24 January 2019) <https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/>, archived at <https://perma.cc/7RS5-VRF4>. For an initiative that seeks to track global progress in AI, see Peter Eckersley and Yomna Nasser, 'Measuring the Progress of AI Research', *Electronic Frontier Foundation* (Web Page) <https://www.eff.org/ai/metrics>, archived at <https://perma.cc/5UA5-5933>; Yoav Shoham et al, *Artificial Intelligence Index: 2018 Annual Report* (Report, December 2018) 59–62.

[2] Nils J Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge University Press, 2010) xiii.

[3] *UN Secretary-General's Strategy on New Technologies* (Report, September 2018) 8.

[4] Michael C Horowitz, 'Artificial Intelligence, International Competition, and the Balance of Power' (2018) 1(3) *Texas National Security Review* 36, 39; MSNBC, 'Google CEO Sundar Pichai: AI More Important to Humanity Than Fire and Electricity' (YouTube, 29 January 2018) 00:00:00–00:01:12 <https://www.youtube.com/watch?v=jxEo3Epc43Y>; Shana Lynch, 'Andrew Ng: Why AI Is the New Electricity', *Stanford Graduate School of Business* (Blog Post, 11 March 2017) <https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity>, archived at <https://perma.cc/WT42-XCCE>.

[5] See generally Allan Dafoe, 'AI Governance: A Research Agenda' (Research Paper, Future of Humanity Institute, University of Oxford, 27 August 2018).

reinforced surveillance capabilities for authoritarian states,[6] increasingly scalable cyberwarfare capabilities,[7] an oligopolistic or mercantilist market structure dominated by a few leading AI companies or principals, disruptive shifts in the balance of national power or in the relative competitiveness of democratic and dictatorial states,[8] a fundamental revolution in the decision-making character of warfare and tactical or strategic (nuclear) instability.[9] What are the causes and effects of 'global disruption' by AI?

On the 'cause' (or input) side, just as AI is itself not one single technology but many, the class of 'globally disruptive' AI ('GDAI') denotes a broad set of AI applications and capabilities. The threshold of 'global disruption' in principle extends to future breakthroughs in AI capabilities, though it does not presume or require them. At its floor, it arguably applies to a number of currently prototyped or plausibly imminent AI systems which can be used to far-reaching political or strategic effect.[10] Near its ceiling, the class would also extend to future, more advanced systems, including hypothetical 'artificial general intelligence',[11] if or when it is developed.[12] Either way, the key criterion is that this concerns AI applications that, once (or if) deployed, require governance, legal or regulatory coordination on a global level to be adequately addressed. This does not mean that such challenges are all revolutionary: for AI to be 'globally disruptive' denotes a lower threshold (though no upper ceiling) for disruption than that invoked by comparisons to 'general-purpose technologies' such as fire or electricity, or that invoked in the definition of 'transformative AI' as AI that

---

6 Yuval Noah Harari, 'Why Technology Favors Tyranny' (October 2018) *The Atlantic* <https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/>, archived at <https://perma.cc/F7GM-3ZWC>.

7 Miles Brundage et al, *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (Report, February 2018) 52–3.

8 Horowitz, 'Artificial Intelligence, International Competition, and the Balance of Power' (n 4) 53; Richard Danzig, 'An Irresistible Force Meets a Moveable Object: The Technology Tsunami and the Liberal World Order' (2017) 5(1) *Lawfare Research Paper Series* 1; Harari (n 6); Michael C Horowitz, 'Who'll Want Artificially Intelligent Weapons? ISIS, Democracies, or Autocracies?', *Bulletin of the Atomic Scientists* (online, 29 July 2016) <https://thebulletin.org/2016/07/wholl-want-artificially-intelligent-weapons-isis-democracies-or-autocracies/>, archived at <https://perma.cc/2AUU-8SMD>. See also the expert survey by Foreign Affairs: 'Does Technology Favor Tyranny?', *Foreign Affairs* (online, 12 February 2019) <https://www.foreignaffairs.com/ask-the-experts/2019-02-12/does-technology-favor-tyranny>, archived at <https://perma.cc/5ZP7-2P93>.

9 Kenneth Payne, 'Artificial Intelligence: A Revolution in Strategic Affairs?' (2018) 60(5) *Survival* 7; Edward Geist and Andrew J Lohn, 'How Might Artificial Intelligence Affect the Risk of Nuclear War?' (Research Paper, RAND Corporation, 2018); Keir A Lieber and Daryl G Press, 'The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence' (2017) 41(4) *International Security* 9.

10 Dafoe (n 5).

11 Phil Torres, 'The Possibility and Risks of Artificial General Intelligence' (2019) 75(3) *Bulletin of the Atomic Scientists* 105, 105; Martina Kunz and Seán Ó hÉigeartaigh, 'Artificial Intelligence and Robotization' in Robin Geiss and Nils Melzer (eds), *Oxford Handbook on the International Law of Global Security* (Oxford University Press, forthcoming) 13.

12 Estimates differ widely as to when, if ever, this benchmark might be achieved. One recent survey indicated that AI researchers assign a 50% chance to AI outperforming humans in all tasks within the next 45 years: Katja Grace et al, 'When Will AI Exceed Human Performance? Evidence from AI Experts' (2018) 62 *Journal of Artificial Intelligence Research* 729, 730.

'precipitates a transition comparable to (or more significant than) the agricultural or industrial revolution'.[13]

On the 'effect' (or output) side, 'global disruption' covers a spectrum of possible impacts; certain AI applications might be disruptive in a simply 'positional' sense, involving changes to which particular states — for example, the United States or China — are dominant or exert the most power within the international system, while leaving the broader rules of that system intact.[14] However, disruption is also manifested through more general shifts in the nature or structural dynamics of (international) society. It might involve, for instance, a shift in the relative balance of power between types of actors, away from states and towards non-state actors (whether civil society or private actors). Finally, disruption may also be manifested through shifts in the manner in which actors seek to exercise power (from military force to computational propaganda) or in changes in the main terms by which actors come to conceive of, measure and pursue power or (national) interest (eg, data, global public opinion, oil reserves, population, soft power or territory).

Admittedly, this definition of global disruption may be leaky and contested: establishing an appropriate baseline against which to measure or mark whether an impact of AI has been 'disruptive' to the existing balance of power or extant societal dynamics is a critical and challenging question. Likewise, it may be difficult to draw a clear line between disruption that is unambiguously 'technological' (directly driven by AI capabilities) and more indirect disruptions that appear 'non-technological' (eg, gradual changes to public values such as the appreciation of data privacy) or at most as a side effect of the use of AI.[15] The existing literature on the regulation of technology has emphasised this difficulty of pinning down a workable definition of 'technology' or drawing a clear line to distinguish 'technology-driven change' from more general societal change. However, for this article, I draw on a broad concept of 'sociotechnical change', which, as articulated by Lyria Bennett Moses, is focused not purely on artefacts but also on behaviour — on the new things, entities and actions that a technology makes possible.[16] To paraphrase a witticism from computer science: all concepts are wrong; some concepts are useful.[17]

---

[13] Holden Karnofsky, 'Potential Risks from Advanced Artificial Intelligence: The Philanthropic Opportunity', *Open Philanthropy Project* (Blog Post, 6 May 2016) <http://www.openphilanthropy.org/blog/potential-risks-advanced-artificial-intelligence-philanthropic-opportunity>, archived at <https://perma.cc/P5F4-WGUN>. For a discussion of AI as a 'General Purpose Technology', see generally Manuel Trajtenberg, 'AI as the Next GPT: A Political-Economy Perspective' (Working Paper No 24245, National Bureau of Economic Research, January 2018). Note that the definition of 'globally disruptive' AI used here is closer to the definition of 'Transformative AI' discussed by Allan Dafoe, who referred to it as 'advanced AI that could lead to radical changes in welfare, wealth, or power': Dafoe (n 5) 8.

[14] On such positional changes, see generally Stephen G Brooks and William C Wohlforth, 'The Rise and Fall of the Great Powers in the Twenty-First Century: China's Rise and the Fate of America's Global Position' (2016) 40(3) *International Security* 7.

[15] I thank one anonymous referee for these points.

[16] Lyria Bennett Moses, 'Regulating in the Face of Sociotechnical Change' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation, and Technology* (Oxford University Press, 2017) 573, 576.

[17] George EP Box, 'Science and Statistics' (1976) 71(356) *Journal of the American Statistical Association* 791, 792.

Whichever labels we choose to use, and however we draw the exact threshold for global disruption, the rise of AI is clearly set to lead to urgent new questions for global governance. Equally important, however, this revolution also raises new challenges for the tools of such governance, including international law. There has been an established body of literature which examines how new technology in general can affect or disrupt law.[18] Others have identified the particular policy and regulatory challenges posed by AI,[19] as well as the relevant institutional competencies of the common law tort system, expert agencies or national legislatures in mitigating risks.[20] At a global level, there is an incipient field of study examining potential global governance approaches to these diverse AI challenges. Of course, much of this work to date has focused on the issue of lethal autonomous weapon systems;[21] although some recent work has extended these discussions into examining a broader range of prospective military AI systems.[22] Recent years have also seen the promulgation of diverse sets of 'AI ethics' and guidelines by a wide range of actors and states.[23] Others have begun

---

[18] See generally David Friedman, 'Does Technology Require New Law?' (2001) 25(1) *Harvard Journal of Law and Public Policy* 71; Lyria Bennett Moses, 'Recurring Dilemmas: The Law's Race to Keep up with Technological Change' [2007] *University of New South Wales Faculty of Law Research Series* 21 ('Recurring Dilemmas'); Roger Brownsword, 'Law and Technology: Two Modes of Disruption, Three Legal Mind-Sets, and the Big Picture of Regulatory Responsibilities' (2018) 14(1) *Indian Journal of Law and Technology* 1 ('Law and Technology').

[19] Ryan Calo, 'Artificial Intelligence Policy: A Primer and Roadmap' (2018) 3(2) *University of Bologna Law Review* 180, 190–213. See generally Michael Guihot, Anne F Matthew and Nicolas P Suzor, 'Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence' (2017) 20(2) *Vanderbilt Journal of Entertainment and Technology Law* 385.

[20] Matthew U Scherer, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies' (2016) 29(2) *Harvard Journal of Law and Technology* 353, 376–92 ('Regulating Artificial Intelligence Systems').

