AI, Governance Displacement, and the (De)Fragmentation of International Law

Matthijs M. Maas, University of Cambridge[*]

Draft for discussion. This version: March 17th, 2021

Submitted to conference panel under working title: *"Artificial Intelligence and the Fragmentation of the Liberal International Order: Five Vectors"*

Abstract:

The emergence, proliferation, and use of new general-purpose technologies can often produce significant political, redistributive, normative and legal effects on the world. Artificial intelligence (AI) has been identified as one such transformative technology. Many of its impacts may require global governance responses. However, what are the direct and indirect effects of AI technologies on the viability, form, or functioning of the international legal order itself? What, if any, are the prospects, peril or promise of AI-driven legal automation at the international level? This paper draws on an 'AI Governance Disruption' framework to understanding AI's impacts on the global governance architecture. Focusing particularly on the potential for legal automation at the international law level, it explores three potential pathways of such 'legal displacement': (1) the automation of rule creation and arbitration; (2) the automation of monitoring & enforcement; or (3) the 'replacement' of international law with new architectural modes of (international) behaviour control. It then focuses on the effects of these trends on the architecture of international law. It distinguishes 10 different roles that AI applications could play, with distinct effects on the international legal order. That is, AI systems can serve as (1) legal 'canary in the coal mine', highlighting the need for greater cross-regime harmonization. However, it can also serve as (2) tough knot or (3) generator of regime fault lines. Under even modest scenarios of legal automation, AI systems may serve variably as a (4) shield, (5) patch, (6) cure, or (7) accelerator of international legal fragmentation. Finally, AI tools may serve as (8) differential enabler; (9) driver of value shifts, or (10) asymmetric weapon, potentially contributing to trends of contestation or erosion in the international legal order. The paper concludes with a brief review of the ways in which international lawyers or regime scholars might approach the risks and opportunities of increasing automation in international law, in order to leverage these trends and tools towards improved efficacy, resilience, and legitimacy of global governance.

Table of Contents

---

[*] Postdoctoral Research Associate, Centre for the Study of Existential Risk & Centre for the Future of Intelligence, University of Cambridge. College Research Associate, King's College (Cambridge). Email: mmm71@cam.ac.uk. ORCID ID: 0000-0002-6170-9393

I.   Introduction

New technologies come and—often fade into the background.[1] However, they stay, and leave their marks on our behaviour, relations, and ways of being. At times, a technology's impact is broad, driving many small changes across society. At other times, its impact is deep and highly disruptive within a single industry or domain of life. A rare few technologies have an impact on the world that is both unusually broad and deep— potentially redrawing or reshaping these domains themselves. To govern a transformative technology is to reckon with questions of change: in technology; in society; and in the tools of governance itself.

Artificial Intelligence ('AI') may prove one such transformative technology.[2] This prospect may be ground for anticipation—but also for caution and reflection. This coming decade will see a growing need and opportunity to ensure that global governance is up to the task of governing the diverse challenges created by AI technology. In doing so, various governance strategies and initiatives will have to reckon with these questions of change.

Scholarship accordingly has paid extensive attention to questions of how AI can create a direct problem for the international order.[3] There are also extensive lists of governance proposals for AI's challenges, drawing on diverse instruments or bodies of international law.[4] However, with some exceptions, there is relatively little examination

---

[1] This paper draws in part on arguments developed in: Matthijs M Maas, 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks' (University of Copenhagen 2020).

[2] Ross Gruetzemacher and Jess Whittlestone, 'The Transformative Potential of Artificial Intelligence' [2020] Communications of the ACM <https://arxiv.org/abs/1912.00747>; Manuel Trajtenberg, 'AI as the next GPT: A Political-Economy Perspective' (National Bureau of Economic Research 2018) Working Paper 24245 <http://www.nber.org/papers/w24245> accessed 22 October 2018.

[3] See for instance Allan Dafoe, 'AI Governance: A Research Agenda' (Center for the Governance of AI, Future of Humanity Institute 2018) <https://www.fhi.ox.ac.uk/govaiagenda/>; Mary L Cummings and others, *Artificial Intelligence and International Affairs: Disruption Anticipated* (Chatham House 2018) <https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial-intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf>; Eleonore Pauwels, 'The New Geopolitics of Converging Risks: The UN and Prevention in the Era of AI' (United Nations University - Centre for Policy Research 2019) <https://i.unu.edu/media/cpr.unu.edu/attachment/3472/PauwelsAIGeopolitics.pdf>.

[4] See for instance Peter Cihon, Matthijs M Maas and Luke Kemp, 'Fragmentation and the Future: Investigating Architectures for International AI Governance' (2020) 11 Global Policy 545; Thorsten Jelinek, Wendell Wallach and Danil Kerimi, 'Policy Brief: The Creation of a G20 Coordinating Committee for the Governance of Artificial Intelligence' [2020] AI and Ethics <https://doi.org/10.1007/s43681-020-00019-y> accessed 30 October 2020; Eugenio V Garcia, 'Multilateralism and Artificial Intelligence: What Role for the United Nations?' in Maurizio Tinnirello (ed), *The Global Politics of Artificial Intelligence* (CRC Press 2020); Martina Kunz and Seán Ó hÉigeartaigh, 'Artificial Intelligence and Robotization' in Robin Geiss and Nils

2

of AI's effects on the legal and institutional scaffolding of the international legal order. *How can we analyse the direct and indirect effects of AI technologies on the viability, form, or functioning of the international legal order?* This paper explores this question, in order to derive both practical and theoretical implications for international law and legal scholarship.

The analysis proceeds as follows: in section II, it provides a background and theoretical underpinning to this analysis. It provides a functional definition of AI, explores why we should expect it to have wide-ranging impacts on the international legal order, and explore some theoretical frameworks for approaching the study of legal automation (LawTech), and the interrelation between technological change and legal systems more broadly (TechLaw). Section III then examines in depth the sources and vectors by which AI systems can affect international law. It sketches a 'governance disruption' framework of situations where AI applications create a need for international legal *development,* and where they can contribute to the potential *destruction* of specific regimes or the erosion of the international legal order. It then focuses in on the vectors of legal *displacement* (the potential effects of the 'automation' of international law). The section explores three potential pathways of such displacement: automation of rule creation; automation of monitoring & enforcement; or the 'replacement' of international law. Section IV then shifts the focus to the effects of these trends on the integration or fragmentation of international law, distinguishing 10 different roles that AI systems may play, with distinct effects on the global governance architecture. Section V briefly reviews some potential responses to these trends for legal scholars, and section VI concludes.

II.   Background and Rationale

What is AI? Unfortunately, there is no settled definition even amongst practitioners.[5] Scientifically, field of AI has been characterised as "making machines intelligent, [where] intelligence is that quality that enables an entity to function appropriately and with foresight in its environment."[6] Technically, however, 'AI' is an umbrella term that includes both traditional rule-based symbolic AI as well as the data-driven machine learning (ML) approaches that are responsible for the recent surge in AI progress and attention. Functionally, AI can be described as a varied suite of computational *techniques* which are able to improve the accuracy, speed, and/or scale of machine decision-making across diverse information-processing or decision-making contexts. The resulting *capabilities* can accordingly be used in order to support, substitute for-, and/or improve upon human performance in diverse tasks, enabling their deployment in *applications* across various domains, and resulting in diverse *societal impacts*.[7]

As a result, AI has in recent years come into its own as a widely applicable set of technologies, with applications in a diverse array of sectors ranging from healthcare to finance, from education to security, and even the scientific process itself.[8] While the

---

Melzer (eds), *Oxford Handbook on the International Law of Global Security* (Oxford University Press 2021) <https://papers.ssrn.com/abstract=3310421> accessed 30 January 2019; Han-Wei Liu and Ching-Fu Lin, 'Artificial Intelligence and Global Trade Governance: A Pluralist Agenda' (2020) 61 Harvard International Law Journal <https://papers.ssrn.com/abstract=3675505> accessed 26 September 2020.