[21] See, eg, Human Rights Watch and International Human Rights Clinic, *Losing Humanity: The Case against Killer Robots* (Report, 2015) <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots#>, archived at <https://perma.cc/G2FE-NVUP>; Myriam Dunn Cavelty, Sophie-Charlotte Fischer and Thierry Balzacq, '"Killer Robots" and Preventive Arms Control' in Myriam Dunn Cavelty and Thierry Balzacq (eds), *Routledge Handbook of Security Studies* (Routledge, 2nd ed, 2017) 457. For an analysis that seeks to generalise this into an overarching model of 'preventive security governance' for new military technologies, see Denise Garcia, 'Future Arms, Technologies, and International Law: Preventive Security Governance' (2016) 1(1) *European Journal of International Security* 94.

[22] Matthijs M Maas, 'How Viable Is International Arms Control for Military Artificial Intelligence? Three Lessons from Nuclear Weapons' (2019) 40(3) *Contemporary Security Policy* 285 ('How Viable Is International Arms Control for Military Artificial Intelligence?'); Matthijs M Maas, 'Innovation-Proof Global Governance for Military Artificial Intelligence? How I Learned to Stop Worrying, and Love the Bot' (2019) 10(1) *Journal of International Humanitarian Legal Studies* 129.

[23] For a brief study, see Yi Zeng, Enmeng Lu and Cunqing Huangfu, 'Linking Artificial Intelligence Principles' (Conference Paper, AAAI Workshop on Artificial Intelligence Safety, 12 December 2018) <https://arxiv.org/ftp/arxiv/papers/1812/1812.04814.pdf>, archived at <https://perma.cc/M2EM-2RQR>. For a discussion of the limits of and tensions between ethical principles, see Jess Whittlestone et al, 'The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions' (Conference Paper, AAAI/ACM Conference on AI, Ethics and Society, 27–28 January 2019) <https://dl.acm.org/citation.cfm?id=3314289>, archived at <https://perma.cc/9BHS-WZ3F>. For a discussion of general macro-strategic 'policy desiderata' for the transition towards more capable AI, see Nick Bostrom, Allan Dafoe and Carrick Flynn, 'Policy Desiderata for Superintelligent AI: A Vector Field Approach' in S Matthew Liao (ed), *Ethics of Artificial Intelligence* (Oxford University Press, forthcoming).

to concretely explore possible avenues or arrangements by which international law might effectively govern the challenges posed by AI systems in the near term,[24] as well as eventual systemic risks posed by advanced AI systems in the longer term.[25]

However, with a few exceptions,[26] there has been less work examining how AI will affect international law. Yet, prior to bringing international law to bear on regulating GDAI systems, it is pertinent to ask how AI will affect the integrity, viability, form and relevance of that international law system itself. How will applications of this technology affect specific domains within international law or the general scaffolding of this order? Will these challenges be trivial (on reflection requiring minimal or no change), tractable (able to be accommodated through some legal change) or terminal (beyond the scope of the instruments or concepts of the existing legal order)?

In what follows, I will not seek to provide any definite answers but instead sketch what I hope is an exploratory framework for exploring these questions. In Part II, I provide a brief background on the longstanding history of technological innovation driving change in international law. In Part III, I provide the body of the argument, assessing three types of legal impacts which AI systems may have on international law. The first is 'legal development' (patching), the second is 'legal displacement' (substitution) and the third is 'legal destruction' (erosion). In unfolding this argument, I will draw on a number of analytical frameworks which have previously been developed in the context of scholarship on the interaction between technology and law.

II TECHNOLOGICAL CHANGE AND INTERNATIONAL LAW: AN ABRIDGED HISTORY

The relative lack of attention to the ways in which AI might affect international law is curious, since, as noted by Colin Picker, technological innovations have driven the 'creation, modification, or destruction of international law' throughout history.[27] Even in pre-modern societies, the technologies that enabled the agricultural revolution created a need for exclusive

---

[24] Olivia J Erdélyi and Judy Goldsmith, 'Regulating Artificial Intelligence: Proposal for a Global Solution' (Conference Paper, AAAI/ACM Conference on AI, Ethics, and Society, 2–3 February 2018) <http://www.aies-conference.com/2018/contents/papers/main/AIES_2018_paper_13.pdf>, archived at <https://perma.cc/UB5L-5DWG>; Kunz and Ó hÉigeartaigh (n 11); Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Palgrave Macmillan, 2019) 237–54.

[25] JG Castel and Matthew E Castel, 'The Road to Artificial Super-Intelligence: Has International Law a Role to Play?' (2016) 14(1) *Canadian Journal of Law and Technology* 1; Grant Wilson, 'Minimizing Global Catastrophic and Existential Risks from Emerging Technologies through International Law' (2013) 31(2) *Virginia Environmental Law Journal* 307, 332–3; Reinmar Nindler, 'The United Nation's Capability to Manage Existential Risks with a Focus on Artificial Intelligence' (2019) 21(1) *International Community Law Review* 5, 15–20; Luke Kemp et al, 'UN High-Level Panel on Digital Cooperation: A Proposal for International AI Governance', Submission to the UN Secretary-General's High-level Panel on Digital Cooperation, 26 February 2019, 2–3 <https://digitalcooperation.org/wp-content/uploads/2019/02/Luke_Kemp_Submission-to-the-UN-High-Level-Panel-on-Digital-Cooperation-2019-Kemp-et-al.pdf>, archived at <https://perma.cc/5RPN-KSNP>.

[26] See, eg, Thomas Burri, 'International Law and Artificial Intelligence' (2017) 60 *German Yearbook of International Law* 91.

[27] Colin B Picker, 'A View from 40,000 Feet: International Law and the Invisible Hand of Technology' (2001) 23(1) *Cardozo Law Review* 149, 156.

and persistent control of land, which led to the first (tacit) expression of concepts of sovereignty and diplomatic relations.[28] Advances in ship and navigation technologies during the 16th and 17th centuries increased coastal states' ability to project military force and exclude rivals from trade routes, which offered the impetus for Hugo Grotius' seminal articulation of the principle of *mare liberum* (freedom of the seas) in response, considered by some to be the founding moment for international law.[29]

The development and proliferation of new military technologies enabled unforeseen brutality in several major systemic wars, which motivated key developments in international law. For instance, the proliferation and use of gunpowder weaponry in the Thirty Years' War contributed to the states party's willingness to forestall a repeat of that war's brutality through the *Treaty of Westphalia* in 1648.[30] In the 20th century, the technology-enhanced nightmares of trench warfare during the First World War and the nuclear shadow of the Second World War contributed to the establishment of the Permanent Court of International Justice and the UN, respectively.[31] In these cases, the new technology-enabled (threat of) unprecedented bloodshed generated at least part of the critical political momentum necessary for landmark legal innovations aimed at mitigating avenues for future turbulence and conflict, or at containing its horror.[32]

Specific technologies can also upset 'narrow' legal regimes. At the 1907 Second Hague Conference, the *Convention Relative to the Creation of an International Prize Court* was proposed to hear cases regarding the capturing of prizes (material, mostly ships) during wartime.[33] As it was, the initiative soon proved abortive: its political prospects were never promising, and by 1912, it was clear the proposed Prize Court was dead in the water as it became evident that major naval powers would refuse to ratify the key 1909 *Declaration concerning the Laws of Naval War*.[34] Nonetheless, the *Convention* marked history's first treaty proposal for a truly international court. Yet, even had it been politically successful, the rapid technological change at the time would have plausibly made its mission and role moot before long. Specifically, the eventual advent of submarine warfare saw a shift in the practical nature of naval warfare that would

---

28 Ibid 158.

29 Ibid 160–3.

30 Ibid 163.

31 Ibid 163–4 n 37.

32 Ibid 163–4. For a survey of the rationales motivating various historical arms control efforts (examined in light of a possible autonomous weapons ban), see also Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (WW Norton, 2018) 331–45 ('*Army of None*').

33 *Convention for the Establishment of an International Prize Court*, signed 18 October 1907, 205 CTS 381 (not in force). See generally Henry B Brown, 'The Proposed International Prize Court' (1908) 2(3) *American Journal of International Law* 476.

34 *Final Protocol of the Naval Conference*, signed 26 February 1909, 208 CTS 338 (not in force). On the stillbirth of the Permanent Court of International Justice, see Manley O Hudson, *The Permanent Court of International Justice 1920–1942: A Treatise* (Macmillan, 1943) 78; Eugene Kontorovich, 'Three International Courts and Their Constitutional Problems' (2014) 99(6) *Cornell Law Review* 1353, 1380. See generally James Brown Scott, 'The International Court of Prize' (1911) 5(2) *American Journal of International Law* 302, 314; James L Tryon, 'The International Prize Court and Code' (1911) 20(8) *Yale Law Journal* 604.

have challenged or superseded many of the Court's key operational assumptions.[35] In contrast with surface navy vessels, submarines relied on surprise and stealth; they found it hard to clearly or reliably distinguish merchant from navy vessels while submerged and were vulnerable to small arms fire if they surfaced to issue warnings or request right of visit and search.[36] Moreover, submarines at any rate did not have the spare space to carry enemy prize crew back to port.[37] Indeed, the tactical and operational demands of the technological shift towards submarine warfare would also prove a thorn in the side of subsequent tools of international law aimed at regulating naval warfare. For example, Rebecca Crootof has discussed how the 1930 *Limitation and Reduction of Naval Armament (London Naval Treaty)* and the 1936 *Limitation of Naval Armament (Second London Naval Treaty)* — which held that submarines were not distinct from surface warships, and as such were also bound by the prohibition against attacking enemy merchant vessels without first ensuring the safety of their sailors — became functionally eroded by a subsequently developed customary international law of submarine warfare,[38] as the abovementioned tactical and logistical constraints on submarine operation came to be far more determinative of actual practices than were the extant treaty instruments nominally in force. In fact, Crootof argues that this customary norm, which matured in the Second World War, effectively expanded states' rights with regard to the lawful use of submarines, rendering the older treaties 'dead letter'.[39] She describes this as an illustration of a broader pattern whereby technological developments continuously (or iteratively) challenge or bypass existing governance approaches, leaving in their wake 'jurisprudential space junk':[40] collections of fragmented, hard-to-amend treaty regimes which 'are theoretically in force but actually simply clutter and confuse the relevant legal regime'.[41]

In the post-war era, technological developments have played decisive roles in the initial formulation and course of development of diverse areas of international law, from the law of the sea to the international law of fisheries, and from non-proliferation regimes for modern weapons of mass destruction to space law.[42] Moreover, in the last three decades, advances in communication technologies — and specifically the internet — have significantly enhanced information exchange, altering the institutional and logistical landscape of international law formation and enforcement. For one, the greater dissemination of information enabled by the internet can make it easier for states to learn about

---

35 Janet Marilyn Manson, 'International Law, German Submarines and American Policy' (MA Thesis, Portland State University, 1977) 7–9.

36 Ibid 9.

37 Ibid 8.

38 *Limitation and Reduction of Naval Armament (London Naval Treaty)*, signed 22 April 1930, 112 LNTS 65 (entered into force 31 December 1930); *Limitation of Naval Armament (Second London Naval Treaty)*, signed 25 March 1936, 197 LNTS 387 (entered into force 29 July 1937); Rebecca Crootof, 'Jurisprudential Space Junk: Treaties and New Technologies' in Chiara Giorgetti and Natalie Klein (eds), *Resolving Conflicts in the Law: Essays in Honour of Lea Brilmayer* (Brill Nijhoff, 2019) 106, 113–14.

39 Crootof (n 38) 114.

40 Ibid 107.

41 Ibid.

42 Picker (n 27) 164–78.

the practices and laws of others, facilitating legal harmonisation. Moreover, it renders it easier for states party to locate and apply evidence for customary international law.[43]

On a process level, communication technologies have also shaped the creation of international law. For instance, travel and communications technologies have made treaty negotiation 'faster and easier',[44] as state representatives can receive more rapid guidance and approval; a shift noted by Louis B Sohn in 1973, when he wrote:

> Now information can be transmitted quickly in both directions, and the process of final approval can be expedited. … Thus, new application of technology and science lead not only to new rules but also to new methods of creating new rules.[45]

Moreover, these communications technologies have not just sped up the development of international law but have also changed its character, as they enable more states and institutions — particularly non-governmental organisations — to mobilise for and participate in the negotiation process of treaties,[46] such as the *Anti-Personnel Mine Ban Convention*[47] or the *Paris Agreement*.[48] On the other hand, if the internet's apparent transparency of information has resulted in an increased democratisation of international law formation processes, its demonstrated opacity of identity also gives it a reverse face. The ability of states to interfere in each other's affairs through deniable cyberattacks and digital (mis)information campaigns (whether by trolls or bots) is one technology-mediated strategy that has been used to considerable political effect[49] and which risks hollowing out cornerstone international law concepts such as 'state sovereignty' or 'state aggression'.

All of this is not to suggest that the effect of technological development on international legal concepts or instruments is always (or even ever) unidirectional, straightforward or irresistible. Nonetheless, given the historical track record of technologies serving as one vector in seminal legal change, what does all this spell for the effect of transformative AI on international law?

---

43 Ibid 198.

44 Crootof (n 38) 107.

45 Louis B Sohn, 'The Impact of Technological Changes on International Law' (1973) 30(1) *Washington and Lee Law Review* 1, 10.

46 Picker (n 27) 199.

47 *Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on Their Destruction*, opened for signature 3 December 1997, 2056 UNTS 211 (entered into force 1 March 1999) ('*Anti-Personnel Mine Ban Convention*').

48 *Paris Agreement*, opened for signature 22 April 2016, [2016] ATS 24 (entered into force 4 November 2016).

49 See, eg, Philip N Howard and Bence Kollanyi, 'Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK–EU Referendum' (Research Note No 2016.1, COMPROP, 2016); Mark Leiser, 'AstroTurfing, "CyberTurfing" and Other Online Persuasion Campaigns' (2016) 7(1) *European Journal of Law and Technology* 1. For a discussion of the interrelation of information and cyber-warfare with international law, see also ME Bowman, 'Is International Law Ready for the Information Age?' (1996) 19(5) *Fordham International Law Journal* 1935; Christopher C Joyner and Catherine Lotrionte, 'Information Warfare as International Coercion: Elements of a Legal Framework' (2001) 12(5) *European Journal of International Law* 825; Sean Kanuck, 'Sovereign Discourse on Cyber Conflict under International Law' (2010) 88(7) *Texas Law Review* 1571.

III AI AND INTERNATIONAL LAW: THREE EFFECTS

AI may change the international legal situation both directly and indirectly. Directly, it generates new legal situations by creating new legal entities or by enabling new behaviour. Indirectly, AI may shift the incentives or values for states interacting with international law. Out of this, we may distinguish three types of legal impacts effected by any sufficiently disruptive technology such as AI. The first is legal development (change of elements leading to a need for legal change to accommodate or address the new situation), the second is legal displacement (systemic substitution of regulatory modality; the 'automation' of international law) and the third is legal destruction (systemic disruption of key premises; erosion). I will examine all three in turn to understand the conditions under which a technology such as AI might produce manageable (non-disruptive) development or change, where it might be vulnerable to legal displacement and where it might lead to international legal destruction.

A *AI and Legal Development*

Bennett Moses has argued in favour of a theory of 'law and technological change'.[50] She argues that while not every technology creates the occasion or need for new litigation or legal scholarship, technological change does often create a recurring dilemma for law by creating new entities or enabling new behaviour.[51] In her analysis, this creates four distinct types of new legal situations which call for legal development: (1) a need for new, special laws; (2) legal uncertainty; (3) incorrect scope (under- or over-inclusiveness of laws); and (4) legal obsolescence.[52] Whilst originally developed in the context of domestic legal systems, Bennett Moses' model nonetheless offers valuable lessons for understanding the space of legal problems generated or provoked by new technologies more broadly. Of course, there are important operational differences between national and international law; yet the insights from her model are not dependent on the specific legal toolset in question (whether domestic or international), but rather describe the features of the legal problems that are to be regulated — situations which pose problems for any normative system of law.[53] I will therefore discuss these four types of legal situations in turn and offer some preliminary examples of how these might apply in the context of AI systems.

1 *The Need for New Laws*

The technology creates a straightforward need for new sui generis rules to deal with new situations or forms of conduct, or to ban a particular technology or particular applications. AI may enable new, morally problematic or politically or strategically disruptive forms of conduct — say, the systematic monitoring and control of populations through enhanced surveillance; the deployment of fully autonomous weapons or (cyber)warfare systems operationally susceptible to

---

50 Lyria Bennett Moses, 'Why Have a Theory of Law and Technological Change?' (2007) 8(2) *Minnesota Journal of Law, Science and Technology* 589, 590, 605–6.

51 Bennett Moses, 'Recurring Dilemmas' (n 18) 4–5.

52 Ibid 8.

53 I thank Bennett Moses for flagging this question.

emergent accidents;[54] or the tracking of rival nuclear assets in ways that threaten deterrence stability.[55] Such conduct may be considered hazardous and undesirable by most or all of the states involved, which therefore creates both a need and the conditions for new treaties to explicitly ban or control the development, deployment or use of these systems. In an international legal context, this might echo past (multilateral or bilateral) arms control efforts, such as the 1968 *Treaty on the Non-Proliferation of Nuclear Weapons*[56] or the 1972 *Treaty on the Limitation of Anti-Ballistic Missile Systems*.[57] While the creation of new technology-specific treaty regimes to address the gaps has not always been politically easy, the international legal system is, at least in principle, clearly capable of proposing and promulgating new legal regimes in order to address the gaps opened up by new technologies — even those as potentially transformative as AI.

2 *Legal Uncertainty*

The technology creates uncertainty regarding how the existing law applies to the new forms of conduct. This includes uncertainty over how a new activity, entity or relationship will be classified, either because no adequate classification exists, because it fits into more than one existing classification and becomes subject to different, conflicting rules or because an existing category becomes ambiguous in light of the new forms of conduct.[58] This may lead to a need for existing legal rules to be clarified or sharpened.[59] Matthew Scherer has argued that the autonomy, opacity and unpredictability of certain AI systems might create uncertainty over concepts such as attribution, control and responsibility.[60] On the other hand, Thomas Burri has argued that the case law of international

---

54 Paul Scharre, 'Flash War: Autonomous Weapons and Strategic Stability' (Conference Presentation, The Weaponization of Increasingly Autonomous Technologies: Understanding Different Types of Risks, 11 April 2016) 16, 18–19 <http://www.unidir.ch/files/conferences/pdfs/-en-1-1113.pdf>, archived at <https://perma.cc/YUJ2-FY3U>; Richard Danzig, Center for a New American Security, *Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority* (Report, June 2018) 40 <https://s3.amazonaws.com/files.cnas.org/documents/CNASReport-Technology-Roulette-DoSproof2v2.pdf?mtime=20180628072101>, archived at <https://perma.cc/ULE6-YT7A>; Matthijs M Maas, 'Regulating for "Normal AI Accidents": Operational Lessons for the Responsible Governance of AI Deployment' (Conference Paper, AAAI/ACM Conference on AI, Ethics and Society, 2–3 February 2018) <http://www.aies-conference.com/wp-content/papers/main/AIES_2018_paper_118.pdf>, archived at <https://perma.cc/C5DA-PDL6>.

55 Lieber and Press (n 9).

56 *Treaty on the Non-Proliferation of Nuclear Weapons*, opened for signature 1 July 1968, 729 UNTS 161 (entered into force 5 March 1970).

57 *Treaty on the Limitation of Anti-Ballistic Missile Systems*, United States of America–Soviet Union, signed 26 May 1972, 944 UNTS 13 (entered into force 3 October 1972) ('*Anti-Ballistic Missile Treaty*'). Cf Emanuel Adler, 'The Emergence of Cooperation: National Epistemic Communities and the International Evolution of the Idea of Nuclear Arms Control' (1992) 46(1) *International Organization* 101. For a recent application of the historical lessons from these cases and for designing arms control regimes for present or future military AI systems, see also Maas, 'How Viable Is International Arms Control for Military Artificial Intelligence?' (n 22).

58 Bennett Moses, 'Recurring Dilemmas' (n 18) 26.

59 Ibid 16.

60 Scherer, 'Regulating Artificial Intelligence Systems' (n 20) 363–6.

courts (such as the International Tribunal for the Former Yugoslavia or the International Court of Justice ('ICJ')) includes more than sufficient precedent for resolving questions of state control, attribution and the limits of delegation.[61] Even if such judicial clarification were not available, new laws, treaties or customary international law could, again, plug the gaps to provide the necessary conceptual clarification around AI systems, thus accommodating these changes within the international legal system.