[5] Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (3rd edn, Pearson 2016); Peter Stone and others, 'Artificial Intelligence and Life in 2030' (Stanford University 2016) <http://ai100.stanford.edu/2016-report> accessed 26 February 2017.

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

[7] See also Hin-Yan Liu and Matthijs M Maas, '"Solving for X?" Towards a Problem-Finding Framework to Ground Long-Term Governance Strategies for Artificial Intelligence' (2021) 126 Futures 22.

[8] See for instance Maithra Raghu and Eric Schmidt, 'A Survey of Deep Learning for Scientific Discovery' [2020] arXiv:2003.11755 [cs, stat] <http://arxiv.org/abs/2003.11755> accessed 29 June 2020. For instance, in

3

efficacy of today's AI technology is not without its problems and limits, there can be little doubt that AI's development and proliferation will impact many or all sectors of society.

What are the stakes? AI has seen a series of high-profile breakthroughs in recent years.[9] As a result, its promise of AI has hardly gone unnoticed. AI, some argue, will be 'the biggest geopolitical revolution in human history'.[10] Its prospective impact on national security has been compared to nuclear weapons, aircraft, and computing.[11] "Whoever leads in artificial intelligence in 2030," it has been claimed, "will rule the world until 2100."[12] Even under more modest readings, AI has been perceived to offer considerable economic and scientific advantages. As such, over the past years, dozens of states have articulated national AI strategies,[13] and many have begun investing vast sums in both research and application, including the development of AI systems in strategic, military

---

one study, the IBM Watson system used language processing algorithms to process thousands of peer-reviewed medical articles on the neurodegenerative disorder amyotrophic lateral sclerosis (ALS). On this basis, it correctly predicted five previously unknown genes related to the disease; Nadine Bakkar and others, 'Artificial Intelligence in Neurodegenerative Disease Research: Use of IBM Watson to Identify Additional RNA-Binding Proteins Altered in Amyotrophic Lateral Sclerosis' (2018) 135 Acta Neuropathologica 227. In another case, an unsupervised Natural Language Processing (NLP) system deployed on the scientific materials science literature, was able to capture complex materials science concepts such as the underlying structure of the periodic table, and on the basis of past literature was able to 'predict' later scientific findings by recommending new materials for application, ahead of their eventual discovery. Vahe Tshitoyan and others, 'Unsupervised Word Embeddings Capture Latent Knowledge from Materials Science Literature' (2019) 571 Nature 95.

[9] For performance and investment trends, see: Daniel Zhang and others, 'Artificial Intelligence Index Report 2020' (AI Index Steering Committee, Human-Centered AI Initiative, Stanford University 2021) <https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report_Master.pdf> accessed 3 March 2021.

[10] Kevin Drum, 'Tech World: Welcome to the Digital Revolution' [2018] *Foreign Affairs* 46.

[11] Greg Allen and Taniel Chan, 'Artificial Intelligence and National Security' (Harvard Kennedy School - Belfer Center for Science and International Affairs 2017) <http://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf> accessed 19 July 2017.

[12] Indermit Gill, 'Whoever Leads in Artificial Intelligence in 2030 Will Rule the World until 2100' (*Brookings*, 17 January 2020) <https://www.brookings.edu/blog/future-development/2020/01/17/whoever-leads-in-artificial-intelligence-in-2030-will-rule-the-world-until-2100/> accessed 22 January 2020. Note, however, that an often-repeated claim, by Russian President Vladimir Putin, that "whoever rules AI rules the world", may have been taken out of context: rather than an official statement of Russian foreign policy, this appears to have been an off-the-cuff comment which Putin made in the context of giving young Russian school children feedback on their science projects. Interview with Robert Wiblin, Keiran Harris and Allan Dafoe, 'The Academics Preparing for the Possibility That AI Will Destabilise Global Politics' (18 March 2018) <https://80000hours.org/podcast/episodes/allan-dafoe-politics-of-ai/> accessed 12 August 2020.

[13] See for overviews Jessica Cussins, 'National and International AI Strategies' (*Future of Life Institute*, February 2020) <https://futureoflife.org/national-international-ai-strategies/> accessed 22 June 2020.; Tim Dutton, Brent Barron and Gaga Boskovic, 'Building an AI World: Report on National and Regional AI Strategies' (CIFAR 2018) <https://www.cifar.ca/docs/default-source/ai-society/buildinganaiworld_eng.pdf?sfvrsn=fb18d129_4>.

or cybersecurity applications.[14] Both the US and China have established AI technology as a lynchpin of their future strategic dominance.[15]

A.   Why study AI's effects on the international order?

From both a practical perspective, as well as a theoretical one, it is important to consider how AI technology may change the tools, processes and assumptions of international law. At present, governance proposals for AI often fail to factor in technology-driven changes to governance itself. To be sure, independent bodies of scholarship exist which explore the use of AI, algorithms and digital technology in legal systems,[16] including a small but growing body of work exploring this phenomenon at the international law level.[17] However, most proposals for the global governance of AI itself often reason in relative isolation from such dynamics.

This is a problem because, as noted by Colin Picker, various technological innovations have throughout history driven the "creation, modification, or destruction of

---

[14] For instance, on the US side, the Pentagon has emphasized its intention to invest up to $2 billion in the next five years to develop programs advancing AI—and spent around $7.4 billion on AI in 2017. Drew Harwell, 'Defense Department Pledges Billions toward Artificial Intelligence Research' *Washington Post* (7 September 2018) <https://www.washingtonpost.com/technology/2018/09/07/defense-department-pledges-billions-toward-artificial-intelligence-research/> accessed 22 June 2020; Julian E Barnes and Josh Chin, 'The New Arms Race in AI' *Wall Street Journal* (2 March 2018) <https://www.wsj.com/articles/the-new-arms-race-in-ai-1520009261> accessed 22 November 2018. For discussion of funding for military AI projects, see: Justin Haner and Denise Garcia, 'The Artificial Intelligence Arms Race: Trends and World Leaders in Autonomous Weapons Development' (2019) 10 Global Policy 331. In March 2021, the US National Security Commission on Artificial Intelligence presented its final report to the US Congress, which recommended an investment of $40 billion. National Security Commission on Artificial Intelligence, 'Final Report' (National Security Commission on Artificial Intelligence 2021) <https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf> accessed 3 March 2021.

[15] In the US, this was initially articulated (in 2016) under the Obama administration; Office of Science and Technology Policy, 'The National Artificial Intelligence Research and Development Strategic Plan' (National Science and Technology Council 2016) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf> accessed 26 February 2017; This has since been rearticulated under the 'American Artificial Intelligence Initiative': Office of Science and Technology Policy, 'American Artificial Intelligence Initiative: Year One Annual Report' (The White House 2020) <https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>. Meanwhile, China's State Council in 2017 issued a plan that anticipated China becoming the world leader in the AI field by 2030: China's State Council, 'A Next Generation Artificial Intelligence Development Plan' (Rogier Creemers and others trs, New America Cybersecurity Initiative 2017) <https://na-production.s3.amazonaws.com/documents/translation-fulltext-8.1.17.pdf> accessed 23 October 2017. See also Jeffrey Ding, 'Deciphering China's AI Dream: The Context, Components, Capabilities, and Consequences of China's Strategy to Lead the World in AI' (Future of Humanity Institute, Governance of AI Program 2018) <https://www.fhi.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf?platform=hootsuite>. For general info on China and AI, one can refer to his 'ChinAI' newsletter, at https://chinai.substack.com/

[16] This literature is extensive. But for a brief survey, see: Roger Brownsword, 'In the Year 2061: From Law to Technological Management' (2015) 7 Law, Innovation and Technology 1; Karen Yeung, '"Hypernudge": Big Data as a Mode of Regulation by Design' (2017) 20 Information, Communication & Society 118; Benjamin Alarie, 'The Path of the Law: Towards Legal Singularity' (2016) 66 University of Toronto Law Journal 443; Anthony J Casey and Anthony Niblett, 'Self-Driving Laws' (2016) 66 University of Toronto Law Journal 429; Christopher Markou and Simon Deakin, 'Is Law Computable? From Rule of Law to Legal Singularity' in Christopher Markou and Simon Deakin (eds), *Is Law Computable? Critical Perspectives on Law + Artificial Intelligence* (Hart 2020) <https://papers.ssrn.com/abstract=3589184> accessed 15 May 2020; Brian Sheppard, 'Warming up to Inscrutability: How Technology Could Challenge Our Concept of Law' (2018) 68 University of Toronto Law Journal 36.