3 *Incorrect Scope*

The new technology creates new contexts that lead to inappropriate over-inclusiveness and under-inclusiveness of existing laws. Previously unproblematic laws are suddenly found to have an inappropriate scope. For instance, some have argued — on purely legal rather than ethical or philosophical grounds — that it may already, today, be possible to grant certain algorithms some semblance of personhood. For instance, Shawn Bayern has argued that loopholes in existing US company law might already allow for the incorporation of a limited liability company ('LLC') whose operating agreement places it under operational control of an AI system.[62] After having every other (human) member withdraw, the LLC would be left with an algorithm wholly and solely in charge, functionally establishing artificially intelligent entities with legal personhood.[63] While others have argued that courts would not interpret the relevant statutes in this way, as this outcome would be considered contrary to legislative intention,[64] Bayern and others have since sought to extend their argument to the legal systems of Germany, Switzerland and the United Kingdom.[65] Burri has argued that, if such an entity were to be established in a European Union member state, the internal market principle of the mutual recognition of national legal personality — as established in the European Court of Justice rulings in *Centros Ltd v Erhvervs-og Selskabsstyrelsen* and *Überseering BV v Nordic Construction Company Baumanagement GmbH*[66] — would mean that this entity would have to be recognised by all EU member states.[67] Such legal ploys or exploits to establish AI personhood, if indeed viable,

---

61 Burri, 'International Law and Artificial Intelligence' (n 26) 101–3, 108.

62 Shawn Bayern, 'The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems' (2016) 7(2) *European Journal of Risk Regulation* 297, 300–4.

63 Ibid 302.

64 Matt Scherer, 'Is AI Personhood Already Possible under US LLC Laws? (Part One: New York)', *Law and AI* (Blog Post, 14 May 2017) <http://www.lawandai.com/2017/05/14/is-ai-personhood-already-possible-under-current-u-s-laws-dont-count-on-it-part-one/>, archived at <https://perma.cc/5DNW-X2TY>. Cf Turner (n 24) 177.

65 Shawn Bayern et al, 'Company Law and Autonomous Systems: A Blueprint for Lawyers, Entrepreneurs, and Regulators' (2017) 9(2) *Hastings Science and Technology Law Journal* 135, 139–53, 160.

66 *Centros Ltd v Erhvervs-og Selskabsstyrelsen* (C-212/97) [1999] ECR I-1459, I-1497; *Überseering BV v Nordic Construction Company Baumanagement GmbH (NCC)* (C-208/00) [2002] ECR I-9943, I-9975–6, cited in Burri, 'International Law and Artificial Intelligence' (n 26) 96. See generally Wulf-Henning Roth, 'From *Centros* to *Ueberseering*: Free Movement of Companies, Private International Law, and Community Law' (2003) 52(1) *International and Comparative Law Quarterly* 177.

67 Thomas Burri, 'Free Movement of Algorithms: Artificially Intelligent Persons Conquer the European Union's Internal Market' in Woodrow Barfield and Ugo Pagallo (eds), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar, 2018) 537, 545.

would create potential for criminal abuse,[68] and arguably create an inappropriate over-inclusiveness of existing laws,[69] such that this gap should be rapidly patched through legal review or legislation.

4 *Legal Obsolescence*

The new technology leads to the obsolescence of particular laws, because the law is no longer needed, justified, or cost-effective to enforce. First, technology can provoke legal obsolescence because conduct that was regulated by the existing laws has itself been rendered obsolete by new technologies, such that the law is no longer needed.[70] To understand such 'obsolescence through obscurity', one might consider laws governing communication by postal pigeon or by telegraph: whilst the precedent they set may certainly still provide key legal metaphors that shape the path of subsequent legal interpretation, especially in common law contexts,[71] the laws themselves are no longer invoked. In principle, they might be struck from the books, and no citizen would even notice. Might international law ever be rendered obsolete in this manner? If, for the sake of argument, AI-steered combat platforms were to fully replace human soldiers on the battlefield, this might render moot certain *jus in bello*/international humanitarian law ('IHL') principles dictating the treatment of prisoners of war. Would such a development be a problem? That depends on whether one takes a view of IHL as being narrowly meant 'to limit the effects of armed conflict'[72] on non-combatants. If so, one can argue that the removal of potential prisoners of war from the theatres of war has successfully achieved at least this sub-goal of IHL. Of course, IHL is hardly limited to the treatment of prisoners of war alone, and a more refined IHL perspective that also considers the effects of these AI combat systems on other non-combatants — such as local civilians — might still hold much force, making those parts of IHL more resilient to legal obsolescence.

Secondly, technology could lead to legal obsolescence because a rule can no longer be justified. For instance, the human right to work, as enshrined in the *Universal Declaration of Human Rights* and the *International Covenant on Economic, Social and Cultural Rights*,[73] appears premised on the notion that a society can provide productive employment opportunities for all its members. If, as some argue, AI systems will be able to surpass human beings in more and

---

68 Lynn M LoPucki, 'Algorithmic Entities' (2018) 95(4) *Washington University Law Review* 887, 890. On the overarching challenges that AI-supported crime may pose for existing (criminal) liability models, and specifically the actus reus or mens rea standards, see also Thomas C King et al, 'Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions' (2019) *Science and Engineering Ethics* (forthcoming) 7–9.

69 See Burri, 'International Law and Artificial Intelligence' (n 26) 95–8.

70 Bennett Moses, 'Recurring Dilemmas' (n 18) 46–8.

71 Cf Gregory N Mandel, 'Legal Evolution in Response to Technological Change' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation, and Technology* (Oxford University Press, 2017) 225, 233–4.

72 ICRC, *What Is International Humanitarian Law?* (Fact Sheet, July 2004) <https://www.icrc.org/en/doc/assets/files/other/what_is_ihl.pdf>, archived at <https://perma.cc/QN5N-V3X5>.

73 *Universal Declaration of Human Rights*, GA Res 217A (III), UN GAOR, UN Doc A/810 (10 December 1948) art 23(1); *International Covenant on Economic, Social and Cultural Rights*, opened for signature 16 December 1966, 993 UNTS 3 (entered into force 3 January 1976) art 6(1).

more skills,[74] simply rendering large swathes of jobs obsolete and a large population permanently and hopelessly unemployable,[75] that cornerstone assumption (admittedly always an aspirational rather than an enforceable right) might have to be substantially revised or reconsidered. More concretely, the changes in the employment market might render large parts of the legal regime built up by the International Labour Organization (hundreds of treaties, as well as extensive soft law) moot.

Thirdly, technology might lead to legal obsolescence because a rule is no longer cost-effective to enforce. To take one (perhaps slightly quixotic) example: while AI could help in the general monitoring, modelling and prediction of human rights abuses or episodes of mass violence,[76] AI-produced 'DeepFakes'[77] techniques could also enable the at-scale forging of video documentation — say, of presumed human rights abuses. Indeed, scholars have raised the possibility that such DeepFakes may adversely affect the probative value of (video) evidence — not just in domestic courts, but also in the sense that it could erode the epistemological foundation of international journalism, human rights investigations or judicial proceedings.[78] The use of DeepFakes could enable malicious (or politically interested) actors to swamp human rights bodies or media observers with fake reports which could only be revealed as such after laborious analysis. Such strategies would simultaneously erode the credibility of genuine reports and impose heavy costs on these organisations, rendering effective monitoring and enforcement of human rights norms more difficult yet.

These situations demonstrate how AI might drive legal change by creating new entities, enabling new behaviour, or by shifting incentives of actors. Bennett Moses and others argue that these four cases — the need for new laws, legal uncertainty, incorrect scope and legal obsolescence — demonstrate how technology can create an intractable recurring dilemma;[79] a problem against which legal systems cannot be future-proofed.[80] Nonetheless, the above discussion also shows that in many cases, abstract legal notions are flexible enough to be invoked, reinterpreted or adapted to these new situations: in that sense, legal development is almost always possible in principle. Discussing the

---

74 Grace et al (n 12).

75 Carl Benedikt Frey and Michael A Osborne, 'The Future of Employment: How Susceptible Are Jobs to Computerisation?' (2017) 114 *Technological Forecasting and Social Change* 254, 265–9. See Erin Winick, 'Every Study We Could Find on What Automation Will Do to Jobs, in One Chart', *MIT Technology Review* (online, 25 January 2018) <https://www.technologyreview.com/s/610005/every-study-we-could-find-on-what-automation-will-do-to-jobs-in-one-chart/>, archived at <https://perma.cc/9XQ7-AEGJ>. See generally McKinsey Global Institute, *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation* (Report, December 2017).

76 Cf Part III(A)(2).

77 See generally Bobby Chesney and Danielle Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107 *California Law Review* (forthcoming).

78 Marie-Helen Maras and Alex Alexandrou, 'Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos' (2019) 23(3) *International Journal of Evidence and Proof* 255, 258. Steven Livingston and Mathias Risse, 'The Future Impact of Artificial Intelligence on Humans and Human Rights' (2019) 33(2) *Ethics and International Affairs* 141, 144.

79 Bennett Moses, 'Recurring Dilemmas' (n 18).

80 Friedman (n 18) 85.

international law context, Picker echoes Bennet Moses' emphasis on legal development, arguing that technological innovations may drive change in international law by causing states to:

> (1) agree to modify their behavior (usually through the device of a treaty); (2) abandon previously agreed behavior (abandonment of treaties); (3) abandon the effort to agree on new behavior (abandonment of treaty formation); (4) engage in new practices that eventually are accepted by the global community (creating customary international law); (5) abandon previously widely accepted customs (abandonment of customary international law); or (6) accept peremptory obligations (creating *ius cogens*).[81]

Many of the above-discussed scenarios of AI, while at times challenging to international law as it exists, might arguably be well-accommodated within that system through routine legal development by new customary international law, peremptory obligations or treaties. The pertinent question then becomes whether such legal development to patch the 'holes' in international law is actually politically feasible. I will turn to that question later in Part III(C), but I will first discuss the prospects of AI *displacing* international law.