[17] Thomas Burri, 'International Law and Artificial Intelligence' (2017) 60 German Yearbook of International Law 91; Ashley Deeks, 'High-Tech International Law' (2020) 88 George Washington Law Review 575.

international law, or the derailment of the creation of new international law".[18] Indeed, scholars working in various governance areas have already begun to identify the adverse effects which digital technologies may have on the continued viability of regimes in areas such as arms control.[19] Likewise, AI technology is likely to affect not just the *substance* of various international regimes, but also their *processes and procedures*, and potentially even their *political scaffolding*.[20] This will have effects not just on the instruments envisaged for governing AI itself, but indeed will likely spread to many other legal domains. AI applications may also exert effects on the very AI governance instruments that are being proposed, or on the broader global legal order within which these would be ingrained.

How do we engage with the questions of how technology alters (international) law? There is an established body of work that has examined the phenomena of 'law and technology' and 'legal automation' in a domestic law context. Over the past decade, such scholarship has studied, through both concrete cases and at a conceptual level, how the use of algorithms and AI systems might alter the coherence, form, and practices of legal systems, or even the underlying values of regulators.[21]

Significantly, if one expects AI technologies to have such a disruptive impact on domestic legal systems, one should arguably expect their disruptive effects to be even steeper in the international domain. This is because in comparison to national law, the global legal order appears less well equipped to anticipate or resolve situations of legal uncertainty, or to keep accountable the insertion of technology in legal processes. After all, at the international level there is no authoritative final legislator to compel *ex ante* consideration of a technology, to guarantee the coherence of legal responses across domains, or to coordinate (and critically scrutinize) the integration of new technologies into law-making, -adjudication, or –enforcement practices. Moreover, amending or altering multilateral treaty regimes to take stock of subsequent technological developments can often prove a more painful, drawn-out, and contested affair than the revision of domestic laws.[22] Furthermore, governance disruption can pose a challenge that is not just legal or procedural, but also political. After all, while international relations scholars are still weighing the precise impact of AI,[23] there is a general appreciation that

---

[18] Colin B Picker, 'A View from 40,000 Feet: International Law and the Invisible Hand of Technology' (2001) 23 Cardozo Law Review 151, 156.

[19] Amy J Nelson, 'The Impact of Emerging Technologies on Arms Control Regimes' (ISODARCO, 2018) <http://www.isodarco.eu/courses/andalo18/paper/iso18-AmyNelson.pdf>; Amy J Nelson, 'Innovation Acceleration, Digitization, and the Arms Control Imperative' (Social Science Research Network 2019) SSRN Scholarly Paper ID 3382956 <https://papers.ssrn.com/abstract=3382956> accessed 29 May 2020. And see generally, Richard Danzig, 'An Irresistible Force Meets a Moveable Object: The Technology Tsunami and the Liberal World Order' (2017) 5 Lawfare Research Paper Series <https://assets.documentcloud.org/documents/3982439/Danzig-LRPS1.pdf> accessed 1 September 2017.

[20] Matthijs M Maas, 'International Law Does Not Compute: Artificial Intelligence and The Development, Displacement or Destruction of the Global Legal Order' (2019) 20 Melbourne Journal of International Law 29.

[21] The literature is vast, and no exhaustive review is here attempted. However, some contributions over the past years include: Roger Brownsword, 'Technological Management and the Rule of Law' (2016) 8 Law, Innovation and Technology 100; Benjamin Alarie, Anthony Niblett and Albert H Yoon, 'Law in the Future' [2016] University of Toronto Law Journal <https://www.utpjournals.press/doi/abs/10.3138/UTLJ.4005> accessed 28 January 2019; Alarie (n 17); Casey and Niblett (n 17); Sheppard (n 17); Woodrow Hartzog and others, 'Inefficiently Automated Law Enforcement' (2016) 2015 Michigan State Law Review 1763; Rebecca Crootof, '"Cyborg Justice" and the Risk of Technological-Legal Lock-In' [2019] Columbia Law Review Forum <https://papers.ssrn.com/abstract=3464724> accessed 18 November 2019; John Danaher, 'The Threat of Algocracy: Reality, Resistance and Accommodation' (2016) 29 Philosophy & Technology 245; Yeung (n 17).

[22] Rebecca Crootof, 'Jurisprudential Space Junk: Treaties and New Technologies' in Chiara Giorgetti and Natalie Klein (eds), *Resolving Conflicts in the Law* (2019) <https://brill.com/view/book/edcoll/9789004316539/BP000015.xml> accessed 15 March 2019.

[23] Michael C Horowitz, 'Artificial Intelligence, International Competition, and the Balance of Power' [2018] *Texas National Security Review* <https://tnsr.org/2018/05/artificial-intelligence-international-competition-

new innovations can redraw the global political map.[24] It would be very surprising if such far-reaching (geo)-political shocks did not have their echoes at the international legal level.

The aim of this paper is not to provide an exhaustive analysis, but rather to provide conceptual clarification—which has been fundamental to much past scholarship at the intersection of 'law', 'regulation' and 'technology'.[25] For instance, Colin Picker has noted that a high-level view of strategic dynamics amongst technology and international law is critical:

> "[e]ven though policy makers must be closely concerned with the "nitty gritty" of their international regimes and negotiations, […] [they] have much to gain from taking a macro or holistic view of the issues raised by technology. Macro-examinations can provide larger theoretical understandings and can reveal previously hidden characteristics that are simply not discernable from the "trenches." Viewing technology from "40,000 feet up" reveals certain patterns, pitfalls, and lessons for policy."[26]

The aim of this paper is to explore the intersection of AI and international law, in the hope that additional conceptual disaggregation can help international lawyers consider ways to 'update' or 'reboot' the international legal order, to ensure it is not only more resilient to AI-driven governance disruption, but can also build on these tools.

B.   AI and (International) Law: a LawTech-Techlaw Tale

The exploration of *AI technology's ability to change the norms, instruments, or processes of governance* can draw on a framework of 'governance disruption'.[27] This departs from the existing scholarship on 'law and technology'. There are two streams to this. On the one hand, recent years have seen a diverse literature on 'LawTech', which focuses on how new technologies can be used in support of existing legal systems, or to expand the capabilities of lawyers.[28] This work however, has predominantly focused on such applications in the domestic law context, rather than at the international law one. Simultaneously, a second and complementary strand is found in the emerging paradigm of 'TechLaw'—"the study of how law and technology foster, restrict, and otherwise shape

---

and-the-balance-of-power/> accessed 17 May 2018; Michael C Horowitz, 'Do Emerging Military Technologies Matter for International Politics?' (2020) 23 Annual Review of Political Science 385.

[24] Daniel W Drezner, 'Technological Change and International Relations' (2019) 33 International Relations 286, 287. ("[a]ny technological change is also an exercise in redistribution. It can create new winners and losers, alter actor preferences, and allow the strategic construction of new norms").

[25] For instance, in reflecting on these topics, Roger Brownsword and others have noted that "debates over these terms, and about the conceptualization of the field or some parts of it, can significantly contribute to our understanding." Roger Brownsword, Eloise Scotford and Karen Yeung, 'Law, Regulation, and Technology: The Field, Frame, and Focal Questions' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology*, vol 1 (Oxford University Press 2017) 6 <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199680832.001.0001/oxfordhb-9780199680832-e-1> accessed 3 January 2019.

[26] Picker (n 19) 151–152.

[27] Previously set out in Maas, 'International Law Does Not Compute' (n 21). A version of this is also developed at greater length in Hin-Yan Liu and others, 'Artificial Intelligence and Legal Disruption: A New Model for Analysis' (2020) 12 Law, Innovation and Technology 205. And see Maas, 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks' (n 2).