B *AI and Legal Displacement*

So far, I have discussed how distinct uses of AI systems could drive legal change and the development of international law. Before I turn to the separate question of whether such development might be politically feasible (or whether AI systems might instead frustrate and erode such developments), an additional question can be asked: to what extent might AI systems be used to 'displace' or 'substitute for' international law? One can consider two subcategories of legal displacement by AI: first, legal automation — using AI in the production and enforcement of 'normative' international law; and, secondly, legal replacement — using AI to facilitate a shift in the fundamental 'regulatory modality' of international law by enabling the non-normative 'technological management' or 'nudging' of state behaviour.

1 *The Automation of International Law*

In the first, more modest case; can we automate international law? Could AI be used in the production or enforcement of 'normative' international law? The latter case of enforcement seems straightforward: for many years, technologies such as signals intelligence, satellites, arrays of (seismic, hydroacoustic, infrasound and radionuclide) monitoring stations and — aided by arrangements such as the 1992 *Treaty on Open Skies* — surveillance or radionuclide 'sniffer' aircraft have all played (stated and unstated) roles in enabling states party to monitor and verify each other's (non)compliance with treaty commitments or

---

[81] Picker (n 27) 156.

peremptory norms under international law.[82] Moreover, as argued by Steven Livingston and Mathias Risse, 21st-century digital technology — particularly the proliferation of mobile phones, commercial high-resolution imaging satellites and social media — has already begun to enable near-ubiquitous surveillance, not just by states and corporations, but also by non-state actors, such as human rights observers, journalists and open-source (citizen) investigation networks such as Bellingcat or the Syrian Archive.[83] Such analyses have already played a role in war crime investigations; in 2017, digital sleuths located Libyan execution sites by triangulating geographical features in execution videos posted to social media to detailed satellite photographs, resulting in a 2017 International Criminal Court indictment for the arrest of Mahmoud Mustafa Busayf Al-Werfalli, a Libyan warlord.[84] While much of this analysis is presently done by humans, substantial portions of it could be automated or enhanced by machine learning. Berenice Boutin has likewise suggested that AI's ability to access, collect and analyse large swathes of data could play a similar role in strengthening our ability to monitor states' compliance with international law, by detecting (or even predicting) violations.[85] For instance, she refers to the examples of Protection Assistant for Wildlife Security ('PAWS') — a system that utilises machine learning to model and predict the activities of poachers[86] — as well as the use of the 'Sentry' system, which provides Syrian civilians with early warning of incoming airstrikes.[87] Developer Hala Systems suggests that the use of Sentry through 2018 has resulted in an estimated 20–27% reduction in

---

[82] *Treaty on Open Skies*, opened for signature 24 March 1992, [2002] CTS No 3 (entered into force 1 January 2002). See generally David A Koplow, 'Back to the Future and up to the Sky: Legal Implications of "Open Skies" Inspection for Arms Control' (1991) 79(2) *California Law Review* 421. Another example is found in the International Monitoring System (IMS) sensor network, which is currently being developed and operated by the Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization: 'Overview of the Verification Regime', *CTBTO* (Web Page) <https://www.ctbto.org/verification-regime/background/overview-of-the-verification-regime/>, archived at <https://perma.cc/YS24-2553>.

[83] Livingston and Risse (n 78) 143; Steven Livingston and Sushma Raman, 'Human Rights Documentation in Limited Access Areas: The Use of Technology in War Crimes and Human Rights Abuse Investigations' (Discussion Paper No 2018-003, Carr Center for Human Rights Policy, Harvard Kennedy School, May 2018). On the 'multi-use' of remote sensing satellites, see also Nathan Edward Clark, 'Blurred Lines: Multi-Use Dynamics for Satellite Remote Sensing' (2019) 10(1) *Journal of International Humanitarian Legal Studies* 171.

[84] Livingston and Risse (n 78) 143, citing Bellingcat Investigation Team, 'How a Werfalli Execution Site Was Geolocated', *Bellingcat* (online, 3 October 2017) <https://www.bellingcat.com/news/mena/2017/10/03/how-an-execution-site-was-geolocated/>, archived at <https://perma.cc/X6U4-A7AM>.

[85] Berenice Boutin, 'Technologies for International Law & International Law for Technologies', *Groningen Journal of International Law* (Blog Post, 22 October 2018) <https://grojil.org/2018/10/22/technologies-for-international-law-international-law-for-technologies/>, archived at <https://perma.cc/4HZD-8FVD>.

[86] Ibid. See also Fei Fang et al, 'Deploying PAWS: Field Optimization of the Protection Assistant for Wildlife Security' (Conference Paper, AAAI Conference on Innovative Applications of Artificial Intelligence, 12–17 February 2016) <http://teamcore.usc.edu/papers/2016/IAAI16_PAWS.pdf>, archived at <https://perma.cc/9ZW3-6QSG>.

[87] Boutin (n 85), citing Danny Gold, 'Saving Lives with Tech Amid Syria's Endless Civil War', *Wired* (online, 16 August 2018) <https://www.wired.com/story/syria-civil-war-hala-sentry/>, archived at <https://perma.cc/FYN5-F62K>.

casualty rates in bombarded Syrian cities.[88] The use of AI in 'massively parallel DNA sequencing' has increased the scientific accuracy and efficiency of forensic investigations, potentially allowing human rights groups to uncover evidence of war crimes otherwise unavailable.[89] Miles Brundage has similarly argued that AI can play a role in monitoring compliance with treaties in areas such as arms control, cybercrime or environmental remediation.[90] Moreover, through monitoring and anomaly detection, AI could play a role in the consolidation of an understanding of state practices that underpin customary international law. While the scaling up of such capabilities is not guaranteed to happen by itself, such use cases seem to be straightforward and compelling examples of ways in which AI might not so much displace international law, but rather strengthen it by helping to address or at least mitigate longstanding problems and by improving its practical functioning and force.

More controversially, could AI systems also be embedded into the processes creating international law? As noted previously, the rise of communication technologies and the internet affected not just the substance of international law, but also its institutional, informational and logistical dynamics, yielding both greater democratisation as well as new challenges to sovereignty.[91] Likewise, unilateral usage of AI systems by various actors will certainly come to play a role in changing the character of diplomacy and international negotiations in the coming decades. For example, in summer 2018, the Ministry of Foreign Affairs of the People's Republic of China reportedly started using an AI system as a strategic decision support system, providing ranges of options and risk assessment assistance to its diplomats.[92] However, such 'unilateral' applications of AI may have some limitations: to accurately model the reactions or negotiation strategies of other states, these AI systems must be trained on relevant data. Such data will often be messy, unstructured and unclear, thus impeding meaningful analysis. Even where the data stream is clear and structured (eg, a rival state official's Twitter feed), there is the risk of that rival spoofing the data channel in order to engage in 'data poisoning', throw off the AI system and affect negotiations in their favour. However, even if such applications were viable, and even if other states and parties would allow themselves to be 'nudged' in such a manner, that would arguably not be international law, but rather a new form of unilateral 'soft power' exertion in order to shape the creation of international law.[93]

---

[88] Dan Henebery, Hala Systems, *Protect Everything That Matters* (Report, 2018) 2.

[89] Livingston and Risse (n 78) 144. See generally Edo D'Agaro, 'Artificial Intelligence Used in Genome Analysis Studies' (2018) 2(2) *EuroBiotech Journal* 78.

[90] Thomas Metzinger et al, 'Should We Fear Artificial Intelligence?' (EPRS Paper No PE 614.547, Scientific Foresight Unit (STOA), March 2018) 14–15.

[91] Picker (n 27) 198.

[92] Stephen Chen, 'Artificial Intelligence, Immune to Fear or Favour, Is Helping to Make China's Foreign Policy', *South China Morning Post* (online, 30 July 2018) <https://www.scmp.com/news/china/society/article/2157223/artificial-intelligence-immune-fear-or-favour-helping-make-chinas>, archived at <https://perma.cc/4MRG-SLAG>; Didi Tang, 'Chinese Diplomats to Get a Helping Hand from Prototype AI Robots', *The Times* (online, 30 July 2018) <https://www.thetimes.co.uk/article/chinese-diplomats-to-get-a-helping-hand-from-prototype-ai-robots-mjr3g8bvr>, archived at <https://perma.cc/L9A5-6V97>.

[93] See also the 'hard' argument for legal destruction: see below Part III(C)(2).

So, is it instead conceivable to use AI to strengthen the production or adjudication processes of international law? It is important to be clear about what this does and does not mean. There has been much discussion, in recent years, on the phenomenon of 'legal automation' in domestic legal practice.[94] In diverse sectors, from contract to administrative law, and from tax to criminal law, AI systems have begun automating routine jobs of the legal trade, and they have begun to outperform legal experts or judges at predicting legally relevant information, either the outcome of cases or matters such as re-offense rates.[95] This has led some scholars to project the image of an increasing automation of the legal systems,[96] with legal rules and standards becoming progressively replaced by algorithmically-tailored 'micro-directives' that can predict, *ex ante*, what an *ex post* judicial decision would have held in every specific case.[97] In some readings, this process may spark a slow but steady development towards a presumptive, eventual 'legal singularity' — a state wherein the laws of a specific area such as tax law are 'completely specified' in an unfathomably complex but queriable model that is in equilibrium with the desires of society.[98]

But if such 'legal automation' is even possible for domestic legal systems, could it ever make the leap to the international law sphere? Indeed, in some supranational contexts, using AI to predict rulings does appear possible. This is evident in the case of an AI algorithm which, after being trained on 584 cases tried by the European Court of Human Rights, managed to predict the outcome of new cases with 79% accuracy.[99] Nonetheless, it appears unlikely that AI will

---

94 See, eg, Joshua P Davis, 'Law without Mind: AI, Ethics, and Jurisprudence' (2018) 55(1) *California Western Law Review* 165; Anthony J Casey and Anthony Niblett, 'Self-Driving Laws' (2016) 66(4) *University of Toronto Law Journal* 429. For a discussion of some of the implications of this for landmark theories of law, see generally Brian Sheppard, 'Warming up to Inscrutability: How Technology Could Challenge Our Concept of Law' (2018) 68 (Supplement 1) *University of Toronto Law Journal* 36.