[28] Agnieszka McPeak, 'Disruptive Technology and the Ethical Lawyer' (2018) 50 University of Toledo Law Review 457. See also Rebecca Crootof and BJ Ard, 'Structuring Techlaw' (2021) 34 Harvard Journal of Law & Technology n 1 <https://papers.ssrn.com/abstract=3664124> accessed 28 August 2020.

each other's evolution."[29] This work has explored at length how new technologies can alter the *concepts, texture and dynamics* of legal systems.

Importantly, the idea that new (AI) technologies are not just objects for regulation[30] but can also change the operation and processes of legal systems, and even the goals of regulators, is not new. Reflections on what a new technology reveals about the changing face of law has extensive precedent in technology law scholarship. For instance, in early debates in the field of cyberlaw, Lawrence Lessig famously examined several legal questions involving the new technology of cyberspace, not merely in order to discuss the relative efficacy of different approaches to regulating certain topics (e.g. zoning or copyright) on the internet, but also to ground and illustrate systemic reflections on the changing nature and workings of the different 'regulatory modalities' of laws, norms, markets and architectures ('code').[31] Likewise, Roger Brownsword has used studies of behaviour-shaping technologies and geoengineering in order to reflect respectively, upon the rising role of non-normative 'technological management' and 'regulatory responsibilities' for the core 'global commons'.[32]

Moreover, this comparison with earlier scholarship on cyberlaw is illuminating in both directions. The internet 'merely' led to the 'informatisation' of infrastructure and public space—which already proved pivotal in altering the ways that regulation operates. In their turn, AI systems may enable the increasing 'intelligentisation' or 'cognitisation' of these infrastructures,[33] suggesting that this technology could in time have an impact on the practices and dynamics of law and governance that is at least as far-reaching. The 'governance disruption' developed in this paper is of use, as exploring the dynamics and legal impacts of AI governance can reveal interesting and important transferable lessons for the changing relation between technology and global governance—and therefore the changing face of international law in the 21st century.

1.   The rise of LawTech

This debate over how AI technologies might provide new affordances for *the regulator*s who produce and enforce law (and what implications flow from this) becomes particularly urgent, given the developments in AI for LawTech in many domestic legal contexts. Over the last decade, new digital technologies, and particularly AI, have seen particularly rapid and enthusiastic uptake in various branches of domestic public administration, judicial decision-making, policing, and the business of government writ

---

[29] Crootof and Ard (n 29) n 1. See also succinctly BJ Ard and Rebecca Crootof, 'The Case for "Technology Law"' (*Nebraska Governance & Technology Center*, 16 December 2020) <https://ngtc.unl.edu/blog/case-for-technology-law> accessed 16 March 2021.

[30] Miriam C Buiten, 'Towards Intelligent Regulation of Artificial Intelligence' (2019) 10 European Journal of Risk Regulation 41. However, for a critique of technology law approaches that aim to regulate specific 'technologies' (and an argument that the emphasis should instead be on 'regulating for sociotechnical change', see notably Lyria Bennett Moses, 'How to Think about Law, Regulation and Technology: Problems with "Technology" as a Regulatory Target' (2013) 5 Law, Innovation and Technology 1; 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* (2017) <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199680832.001.0001/oxfordhb-9780199680832-e-49> accessed 13 May 2017.

[31] Lawrence Lessig, 'The Law of the Horse: What Cyberlaw Might Teach' (1999) 113 Harvard Law Review 501; Lawrence Lessig, *Code: And Other Laws of Cyberspace, Version 2.0* (2nd Revised ed. edition, Basic Books 2006) <http://codev2.cc/download+remix/Lessig-Codev2.pdf>.

[32] Roger Brownsword, *Law, Technology and Society: Re-Imagining the Regulatory Environment* (1 edition, Routledge 2019); Roger Brownsword, 'Law and Technology: Two Modes of Disruption, Three Legal Mind-Sets, and the Big Picture of Regulatory Responsibilities' (2018) 14 Indian Journal of Law and Technology 1.

[33] This echoes the terminology used by the Chinese PLA. Elsa Kania, 'AlphaGo and Beyond: The Chinese Military Looks to Future "Intelligentized" Warfare' (*Lawfare*, 5 June 2017) <https://www.lawfareblog.com/alphago-and-beyond-chinese-military-looks-future-intelligentized-warfare> accessed 10 June 2017.

large. At the same time, from the perspective of (private) legal practice, AI systems can be used to automate or contribute to a range of routine legal tasks, such as checking contracts,[34] case law research,[35] or the faster provision of legal services.[36] Provided with adequate databases, some machine learning systems have even begun to make headway in predicting the outcomes of legal disputes in court.[37] This has led some scholars to anticipate an increasing automation of legal systems,[38] 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, and so can offer perfectly clear yet tailored legal clarity for individuals.[39]

That is not to say that such scenarios—or indeed even more modest cases of legal automation—have been uncritically welcomed. Indeed, recent years have seen extensive scholarship exploring the practical, normative, and legal implications of 'legal automation' in domestic practice.[40] There are extensive critiques of the near-term prospects for legal automation, given the limits of current machine learning approaches,[41] as well as in light of the various 'interface design errors' that often plague hybrid human and AI 'cyborg justice' arrangements, and which can produce inappropriate 'overtrust'.[42] There have been particularly acute cases involving bias in algorithms used for recidivism prediction;[43] and

---

[34] LawGeex, 'Comparing the Performance of Artificial Intelligence to Human Lawyers in the Review of Standard Business Contracts' (LawGeex 2018) <https://images.law.com/contrib/content/uploads/documents/397/5408/lawgeex.pdf> accessed 27 August 2020.

[35] See for instance the CARA AI system. 'CARA A.I.' (*Casetext*) <https://casetext.com/cara-ai/> accessed 8 September 2020.

[36] See for instance the app 'DoNotPay'. RJ Vogt, 'DoNotPay Founder Opens Up On "Robot Lawyers"' (*Law360*, 9 February 2020) <https://www.law360.com/articles/1241251/donotpay-founder-opens-up-on-robot-lawyers> accessed 8 September 2020.

[37] For instance, a system trained on 584 judicial decisions of the European Court of Human Rights managed to predict the outcome of new cases with 79% accuracy. Nikolaos Aletras and others, 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective' (2016) 2 PeerJ Computer Science e93. 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, 'Using Machine Learning to Predict Decisions of the European Court of Human Rights' (2020) 28 Artificial Intelligence and Law 237. Although for a general critique of such projects, see Frank A Pasquale and Glyn Cashwell, 'Prediction, Persuasion, and the Jurisprudence of Behaviorism' [2017] University of Maryland Francis King Carey School of Law Legal Studies Research Paper <https://papers.ssrn.com/abstract=3067737> accessed 8 September 2020.

[38] Alarie, Niblett and Yoon (n 22) 424.

[39] Casey and Niblett (n 17) 430; Anthony J Casey and Anthony Niblett, 'The Death of Rules and Standards' (2017) 92 Indiana Law Journal 1401, 1410–1411.

[40] Joshua P Davis, 'Law Without Mind: AI, Ethics, and Jurisprudence' (Social Science Research Network 2018) SSRN Scholarly Paper ID 3187513 <https://papers.ssrn.com/abstract=3187513> accessed 23 June 2018. For a discussion of the implications of unintelligible legal automation for core theories of law—notably HLA Hart's account which requires critical officials; Joseph Raz's concept grounded in reason-based tests of legitimacy, or Ronald Dworkin's work on deep justifications for coercion—see Sheppard (n 17).

[41] See for instance: Frank Pasquale and Glyn Cashwell, 'Four Futures of Legal Automation' (2015) 26 UCLA Law Review Discourse 23; Frank A Pasquale, 'A Rule of Persons, Not Machines: The Limits of Legal Automation' (2019) 87 George Washington Law Review 1; Mireille Hildebrandt, 'Law As Computation in the Era of Artificial Legal Intelligence. Speaking Law to the Power of Statistics' (Social Science Research Network 2017) SSRN Scholarly Paper ID 2983045 <https://papers.ssrn.com/abstract=2983045> accessed 9 July 2018.