95 On predicting re-offense rates, see, eg, Jon Kleinberg et al, 'Human Decisions and Machine Predictions' (2018) 133(1) *Quarterly Journal of Economics* 237. However, the accuracy of such algorithms has received notable scrutiny, with many studies demonstrating how they are often biased along racial lines, see generally Richard Berk et al, 'Fairness in Criminal Justice Risk Assessments: The State of the Art', (2019) 48 *Sociological Methods and Research* (forthcoming); Julia Dressel and Hany Farid, 'The Accuracy, Fairness, and Limits of Predicting Recidivism' (2018) 4(1) *Science Advances* eaao5580:1–5. For an excellent discussion on possible sources of bias throughout the design and training process of a machine learning system, see David Lehr and Paul Ohm, 'Playing with the Data: What Legal Scholars Should Learn about Machine Learning' (2017) 51(2) *UC Davis Law Review* 653. For a discussion of three competing notions of 'fairness' in these debates, and how these are often incompatible with one another, see Jon Kleinberg, Sendhil Mullainathan and Manish Raghavan, 'Inherent Trade-Offs in the Fair Determination of Risk Scores' (Conference Paper, Innovations in Theoretical Computer Science, 2017).

96 Benjamin Alarie, Anthony Niblett and Albert H Yoon, 'Law in the Future' (2016) 66(4) *University of Toronto Law Journal* 423, 424.

97 Casey and Niblett, 'Self-Driving Laws' (n 94) 430; Anthony J Casey and Anthony Niblett, 'The Death of Rules and Standards' (2017) 92(4) *Indiana Law Journal* 1401, 1410–11.

98 Benjamin Alarie, 'The Path of the Law: Towards Legal Singularity' (2016) 66(4) *University of Toronto Law Journal* 443, 451–5. Cf Sheppard (n 94) 62.

99 Nikolaos Aletras et al, 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective' (24 October 2016) *PeerJ Computer Science* 2. Another study achieved an average accuracy of 75% in predicting the violation of nine articles of the European Convention on Human Rights: Masha Medvedeva, Michel Vols and Martijn Wieling, 'Judicial Decisions of the European Court of Human Rights: Looking into the Crystal Ball' (Conference Paper, Conference on Empirical Legal Studies in Europe, 31 May – 1 June 2018).

displace humans in many aspects of the production and adjudication of international law. As Burri has argued, domestic legal areas such as national tax law are susceptible to legal automation because the case law in these areas is 'extensive, dense, homogeneous, and structured' and has defined variables.[100] By contrast, the sources to which international law refers — such as (broadly worded) conventions, customary international law and decisions of the ICJ[101] — constitute datasets that are too small, fuzzy and heterogeneous for any AI algorithm (short of hypothetical, future, near-human level systems) to work with.[102] Moreover, the use of such systems to generate evidence — let alone adjudicate — on politically contested subjects is likely to be highly curtailed, at best. This is somewhat problematic for advocates of AI, as international law is hardly short on politically contested subjects. At the same time, care should be taken not to cast legal automation in too binary terms. Even if fully automated legislation or adjudication seems unlikely for the reasons enumerated above, it would seem an even greater stretch to pretend that international law has absolutely nothing to be gained from the use of AI. Instead, it seems plausible that there are subsets of tasks suited to AI within the treaty-making and adjudication processes of international law, where such automation can lead to functional improvements and therefore some limited measure of legal displacement (or, in another framing, legal 'upgrading').[103]

2 *The Technological Replacement of International Law*

So far, I have discussed the 'automation' of international law. However, even if AI were to be used to change the 'input' of international law (eg, the process of treaty negotiation or adjudication), or strengthen its enforcement, this would not yet change the nature of the 'output' of those processes (ie normative rules). 'Legal automation' therefore would not change the core, normative regulatory modality of international law. It would not change the manner in which this system seeks to regulate and change state behaviour. I therefore turn to the second, more speculative case of legal displacement — that is, 'replacement', involving a shift in the 'regulatory modality' of international law, away from normative rules, towards the technologically-mediated 'management' or 'nudging' of state behaviour, by rendering certain types of undesired behaviour technologically difficult or even impossible.

This pertains to debates over technologically-driven shifts in 'regulatory modalities': Roger Brownsword has argued that new technologies introduce a 'double disruption' to law.[104] First, they affect the substance of legal rules.[105] Secondly, they drive a regulatory shift, away from seeking to shape behaviour by normative codes or laws and towards the use of non-normative 'technological

---

100 Burri, 'International Law and Artificial Intelligence' (n 26) 93

101 *Charter of the United Nations* arts 38(1)(a)–(b), (d).

102 Burri, 'International Law and Artificial Intelligence' (n 26) 93–5.

103 I thank one anonymous referee for this observation.

104 Brownsword, 'Law and Technology' (n 18) 6–15.

105 See above Part III(A).

management' as a new possible and dominant regulatory modality.[106] This is not the same as mere use of technology for surveillance (whether centralised or through decentralised social credit), which would be to use technology to monitor 'regulatee' compliance with the normative laws. In contrast, technological management involves the use of technology to change the choice 'architecture' of regulatees in such a way as to render undesired behaviour functionally impossible (or very difficult), rather than illegal.[107] It should be made clear that 'non-normative', here, does not mean 'objective' or somehow removed from social or political goals. Rather, it draws on an operational distinction. Laws or social norms are 'normative' — they involve an explicit appeal to the regulatee to follow a certain norm — whereas systems of 'technological management' (or, less dramatically, nudge architectures) are 'non-normative' in so far as they no longer explicitly invoke or present a social norm with which the regulatee should align their behaviour. Rather, such systems simply present a (technologically) shaped environment which renders certain choice options impossible (or at least invisible). At a domestic level, this raises the spectre of AI systems (and other information technologies) being deployed in order to 'hypernudge'[108] citizens into what has been called an 'algocracy'.[109]

The ethical and political ramifications of such systems aside, what does this connote for the very concept of international law as law? In contrast to Bennett Moses' earlier point about technological change resulting in the 'obsolescence' of specific laws, the practice of 'technological management' would entail a structural obsolescence (by displacement) of normative law in general, and it might even challenge the way that societies have theorised the essential features, forms and foundations of legality and of laws.[110] Within a domestic legal context, such shifts to technological management may turn out to be fundamental and drive a shift away from traditional legal 'coherentist' accounts of law (which focus on the internal consistency and integrity of legal doctrine) towards 'regulatory-instrumentalist' or 'technocratic' approaches (which involve considerations of instrumentally effective responses).[111]

What about an international law context? Could a similar shift occur at the international level, with normative international law experiencing structural obsolescence as it is displaced by (AI-mediated) technological management? The idea that certain technologies might compel or render impossible behaviour even amongst states is not entirely impossible — compare the possible use of blockchain ledgers as a 'trustless' technology that would bypass the need for trusted third-parties, and which would preclude tampering even by (most)

---

106 Roger Brownsword, 'In the Year 2061: From Law to Technological Management' (2015) 7(1) *Law, Innovation and Technology* 1; Roger Brownsword, 'Technological Management and the Rule of Law' (2016) 8(1) *Law, Innovation and Technology* 100, 106–38.

107 Brownsword, 'Law and Technology' (n 18) 11.

108 See generally Karen Yeung, '"Hypernudge": Big Data as a Mode of Regulation by Design' (2017) 20(1) *Information, Communication and Society* 118.

109 John Danaher, 'The Threat of Algocracy: Reality, Resistance and Accommodation' (2016) 29(3) *Philosophy and Technology* 245, 246–7.

110 Sheppard (n 94).

111 Brownsword, 'Law and Technology' (n 18) 15–23.

states.[112] Hypothetically, one could entertain the idea of scaling up AI cyberwarfare systems into some type of independent 'AI leviathan', charged with tracking, tracing and isolating or disabling any computer systems engaged in large-scale cyberattack against a state. However, given the obvious operational and political challenges of such a fraught experiment, it is unclear if 'technological management' could scale. States are not citizens, and it seems hard to conceive of embedded choice architectures that would render something truly technologically impossible for states (although surveillance and intelligence technologies might well render many things impossible for states to do *undetected*). It is also unclear when, if ever, such an explicit and irrevocable forfeiture of sovereignty might be politically acceptable to most states. Of course, this might be the case if certain states or stakeholders in international law (perceive that they) are set to lose ground because of the proliferation of certain AI capabilities and if it also might be rational to forfeit certain privileges or power in order to secure their position against further disruption.[113]

The above discussion does not mean that AI cannot informally and indirectly change the dynamics of international law creation for the better. The ability of AI systems to identify novel, hidden patterns and to predict emerging challenges or trends can help aggregate evidence of state practice to speed up the creation or accumulation of customary international law. At any rate, such systems could well be used to inform 'smarter' global governance approaches to complex problems in areas where international interests are broadly aligned.[114] Even if it does not directly automate or apply to international judicial decision-making, AI could still strengthen international law across many regimes or areas of implementation and enable it to better achieve the goals of maintaining peace, protecting human rights and resolving interstate conflict. At the problem-formulation stage, for instance, applications of AI can help in improving our understanding of the problems to be addressed, including underappreciated second-order aspects of existing challenges (such as the link between regional climate change and civil conflict).[115] They can help improve the effectiveness, speed and quality of international negotiations by evening out the diplomatic playing field for small states,[116] or even by enabling more direct citizen participation or input in the debates and processes of global governance.

In sum, AI may only drive limited processes of legal displacement; its clearest use may be in strengthening the monitoring and enforcement of international law, in supporting some aspects of adjudication and in guiding more informed policymaking or governance approaches more generally. In these limited (but important) ways, AI may in fact strengthen the international law system, not weaken it.

---

112 'Why Is Blockchain Trustless?', *Lisk* (Web Page, 2019) <https://lisk.io/academy/blockchain-basics/benefits-of-blockchain/why-is-blockchain-trustless>, archived at <https://perma.cc/S7ND-4NGR>.

113 I thank one anonymous referee for this point.

114 Boutin (n 85).

115 Pernilla Nordqvist and Florian Krampe, 'Climate Change and Violent Conflict: Sparse Evidence from South Asia and South East Asia' (Research Paper No 2018/4, SIPRI Insights on Peace and Security, September 2018).

116 Katharina E Höne, *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Report, January 2019) 28.

C *AI and Legal Destruction*

A popular adage in Silicon Valley is 'Amara's Law' (after US futurologist Roy Amara), which holds that '[w]e tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run'.[117] It is easy to overstate the impact of a technology, whether on global society or on its ordering principles.