[42] Crootof, '"Cyborg Justice" and the Risk of Technological-Legal Lock-In' (n 22).

[43] Lauren Kirchner and others, 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks.' (*ProPublica*, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 24 May 2017. But see also Sam Corbett-Davies and others, 'A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased against Blacks. It's Actually Not That Clear.' *Washington Post* (17 October 2016) <https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/> accessed 21 May 2017.

9

others have highlighted likely human rights violations that are likely when such algorithms are deployed on or to particularly vulnerable populations, such as migrants.[44]

More conceptually, others have cautioned that the automation of law enforcement systems might close the loop on the legal qualities of 'inefficiency' and 'indeterminacy', arguing that these are both key safeguards against the perfect enforcement of laws that were in fact originally drafted on the implicit assumption of a certain degree of lenience or discretion.[45] Finally, there are concerns that the integration of such systems may produce an increasingly opaque 'algocracy'.[46]

However, while these technologies therefore no doubt raise diverse grounds of concern, barring a major public backlash, it may well be the case that they see continued development and deployment of automation in public services and administration, potentially signalling a general shift towards 'technocratic' regulatory attitudes in domestic legal systems.[47]

2.   Taking legal disruption to the international level

Given its foundational implications, this breadth of work and attention on 'legal displacement' in domestic legal systems is surely warranted. However, by comparison, there has been relatively less attention paid to the prospects of such a shift at the international legal level.

There is indeed some recent work on how AI technology may come to affect and disrupt international law. Thomas Burri, Berenice Boutin, and Ashley Deeks have provided general treatments of the ways in which AI technology could affect or play a role in international law.[48] Meggido has explored the potential contribution of big data to customary international law.[49] Alschner and colleagues have explored the production of international legal materials (such as draft treaties) by algorithms,[50] and have explore some of the implications of the 'data-driven future' of the (text-corpus-rich) international economic law.[51] Deeks and others explore the specific implications of conflict prediction algorithms on the international law on the use of force.[52] Livingston & Risse have discussed the implications of AI for potential monitoring of human rights.[53] More generally, Nelson has evaluated the effects of digitisation (including the spread of AI) on

---

[44] Petra Molnar, 'Technology on the Margins: AI and Global Migration Management from a Human Rights Perspective' (2019) 8 Cambridge International Law Journal 305.

[45] Hartzog and others (n 22).

[46] Danaher (n 22).

[47] See: Roger Brownsword, 'Law Disrupted, Law Re-Imagined, Law Re-Invented' [2019] Technology and Regulation 10.

[48] Burri (n 18); Berenice Boutin, 'Technologies for International Law & International Law for Technologies' (*Groningen Journal of International Law*, 22 October 2018) <https://grojil.org/2018/10/22/technologies-for-international-law-international-law-for-technologies/> accessed 31 October 2018; Deeks (n 18); Ashley Deeks, 'Introduction to the Symposium: How Will Artificial Intelligence Affect International Law?' (2020) 114 AJIL Unbound 138.

[49] Tamar Megiddo, 'Knowledge Production, Big Data and Data-Driven Customary International Law' (Social Science Research Network 2019) <https://papers.ssrn.com/abstract=3497477> accessed 21 January 2020.

[50] Wolfgang Alschner and Dmitriy Skougarevskiy, 'Towards an Automated Production of Legal Texts Using Recurrent Neural Networks', *Proceedings of the 16th edition of the International Conference on Articial Intelligence and Law* (Association for Computing Machinery 2017) <https://doi.org/10.1145/3086512.3086536> accessed 14 May 2020; Wolfgang Alschner and Dmitriy Skougarevskiy, 'Can Robots Write Treaties? Using Recurrent Neural Networks to Draft International Investment Agreements' in F Bex and S Villata (eds), *JURIX: Legal Knowledge and Information Systems* (IOS Press 2016) <https://papers.ssrn.com/abstract=2984935> accessed 14 May 2020.

[51] Wolfgang Alschner, Joost Pauwelyn and Sergio Puig, 'The Data-Driven Future of International Economic Law' (2017) 20 Journal of International Economic Law 217.

[52] Ashley Deeks, Noam Lubell and Daragh Murray, 'Machine Learning, Artificial Intelligence, and the Use of Force by States' (2019) 10 Journal of National Security Law & Policy 1.

[53] Steven Livingston and Mathias Risse, 'The Future Impact of Artificial Intelligence on Humans and Human Rights' (2019) 33 Ethics & International Affairs 141.

10

existing arms control and export control regimes.[54] However, this scholarship remains incipient and at an early stage.

As such, this article finds itself at the intersection of the LawTech and TechLaw approaches. Drawing on the work of Rebecca Crootof and Ashley Deeks, it aims to extend these paradigms to the international legal level, by exploring how the use of AI may affect the form and viability of international law—and what are the resulting practical, normative, and theoretical implications.

III.   AI and the international order: sources of disruption, vectors of displacement

In previous work, I have mapped a taxonomy of ways in which AI could affect the integrity, viability, form or relevance of the international law system itself.[55] Based on historical accounts and theoretical frameworks, I argued that the deployment and use of AI systems might produce three types of global legal impacts: 'development', 'displacement', or 'destruction' (see Table 1).

This model maps how AI applications can and may under various circumstances produce situations of legal uncertainty (i.e. new gaps, ambiguity, over- or under-inclusiveness, or obsolescence) around existing norms or regimes in international law. Whenever this happens, it sparks a moment of reflection—an 'international law step zero'[56]—over whether or how these can be accommodated into international law; that is, there is a need to accommodate or address the uncertainty through some form of legal *development*. However, the international law system may not always be able to carry out such developments, given certain features of- and tensions within its usual tools.[57]

Indeed, it was argued that certain technical and political features of AI technology may in practice render it destructive or erosive to key areas of international law (legal *destruction*). This is because the legal 'gaps' that AI systems reveal on the international level may be conceptually or politically hard to patch through the usual avenues of international law 'development' (adaptive interpretation; treaty amendments; new treaties; customary international law development; or jurisprudence by international courts).[58] At the same time, the strategic capabilities it offers chip away at the rationales for powerful states to engage fully in, or comply with, international law regimes.

---

[54] Nelson (n 20).

[55] Maas, 'International Law Does Not Compute' (n 21); see also Liu and others (n 28).

[56] Kristen E Eichensehr, 'Cyberwar & International Law Step Zero' (2015) 50 TEXAS INTERNATIONAL LAW JOURNAL 24; see also Rebecca Crootof, 'Regulating New Weapons Technology' in Eric Talbot Jensen and Ronald TP Alcala (eds), *The Impact of Emerging Technologies on the Law of Armed Conflict* (Oxford University Press 2019).

[57] See also Crootof, 'Jurisprudential Space Junk' (n 23).

[58] For a comparative argument that the usual (state consent-based) toolset of international law may not be well equipped to deal with certain types of global commons challenges, see also Anne van Aaken, 'Is International Law Conducive To Preventing Looming Disasters?' (2016) 7 Global Policy 81. And for reflections on the way forward, see Nico Krisch, 'The Decay of Consent: International Law in an Age of Global Public Goods' (2014) 108 American Journal of International Law 1. But see also Eyal Benvenisti and George W Downs, 'Comment on Nico Krisch, "The Decay of Consent: International Law in an Age of Global Public Goods"' (2014) 108 AJIL Unbound 1.

	Type		Example	Outcome
Need for Development	**New governance gaps**		AI-enabled swarm warfare (possibly) not covered by existing international regimes	Need for new law.
	Conceptual uncertainty or ambiguity		LAWS highlight potential ambiguity or inadequacy of concepts such as 'intent', 'effective control', etc.	Need for new law or adaptation of law, to sharpen existing rules or clarify concepts.
	Incorrect scope of application		Underinclusive application of Convention Against Torture to use of autonomous robots for interrogation.	Need for new law or adaptation of law, to demarcate scope and applicability of existing instruments.
			Overinclusive applicability of company law enabling incorporation of 'algorithmic entities' with corporate legal personhood.	
	Obsolescence	Behaviour obsolete (necessity)	New types of AI-supported remote biometric surveillance (gait or heartbeat identification) replace face recognition.	Need for new law or adaptation of existing instruments, involving a (re)consideration of the underlying aims and purposes of the regime.
		Justifying assumptions no longer valid (adequacy)	Structural unemployability through technological unemployment puts pressure on right to work, ILO regimes.	
		No longer cost-effective (enforceability)	Use of DeepFakes or computational propaganda raises monitoring and compliance enforcement costs for various regimes.	
	Altered problem portfolio		Military AI regime tailored to respond to *ethical challenges* of LAWS (e.g. maintaining meaningful human control over lethal force) might not be oriented to address risks of later adjacent AI capabilities (e.g. cyberwarfare) creating *structural shifts*.	Need for new regime or adaptation or amendment of existing instrument, to shift focus or centre of gravity.
Displacement	**Automation**	International Law Creation & Adjudication	Use of AI text-as-data tools to generate draft treaties, predict arbitral panel rulings, identify state practice, identify treaty conflicts.	Potential speeding-up of international law creation, though distributional and legitimacy concerns.
		Monitoring & enforcement	Use of various AI tools in monitoring and verifying treaty compliance.	Increase deterrent effect of existing regimes; potentially create transparency pre-conditions for new regimes
	Replacement	Changes in regulatory modality	Use of AI tools such as emotion-recognition, social media sentiment analysis, or computational propaganda by states, resulting in increased state preference to resolve disputes in diplomatic channels.	Distributional changes; more unilateral and strategic action by states; potential legal destruction.
Destruction	**Erosion**	Conceptual friction	Attempted extension of existing regimes or norms to new technology cannot pass 'laugh test'.	Inhibits needed developments within existing law; new regimes needed.
		Political 'knots'	Attempted extension of existing regimes or creation of new law, intractable because of political gridlock.	Inhibits needed developments through creation of new regimes; problem remains unsolved.
	Decline	Increasing the spoils of noncompliance	Innovations increase strategic stakes or ability to bypass monitoring, or lower proliferation thresholds or (political) noncompliance costs.	Decays political foundations of specific regimes.
		Active weapon	AI-enabled computational propaganda enables contestation of international law; Suspected use of AI negotiation tools subverts legitimacy of resulting agreements.	Scaling up contestation of international law.
		Shift of values	AI capabilities perceived as enabling unilateralism, alternative to multilateralism	Additional pressure on global legal order

Table 1. A Governance Disruption Framework[59]

---

[59] Reproduced from Maas, 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks' (n 2) 196.

Nonetheless, a third major category of AI's governance disruption concerns the ways in which AI tools may drive *displacement* in the *processes* or instruments of international governance themselves.[60] That is, more even than many previous technologies, it may drive considerable sociotechnical change in the 'legal' domain itself, as various tools might potentially lend themselves to the 'automation' or even the 'replacement' of important (rule-making, adjudication, monitoring & enforcement) functions of international law.

In particular, I have argued that there are three distinct sub-categories of AI-driven governance 'displacement'.[61] Specifically, these are the automation of processes of rule creation and adjudication; the automation of monitoring and enforcement of international regimes; and potential shifts in the 'regulatory modality' of international law, resulting in the gradual replacement of some forms of conflict mediation (e.g. arbitration) with other avenues (e.g. bilateral state diplomacy; or unilateral AI-supported 'lawfare').

A.   *Automation of rule creation, adjudication or arbitration*

In the first place, it has long been recognised that new technologies can change the *processes* by which international law is created. For instance, technologies can *speed up* international law formation. Already in 1973, Louis B. Sohn noted how new communication and travel infrastructures were making treaty negotiation faster and easier.[62] Of course, while promising, this may not be an unalloyed good, because the increased presence of communications technologies has also, in some readings, rendered strategies of 'lawfare' more viable.[63] Others have recently argued that the deluge of big data, and the falling thresholds to collecting it, could facilitate the process of collating and consolidating evidence of state practice, fostering an era of 'Data-Driven Customary International Law'.[64]

Along with speed, new communication technologies have also qualitatively altered the processes of international law, for instance by *expanding participation* to more actors. Indeed, in their discussion of the phenomenon of 'norm cascades', Finnemore & Sikkink have discussed how the pace and scope of (general) norm acceptance in international relations can be directly traced to technological change, since;

> "changes in communication and transportation technologies and increasing global interdependence have led to increased connectedness and, in a way, are leading to the homogenization of global norms […] the speed of normative change has accelerated substantially in the later part of the twentieth century."[65]

For instance, in a comparison of the drafting processes behind UNCLOS and the Mine Ban Treaty, Gamble and Ku have argued that modern communications technologies

---

[60] Maas, 'International Law Does Not Compute' (n 21) 43–49.

[61] ibid. It should be noted that this paper distinguished between just two categories of displacement: 'The automation of International Law', comprising both the use of AI in rule creation or adjudication as well as in monitoring and enforcement; and 'The Technological Replacement of International Law'. In this paper, I have split out the former category for clarity.

[62] Louis Sohn, 'The Impact of Technological Changes on International Law' (1973) 30 Washington and Lee Law Review 1, 10.

[63] Charles Dunlap, 'Lawfare Today: A Perspective' [2008] Yale Journal of International Affairs 146. As discussed by Crootof, 'Regulating New Weapons Technology' (n 58) 6–7. On lawfare generally, see also Jill I Goldenziel, 'Law as a Battlefield: The U.S., China, and Global Escalation of Lawfare' (2020) 106 Cornell Law Review <https://papers.ssrn.com/abstract=3525442> accessed 8 March 2021.

[64] Megiddo (n 50).

[65] Martha Finnemore and Kathryn Sikkink, 'International Norm Dynamics and Political Change' (1998) 52 International Organization 887, 909.

greatly increased the ability of NGOs to get involved in international rule creation, resulting in laws that were less narrowly tailored to state interests.[66]

Accordingly, the first type of legal displacement envisions AI systems changing the *practices* of (international) law *creation* or *adjudication*. Ideas to this effect draw an analogy to growing work on the use of AI systems in domestic legal systems, where algorithms are used in support of the adjudication of cases, the monitoring of crimes and violations, or even the prediction of court rulings. While this rise of 'AI Justice'[67] has received considerable scrutiny and critique even in the domestic realm, could we expect AI systems to see usage in support of international law? One can consider a 'strong' scenario, and a 'modest' scenario.[68]

1.   Strong scenario: a global digital arbiter

The 'strong' scenario posits that advances in AI could, in time, posit some sort of 'global digital arbiter'; a queryable legal model that would facilitate the creation of a 'completely specified' international law.[69] If possible, such systems could serve as a powerful 'integrator' of global governance, as they would be able to identify and resolve conflicts between norms, or goals—from human rights to environmental law—in adjudicating a 'global constitution'—therefore reducing the challenges of fragmentation and regime complexity.[70] At the limit, such systems would promise the fully automated adjudication of international law.[71]

Of course, while an interesting thought experiment or extrapolation of trends, for the purposes of governance this strong scenario may be less relevant. Technically, such a system remains beyond existing machine learning approaches, and might instead require some forms of future 'high-level machine intelligence' capabilities.[72] While it is not clear that such capabilities are categorically out of reach—or even more than a few decades away[73]—the fact remains that if they were to be created and implemented, their general societal impact would likely far exceed their direct impact on international law.

---

[66] John Gamble and Charlotte Ku, 'International Law - New Actors and New Technologies: Center Stage for NGOs' (2000) 31 Law and Policy in International Business 221, 249–251; Crootof, 'Regulating New Weapons Technology' (n 58) 6.

[67] Richard M Re and Alicia Solow-Niederman, 'Developing Artificially Intelligent Justice' (2019) 22 STANFORD TECHNOLOGY LAW REVIEW 48.