Nonetheless, while the discussion on legal development in Part III(A) suggested that a number of the disruptions (new entities or behaviours; shifted incentives) created by AI could in principle be accommodated into international law through legal development, in this part I will argue that AI may, in practice, prove to be a technology resistant to such accommodation and that it may even contribute, indirectly or directly, to the erosion of certain operating conditions for international law as a system.

There is a soft and a hard version of the argument pertaining to legal destruction.

1 *Legal Erosion: AI as Intractable Puzzle for International Law*

The 'soft' argument holds that AI systems combine a number of features that ensure they (and the changes they create) will not be easily handled, and that there may be insufficient international political agreement or leeway to carry out at least some of the important international legal developments or 'patches' specified above,[118] functionally making it an intractable challenge or puzzle to international law. This is not to suggest that AI systems would be the only or first technology to threaten such non-incorporation. For instance, Michael Glennon has argued that both the laws of war and international regulation of weapons are premised on the notion of attributability of attacks, and that in so far as cyber weapons render it easy to sidestep easy, reliable or timely identification, it is not surprising that there has been relatively little meaningful international law regulation of this technology.[119]

Likewise, Picker has identified a number of common issues that arise at the interface of technological change and international law.[120] One problem is timing: international law can be very slow to accrete and react to challenges. Under the best of circumstances, customary international law may be created in as little as 10 to 15 years.[121] While treaty processes are somewhat faster,[122] both are still extremely slow compared to technologies that can rapidly bypass the status quo, forcing a re-evaluation of treaties.

However, while failure to be proactive is problematic, the problem of stepping into a regime before it is clear how a technology works also creates problems.[123] Examples are found in the early legal scholarship during the 1950s and '60s on

---

117 Roy Amara, 'Roy Amara 1925–2007: American Futurologist' in Susan Ratcliffe (ed), *Oxford Essential Quotations* (Oxford University Press, 6th ed, 2018).

118 See above Part III(A).

119 Michael J Glennon, 'The Dark Future of International Cybersecurity Regulation' (2013) 6(2) *Journal of National Security Law and Policy* 563, 564.

120 Picker (n 27) 182–203.

121 Ibid 185.

122 Ibid.

123 Ibid 186.

regulating what was expected to be the imminent technological capabilities of weather control,[124] or the deep seabed mining provisions of the *United Nations Convention on the Law of the Sea*, which ultimately may have jumped the gun on the imminence and maturity of such seafloor mineral exploitation technologies.[125] The problem with AI is that, rather than a single technology, its performance in diverse applications can depend on parallel developments in computer hardware, data availability and (various schools of) software, rendering reliable, trusted predictions difficult.

Indeed, fast-moving emerging technologies — such as AI, nanotechnology and synthetic biology — create their own, unique sets of challenges, which potentially render them politically contentious to international law regimes (and perhaps to many other governance strategies in general). For one, many of these technologies, far from being narrow and domain-specific, offer breadth and potential power. Their very versatility introduces pervasive uncertainty as to the future course of their development and use. This makes it hard for different states to accurately project and anticipate the social impacts of the technology. Indeed, this is exacerbated by the fact that rather than a single 'technological trajectory', these technologies often comprise a host of diverse clusters in schools of development and application, which can advance at different rates (creating uneven, stop-start advancement in different applications), or suddenly 'unlock' new capabilities in and for each other.[126] The convergence of multiple technologies (that may speed up one another's development, or affect one another's modes of application), or even simply their simultaneous development and introduction, render regulatory oversight difficult at a national level,[127] and treaty negotiation perilous at an international one.

In addition to these general factors of technological change and features specific to general 'transformational' technologies, there are features specific to AI technology that further complicate its effective regulation under international (and national) law. As Scherer has argued, AI research and development processes are distinguished by problematic features, namely, that it is '*Discreet*, *Diffuse*, *Discrete*, and *Opaque*'.[128] It will be useful to take these in turn.

First, AI development is often *discreet*, as little physical infrastructure is required. AI projects can be developed without the mass institutional frameworks that were necessary for building industrial capacity in the last century, and without the signature ingredients — such as uranium or certain chemical agents — necessary for building strategically pivotal weapons of mass destruction. Of course, this difference should not be overstated: cutting-edge AI research and

---

124 Edith Brown Weiss, 'International Responses to Weather Modification' (1975) 29(3) *International Organization* 805.

125 Picker (n 27) 195. See generally *United Nations Convention on the Law of the Sea*, opened for signature 10 December 1982, 1833 UNTS 3 (entered into force 16 November 1994).

126 Kenneth W Abbott, 'Introduction: The Challenges of Oversight for Emerging Technologies' in Gary E Marchant, Kenneth W Abbott and Braden Allenby (eds), *Innovative Governance Models for Emerging Technologies* (Edward Elgar, 2013) 1, 1–2, 5–7.

127 Ibid 2–3.

128 Scherer, 'Regulating Artificial Intelligence Systems' (n 20) 369 (emphasis added).

applications still require large amounts of computational power or hardware.[129] For instance, recent years have seen extremely large increases in the level of compute used in the top AI training runs — an exponential increase with a three and a half month doubling time (compared to an 18-month doubling time for Moore's Law),[130] resulting in a staggering 300,000-time increase between 2012 and 2018.[131] On the other hand, the amount of compute necessary for deploying already-trained systems is far lower than that used in the training; not all competitive or useful AI capabilities will come with such prohibitively restrictive hardware requirements;[132] the cost of computing continues to steadily fall; and, at any rate, the sheer ubiquity of computing hardware in modern society, and its dual-use nature, renders it far more difficult to restrict.[133]

Secondly, AI development is *discrete*, as separate components may be designed in a decentralised manner, without top-down coordination, with the full potential not becoming apparent until they are brought together in a new application — although this may not be the case for certain (especially military) AI applications.[134]

Thirdly, AI development is *diffuse*, as software development can be geographically and organisationally dispersed and may involve actors in diverse and potentially unknown jurisdictions — as is often the case with open-source software.[135]

Fourthly, AI development is *opaque*, as the technologies are not well understood by regulators, and outsiders or inspectors cannot reliably detect

---

129 See generally Tim Hwang, 'Computational Power and the Social Impact of Artificial Intelligence' (Research Paper, Massachusetts Institute of Technology, 23 March 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3147971>, archived at <https://perma.cc/W5S6-4YVX>.

130 Gordon E Moore, 'Cramming More Components onto Integrated Circuits' (1965) 38(8) *Electronics* 114 <https://newsroom.intel.com/wp-content/uploads/sites/11/2018/05/moores-law-electronics.pdf>, archived at <https://perma.cc/J5NN-BUSC>. Note that in his initial prediction, Moore predicted an annual doubling, which he revised to a two-year doubling period in 1975: Gordon E Moore, 'Progress in Digital Integrated Electronics' (1975) *IEEE Technical Digest* 13, reproduced in Gordon E Moore, 'Progress in Digital Integrated Electronics' (September 2006) *IEEE SSCS Newsletter*. The 18-month prediction was made by Intel executive, David House, by considering not just the number of transistors but also improvements in transistor speed: see Michael Kanellos, 'Moore's Law to Roll on for Another Decade', *CNET* (online, 11 February 2003) <https://www.cnet.com/news/moores-law-to-roll-on-for-another-decade/>, archived at <https://perma.cc/94H8-WAEP>. In spite of continuous warnings of impending limits, the 18-month doubling-time version has more or less held: see generally Hassan N Khan, David A Hounshell and Erica RH Fuchs, 'Science and Research Policy at the End of Moore's Law' (2018) 1(1) *Nature Electronics* 14.

131 Dario Amodei and Danny Hernandez, 'AI and Compute', *OpenAI* (Blog Post, 16 May 2018) <https://blog.openai.com/ai-and-compute/>, archived at <https://perma.cc/8BRU-RRFK>.

132 For instance, the 2016 'ALPHA' AI developed by the University of Cincinnati proved able to defeat expert United States tacticians in simulated aerial combat, using no more processing power than that afforded by a small, $60 'Raspberry Pi' computer: Nicholas Ernest et al, 'Genetic Fuzzy Based Artificial Intelligence for Unmanned Combat Aerial Vehicle Control in Simulated Air Combat Missions' (2016) 6(1) *Journal of Defense Management* 1000144:1–7, 4; MB Reilly, 'Beyond Video Games: New Artificial Intelligence Beats Tactical Experts in Combat Simulation' (27 June 2016) *University of Cincinnati Magazine*.

133 Cf Scherer, 'Regulating Artificial Intelligence Systems' (n 20) 369.

134 Ibid.

135 Ibid.

features in an AI system under development.[136] One can also compare the so-called 'Volkswagen problem'; a 2015 scandal involving the reveal of 'defeat device' algorithms in Volkswagen on-board vehicle computers. These algorithms detected when a car was undergoing official tests and accordingly altered the engine's nitrogen oxide emissions to feign compliance with environmental standards.[137] Similar modules could be installed in military or cyberwarfare AI systems, rendering effective and credible verification of arms control treaties precarious.

Moreover, AI regulation is complicated by the difficulty of fixing a single definition of what AI actually is.[138] The opacity of AI systems, their heterogeneity in architectures and the deniability of their use in certain contexts (eg, cyberwar) may impede the formation of customary international law on these uses, in a similar way as has been observed (or rather has not been observed) with 'conventional' cyber weapons or intelligence operations.[139]

In contrast, relying on treaties or negotiations may not suffice either, since at a political level, AI technology is set to be particularly contested. Picker has noted that one recurring challenge to international law involves 'unevenly shared technology'.[140] This creates both epistemic problems (not all states have the requisite technological know-how to understand what regulation is needed, or even to appreciate that it is needed) as well as political problems (states can have different stakes and interests).[141] As it is, leading AI systems are set to be developed by or within a relatively small subset of states; the extent to which it can proliferate further is unclear and may sensitively depend on the degree to which performance is differentially determined and constrained by hardware or software.[142] The sheer range of AI applications renders it hard for states party to come to an agreement on the risks of a strategically appealing technology, if that technology does not have a clear 'type specimen' application, or if certain capabilities are hidden.