[68] Because the latter is more likely and analytically interesting, we will briefly discuss the former first.

[69] Along the lines of Benjamin Alarie's 'legal singularity', Alarie (n 17).

[70] We will discuss the intersection of (even more modest) AI tools with the lens of 'regime complexity' near the end of this paper.

[71] Although interestingly, while a dramatic change, such a 'full automation' of international law would not be the same as the 'replacement' of international law, since it would still at its core be reliant on presenting norms to human parties, and adjudicating or arbitrating between them.

[72] In which case global society might have more pressing problems to contend with. JG Castel and Mathew E Castel, 'The Road to Artificial Superintelligence - Has International Law a Role to Play?' (2016) 14 Canadian Journal of Law & Technology <https://ojs.library.dal.ca/CJLT/article/download/7211/6256>; Reinmar Nindler, 'The United Nation's Capability to Manage Existential Risks with a Focus on Artificial Intelligence' (2019) 21 International Community Law Review 5.. For expert estimates of timelines of such developments, see Katja Grace and others, 'When Will AI Exceed Human Performance? Evidence from AI Experts' (2018) 62 Journal of Artificial Intelligence Research 729. As well as Ross Gruetzemacher, David Paradice and Kang Bok Lee, 'Forecasting Extreme Labor Displacement: A Survey of AI Practitioners' (2020) 161 Technological Forecasting and Social Change 120323; Ross Gruetzemacher and others, 'Forecasting AI Progress: A Research Agenda' [2020] arXiv:2008.01848 [cs] <http://arxiv.org/abs/2008.01848> accessed 11 August 2020.

[73] Gwern Branwen, 'On GPT-3 - Meta-Learning, Scaling, Implications, And Deep Theory' <https://www.gwern.net/newsletter/2020/05> accessed 21 September 2020; Rich Sutton, 'The Bitter Lesson' (*Incomplete Ideas*, 2019) <http://www.incompleteideas.net/IncIdeas/BitterLesson.html> accessed 16 December 2019; but see Rodney Brooks, 'A Better Lesson' (19 March 2019) <https://rodneybrooks.com/a-better-lesson/> accessed 18 May 2020.

More pragmatically, proposing the use of an advanced AI system to create or adjudicate international law in this strong sense would be 'passing the buck'. The 'problem' of global governance (from a coherentist or global constitutionalist perspective) is not that we presently cannot conceive of an authoritative body that could reconcile or decide amongst norm conflicts. It is rather that such hierarchical authorities currently only exist at the state level, whereas the international system notably lacks such bodies. There would certainly be distinct and considerable political (and symbolic) sensitivities involved in handing over the international legal order over to a machine, but these seem modest compared to the extant political difficulty of installing *any* final authority above the international system.

In that sense, the 'strong' scenario of international legal automation might be somewhat of a red herring, since if one or more AI systems were deployed to integrate and resolve the problem of authority at the international level, by far the bigger achievement would be that this was achieved at all, not that it was done in silicate. As such, a far-reaching 'automation' of international law might counteract the trend towards the increasing international legal and normative fragmentation;[74] but its prospects appear technologically limited in the near term, and strategically moot in the hypothetical long-term.[75]

2. Modest scenario: text-as-data tools for 'intelligentized' lawmaking

By contrast, the 'modest' scenario of using AI in the automation for rule creation is more analytically fruitful. Of course, a first question to ask is whether even this scenario of 'displacement' is anywhere on the horizon. Can international law be even partially automated in this more modest manner? Thomas Burri has been broadly sceptical: he argues that whereas domestic legal areas such as tax law are susceptible to legal automation because AI systems in those areas can draw on large, dense, structured and homogeneous datasets;[76] the key legal 'datasets' at the international law level are either far too small (e.g. ICJ decisions), or far too heterogeneous and ambiguous to allow for this.[77]

By contrast, Ashley Deeks has been more optimistic about the conditions for gradual but eventually wide-spread take-up of AI technology into 'high-tech international law'.[78] Contra Burri, she argues that in many areas in international law, there are significant digital sources of texts (covering thousands of documents) which provide extensive text corpora that can allow machine learning 'text-as-data' tools to perform various functions. She notes how:

> "[o]ne key reason to think that international legal technology has a bright future is that there is a vast range of data to undergird it. […] there are a variety of digital sources of text that might serve as the basis for the kinds of text-as-data analyses that will be useful to states. This includes U.N. databases of Security Council and General Assembly documents, collections of treaties and their travaux preparatoires, European Court of Human Rights caselaw, international arbitral awards, databases of specialized agencies

---

[74] See Infra, section IV(B).

[75] For an exploration of a 'problem-finding' approach to such macro-strategic trajectory considerations around AI governance, see also Liu and Maas (n 8).

[76] This may be too easy a reading of the difficulties involved in tax law automation, especially in civil law systems where precedent does not hold the same value. I thank Luisa Scarcella for this point. For further discussion, see Joshua D Blank and Leigh Osofsky, 'Automated Legal Guidance' (2021) 106 Cornell Law Review <https://papers.ssrn.com/abstract=3546889> accessed 8 September 2020.

[77] Burri (n 18) 93–95.

[78] Deeks (n 18).

such as the International Civil Aviation Organization, state archives and digests, data collected by a state's own intelligence agencies and diplomats (memorialized in internal memoranda and cables), states' notifications to the Security Council about actions taken in self-defense, legal blogs, the U.N. Yearbook, reports by and submission to U.N. human rights bodies, news reports, and databases of foreign statutes."[79]

Given this, text-as-data machine learning systems could be trained to generate new treaty texts (e.g. draft extradition treaties),[80] to predict how an international arbitral panel might rule, to gauge the course and likely outcomes of treaty negotiations, or to identify possible treaty conflicts.[81] Moreover, far from being limited only to legal texts, Deeks notes how other uses of AI—from systems that aggregate intelligence information about the preferences or negotiation strategies of negotiation partners, to emotion-recognition systems or social media sentiment analysis—might also play a role in broader diplomatic processes.[82]

Focusing on the *procedural* contributions that machine learning could make to the creation of new international law, she argues that international lawyers in service of a state's foreign ministry in practice have three roles—negotiating agreements; dispute resolution (whether in judicial or arbitral fora); and advising policymakers about the existence and meaning of international law—and argues that all of these can, in distinct ways, gain from automation.[83]

Contrary to the 'strong' scenario discussed above, this modest' image of legal automation does not envision that these lawmaking processes will be entirely given over to machine decision-making, but rather that many actors will come to find significant benefits in using machine learning tools to support human legal and diplomatic decision-making processes.

To be certain, that does not mean this process of displacement will occur rapidly. There are a number of barriers which can slow procurement and integration of AI tools, and which ensure a lag between the development of state-of-the-art tools in labs, and their use in the field.[84] Deeks does admit that international law can often be 'conservative' about taking up new technologies, and that at present, states and their advisors in international legal issues still lag far behind the private sector in contemplating how AI could change their work. As a result, with a few exceptions, "international law generally has been a stranger to a new wave of technological tools – including computational text analysis, machine learning, and predictive algorithms – that use large quantities of data to help make sense of the world."[85]

Furthermore, there remain challenges to the incorporation of AI technologies in international law, including technical challenges around the format of some types of international law data; the fact that international law questions have relatively less precedent to guide predictions; and civil liberties concerns over applications in emotion detection or social media scraping software, amongst others.[86]

In spite of this, Deeks does identify several trends that will, in her view, steadily increase the pressure on international lawyers to take up and reckon with these new tools.

---

[79] ibid 596–597. [citing sources]

[80] ibid 605.

[81] ibid 604–606, 616-622 628-630.

[82] ibid 613–615.

[83] ibid 589–593.

[84] Maaike Verbruggen, 'AI & Military Procurement: What Computers Still Can't Do' (*War on the Rocks*, 5 May 2020) <https://warontherocks.com/2020/05/ai-military-procurement-what-computers-still-cant-do/> accessed 12 May 2020; Maaike Verbruggen, 'The Role of Civilian Innovation in the Development of Lethal Autonomous Weapon Systems' (2019) 10 Global Policy 338; Horowitz (n 24).