Moreover, software has features that have, at least in some cases, made it resistant to regulation. It has been argued, for instance, that traditional arms control regimes have not been successful in the realm of cyberspace because of several different strategic features of the technologies in question.[143] As also noted by Paul Scharre, historical non-proliferation or arms control efforts have seen the greatest success when they sought to completely outlaw a weapon,

---

136 Ibid 371.

137 Theo Leggett, 'How VW Tried to Cover up the Emissions Scandal', *BBC News* (online, 5 May 2018) <https://www.bbc.com/news/business-44005844>, archived at <https://perma.cc/5XQ8-NQ2H>.

138 Scherer, 'Regulating Artificial Intelligence Systems' (n 20) 359.

139 Glennon (n 119) 563–4.

140 Picker (n 27) 191–4.

141 Ibid 192–3.

142 Horowitz, 'Artificial Intelligence, International Competition, and the Balance of Power' (n 4) 49.

143 Erica D Borghard and Shawn W Lonergan, 'Why Are There No Cyber Arms Control Agreements?', *Council on Foreign Relations* (Blog Post, 16 January 2018) <https://www.cfr.org/blog/why-are-there-no-cyber-arms-control-agreements>, archived at <https://perma.cc/K4PB-XZQD>; Glennon (n 119) 563.

rather than set up complicated rules about how such systems may and may not be used.[144] This echoes Thomas Schelling's observation that

> the most powerful limitations, the most appealing ones, the ones most likely to be observable in wartime, are those that have a conspicuousness and simplicity, that are qualitative and not a matter of degree, that provide recognizable boundaries.[145]

Digital or virtual technologies, which are not a single 'technology', do not lend themselves so naturally to the articulation of such a simple, discrete or complete ban.

Likewise, the various non-proliferation and arms control regimes for nuclear weapons were able to invoke unambiguous red lines — an outright ban on military development or use — against horizontal proliferation. Simultaneously, they were able to utilise measurable, objective benchmarks for restraining vertical proliferation, in the form of clearly quantified and comparable caps on any of a range of metrics, such as the number of deployed missiles, the maximum number of warheads mountable on each missile, or their total throw-weight or yield.[146] By contrast, it is difficult to measure the relative strength of states in cyberspace or with AI systems.[147] Moreover, as with cyber weapons, there may be uncertainty regarding the efficacy of military AI technology, challenges with monitoring compliance and difficulties with enforcement.[148]

We can now sum up the 'soft' argument for legal destruction. Even though international law is in principle capable of the developments necessary to fix the legal problems created by AI systems,[149] in practice, these systems will likely prove highly resistant to the sources and tools available to international law. Customary international law is slow and requires clear evidence of state practice (which is not present with 'hidden' capabilities such as cyberwarfare AIs, or which is hindered by definitional problems around AI). Treaties usually require that both states party have roughly even stakes in the technology, clear expectations of benefit for abiding by the treaty, the ability to jointly agree on clear definitions and the ability to effectively verify compliance (all of which are difficult in the context of AI development). Finally, international courts are often slow, reactive to specific cases and non-expert on technologies. Effective international law regimes on new, emerging technology have historically relied on a range of ingredients, including an ability to anticipate and agree, to some extent, on the path and applications of development, an ability to agree on definitions of the technology and the ability to effectively verify compliance. Many or all of these ingredients will not apply in the context of many AI systems, suggesting that this technology will lead to disruption of that order.

This may not inhibit international law from carrying out legal developments to resolve *all* situations of legal uncertainty,[150] but it may do so in some key

---

144 Scharre, *Army of None* (n 32) 342.

145 Thomas C Schelling, *Arms and Influence* (Yale University Press, 1966) 164.

146 See, eg, *Anti-Ballistic Missile Treaty* (n 57) arts III–IV. Cf Borghard and Lonergan (n 143).

147 Borghard and Lonergan (n 143).

148 Ibid.

149 See above Part III(A).

150 Cf discussions over the 'legal personhood' of 'algorithmic entities': Bayern (n 62); Bayern et al (n 65).

cases (eg, arms control of destabilising AI systems),[151] and such situations may leave semi-permanent 'holes' in international law. While perhaps not an existential threat to international law, such holes are likely to impede effective governance and undercut the legitimacy of the global legal order in the eyes of the public.

Before proceeding to the 'hard' version of AI-driven legal destruction, it is important, briefly, to compare this version of 'soft' legal destruction with the notion of legal obsolescence covered in our earlier discussion of legal development in Part III(A). It might be argued, for instance, that there is a lot of conceptual overlap between (a) an international legal regime being rendered *obsolete* because technological change now ensures that it is no longer needed, that some of its assumptions are no longer justifiable, or that it is no longer cost-effective to enforce (legal development) and (b) a new technology having certain features that make it operationally or politically intractable to achieve adequate international regulation, ensuring legal erosion (destruction).[152] In some sense, the two indeed overlap, or can be seen as reverse faces of the same coin: for instance, it is a (perhaps *the*) basic assumption of any prospective legal regime that regulation is in fact politically and operationally possible. If that assumption is no longer justified for a new technology, then the resulting situation can be understood as legal *obsolescence* with respect to existing legal frameworks, and as legal *erosion* (or soft legal destruction) with respect to any future or extended legal frameworks which might have been considered to fill the place: the old is no longer fit to serve, but the new is out of reach. However, drawing out the conceptual distinctions between these categories remains complex, and would be a valuable future extension of this model.

2 *Legal Decline: AI as Political Threat to International Law*

Finally, there is a 'hard' version of the argument that AI will drive legal destruction. This is grounded in the idea that, especially at the international level, technological change can alter core conditions or operational assumptions, not just of specific international laws or provisions, but in the scaffolding of entire legal frameworks. This relates to a more general point: as Remco Zwetsloot and Allan Dafoe have pointed out, when we examine risks from AI, we implicitly or explicitly bucket problems as coming from either 'accident' or 'misuse'.[153] However, they argue that this dichotomy should be expanded to also take stock of a 'structural perspective'.[154] Rather than just examining how new technology can afford agents with new capabilities — that is, new opportunities for (mis)use — this perspective asks us to consider how the introduction of AI systems may unwittingly shape the environment and incentives (the 'structure') in which decision-makers operate.[155] As an example, they refer to the prominent historical

---

151 See generally Borghard and Lonergan (n 143).

152 I thank an anonymous referee for this point.

153 Remco Zwetsloot and Allan Dafoe, 'Thinking about Risks from AI: Accidents, Misuse and Structure', *Lawfare* (Blog Post, 11 February 2019) <https://www.lawfareblog.com/thinking-about-risks-ai-accidents-misuse-and-structure>, archived at <https://perma.cc/CX9L-7HL5>.

154 Ibid

155 Ibid.

interpretation of the origin of the First World War as at least partly deriving from the specific operational or logistical features of the contemporary European railroad system — features such as tight mobilisation schedules, which promoted or required rapid, all-or-nothing mass mobilisation decisions over more muted moves and which therefore, paradoxically, reduced states' manoeuvre room and pitched the dominos of general war.[156] In a like manner, certain use of AI could 'unintentionally' and structurally shift states' incentives — possibly creating overlap between offensive and defensive actions, thus driving security dilemmas; creating greater uncertainty or space for misunderstanding; or generally making the inter-state dynamic appear more like a winner-take-all dynamic — in ways that create opportunity for conflict, escalation and crisis.[157]

As such, the 'hard' argument for legal destruction holds that the deployment of AI capabilities may lead to a relative decline of the global legal system, as the capabilities afforded by these AI systems gradually shift the environment, incentives, or even values of key states. For instance, AI systems might strengthen the efficacy of more authoritarian states vis-a-vis more liberal ones,[158] accelerate the current trend towards state unilateralism, or feed into the perceived 'backlash' against international law and multilateralism. One rationale here is that whatever benefits a state believed it previously secured through engagement in, or compliance with, international law (eg, security, domestic legitimacy, soft power or cooperation), if it now perceives (whether or not correctly) that it might secure these goals unilaterally through application of AI, this may erode the broader legitimacy and regulatory capacity of international law. For instance, governments might be tempted (and, perhaps, warranted) to believe that, in the near-term future, they might be able to achieve internal security through AI surveillance capabilities, domestic legitimacy through computational propaganda (rather than through public adherence to human rights norms) or global soft power through predictive modelling of other states' negotiation strategies (rather than reciprocal engagement and compromise). Such prospects are particularly frightening given that the powerful states — on whose (at times fickle) acquiescence much of the operation of, for instance, UN bodies, might currently depend — are also leaders in developing such AI capabilities.

All this is not to say that the prospect of unilateral AI power is the only force eroding international law's multilateralist 'hardware' (institutions) or 'software' (norms), nor that it is a decisive force or even that its effects might be irresistible or irreversible. However, in so far as we are seeing an erosion of the foundations of international law, AI may speed up that decline — with all that this entails.

IV CONCLUSION

Does international law compute? How could 'globally disruptive' AI affect the institutions, instruments and concepts of the global legal order? I have discussed ways in which applications of AI may drive legal development, disruption or displacement within the system of international law.

---

156 Stephen Van Evera, 'The Cult of the Offensive and the Origins of the First World War' (1984) 9(1) *International Security* 58, 80.

157 See Zwetsloot and Dafoe (n 153).

158 Danzig (n 8) 12; Harari (n 6); Horowitz, 'Who'll Want Artificially Intelligent Weapons?' (n 8).

Specifically, I have argued that while many of the challenges raised by AI could, in principle, be accommodated in the international law system through legal development, features of the technology suggest that it will, in practice, be destructive to certain areas or instruments of international law. This ensures that there appears to be a large risk of practical erosion of certain international law structures as a result of practical and political difficulties introduced by AI systems.

The prospects for legal displacement appear more chequered. Extensive automation of the negotiation or adjudication processes of international law seems somewhat unpromising, as does substituting a technologically based system of regulating states' behaviour through non-normative behaviour control. Nonetheless, it appears plausible that more modest applications of AI may strengthen international law in areas such as monitoring, enforcement, or the development of better scientific models and a more refined evidence base to guide diverse governance initiatives.

This article has examined ways that AI challenges a state-based global legal system. However, one angle that I have under-examined is the reverse face, which concerns the possible sources of these threats — the actors (states, but also non-states) that use AI to challenge, disrupt or alter the global legal system. Such an examination would be an important complement to the picture presented here. All of this has a range of implications both for the governance of AI, as well as for general understandings of the interactions between transformative technological change and extant global legal orders.