[85] Deeks (n 18) 576.

[86] ibid 598–599.

These are: (1) near-peer ambitions; (2) proofs of concept in private law; (3) client pressure; (4) necessity in the line of work.[87] As such, she argues that while the challenges to adoption are real, they are not insurmountable, and that there are potentially many use cases of AI, whether in the preparation of treaty negotiations, conducting negotiations, or identifying customary international law.[88]

In addition to these, there might of course be a more fundamental limit to the significance of governance displacement-through-automation: that is, even if there are many domestic areas where there are significant sources for text-as-data tools to be trained on, these may not be the highest-stakes scenarios that international law faces. After all, international law may be especially engaged by large historical shocks and changes.[89] That would not prevent such AI systems from seeing wide usage in many rote tasks, but it would imply that international legal history would continue to be written by humans for some time.

Nonetheless, even this modest scenario might have significant effects. While some AI tools might support greater coordination, the fact that they are developed and owned by some parties may, as Deeks has argued, mean that "states with a high level of technological sophistication are likely to treat some of these tools as proprietary and critical to their national security, and so may use them in a way that exacerbates existing power differentials."[90] This could have erosive effects, and markedly alter dynamics of contestation and the legitimacy of the global governance architecture.

B.   *Automation of monitoring & enforcement*

The second form of governance displacement envisions the use of AI technologies in strengthening the *enforcement* of existing or future international law and global governance instruments.

To be certain, it has long been recognised that new technologies can play a key role in changing the modes and anticipated effectiveness of compliance monitoring and enforcement. Indeed, the role of various forms of 'National Technical Means'[91] in supporting the monitoring of compliance with international commitments, has been well-chronicled in the high-stakes area of arms control. For many decades, a range of technologies including signals intelligence, satellites, networked arrays of seismic, hydroacoustic, infrasound or radionuclide monitoring stations,[92] and (aided by legal arrangements such as the 1992 Treaty on Open Skies) surveillance or radionuclide 'sniffer' aircraft,[93] have all played roles in enabling states parties to monitor and verify

---

[87] ibid 593–597.

[88] ibid 599.

[89] I thank Laura Emily Christiansen for this point. See also Michael P Scharf, 'Seizing the "Grotian Moment": Accelerated Formation of Customary International Law in Times of Fundamental Change' (2010) 43 31.

[90] Deeks (n 18) 582.

[91] Eric H Arnett, 'Science, Technology, and Arms Control' in Richard Dean Burns (ed), *Encyclopedia of Arms Control and Disarmament*, vol 1 (Charles Scribner's Sons 1993).

[92] See for instance 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, and which deploys four complementary verification methods (seismic, hydroacoustic, infrasound and radionuclide sensors) at 321 monitoring stations spread over 89 countries, in order to detect any signs of nuclear testing. CTBTO Preparatory Commission, 'Overview of the Verification Regime' (*Comprehensive Nuclear-Test-Ban Treaty Organization*) <https://www.ctbto.org/verification-regime/background/overview-of-the-verification-regime/> accessed 9 September 2020.

[93] Treaty on Open Skies 1992 (CTS No 3). See generally also David A Koplow, 'Back to the Future and up to the Sky: Legal Implications of Open Skies Inspection for Arms Control' (1991) 79 California Law Review 421. However, in May 2020, President Donald Trump announced an intention for the US to withdraw from the Open Skies Treaty. Bonnie Jenkins, 'A Farewell to the Open Skies Treaty, and an Era of Imaginative

17

each other's (non)compliance with treaty commitments or various peremptory norms under international law.[94] Indeed, aerial and spatial observation alike continue to play an important role in weapons inspections that underpin arms control.[95] Notably, such tools are critical, not only because they can improve compliance with treaties that are in force (by changing the incentives and calculations of actors regarding the expected results of violating the treaty), but also because in some cases, these technological means can provide the guarantees by which parties are willing to bind themselves to such agreements (that is, set up a regime) in the first place.

For instance, Coe and Vaynmann have argued that historically, a critical barrier to reaching arms control agreements has been found in a so-called 'security-transparency' tradeoff.[96] This refers to the dilemma whereby an arms control agreement between A and B to cap B's capabilities must on the one hand have sufficient provisions in place to ensure transparency of B's actions (to ensure A that any cheating can and will be detected), but on the other hand, must not offer A so much transparency of B's military actions, that B holds the arrangement to unacceptably erode its security. In some cases, it may be impossible to satisfy both requirements at once, for, as Coe and Vaynman note:

> "[a]ny deal that is transparent enough to assure that one side complies with the deal may also shift the balance of power so much that the other side reneges to exploit this shift. Any deal that preserves the balance of power well enough to be safe for the arming side may not be transparent enough to assure the monitoring side of its compliance. When this is true, no arms control deal will be viable."[97]

While the 'security-transparency' trade-off might appear to provide grounds for pessimism regarding the prospects of arms control, however, this model also allows that whenever this trade-off is less steep, or where one side can assure itself of the ability to unilaterally monitor the other's compliance, agreements can be struck. Importantly, this highlights structural 'levers' that can be affected, not just through institutional arrangements or confidence-building measures, but also through new technologies, in order to mitigate or sidestep this trade-off. As such, technologies can structurally shift the conditions for an arms control treaty if they increase unilateral compliance monitoring, or if they can reduce the 'transparency-security' trade-off.

In the first case, technological advances that *increase either party's ability to unilaterally monitor* certain capabilities can render a 'closed deal' (whereby parties agree on an arms control deal, but refuse to allow access to facilities for the purposes of mutual verification) viable nonetheless, because either party can be assured that they are able to unilaterally monitor compliance. For instance, in the domain of nuclear arms control, Coe and Vaynmann suggest that the Strategic Arms Limitations Talks (SALT I) negotiations succeeded where the earlier 1964 Freeze negotiations had failed, in large part because the development of satellite surveillance in the intervening years had improved unilateral monitoring capabilities sufficiently to make a 'closed deal' viable.[98]

In the second place, there may be ways to take measures that create mutual perceptions that the *trade-off between transparency and security is mild*. For instance,

---

Thinking' (*Brookings*, 16 June 2020) <https://www.brookings.edu/blog/order-from-chaos/2020/06/16/a-farewell-to-the-open-skies-treaty-and-an-era-of-imaginative-thinking/> accessed 22 June 2020.

[94] Maas, 'International Law Does Not Compute' (n 21) 43.

[95] David A Koplow, 'What Are You Lookin' at? Aerial and Space Observation for Arms Control' (2021) 115 American Journal of International Law 89.

[96] Andrew J Coe and Jane Vaynman, 'Why Arms Control Is So Rare' (2020) 114 American Political Science Review 342.

[97] ibid 343. (arguing that this theret trade-off has driven Iraq's nuclear weapons programs after the Gulf War, great power competition in arms in the interwar period, and superpower military rivalry during the Cold War, noting that this accounts for almost 40% of all global arming in the past 2 centuries).

[98] ibid 352.

during the first round of negotiations over the Intermediate Nuclear Forces (INF) Treaty (1980-1983), US negotiators initially proposed 'anytime anywhere inspections', which were rejected not only by the USSR but also, before long, by many on their own side.[99] Instead, during the second round of negotiations (1985-1987), the US put forward proposals for more limited inspections with access to only one missile production facility. Moreover, the US devised a way to outfit Soviet facilities with a sensor that would reveal whether an exiting missile was of the banned type, without revealing the technical characteristics of non-banned missiles.[100] Since the Cold War, there has been continued scientific research into new avenues to, for instance, validate nuclear warheads that are slated for retirement, in ways that do not reveal key engineering features.[101]

Indeed, such applications might be recursively applied to the governance of AI itself: while it has been argued that there is historical precedent for global arms control regimes for high-stakes emerging technologies,[102] there remains uncertainty of the viability of monitoring such regimes effectively.[103] However, scholars have recently begun to propose various technical 'sustained verification' mechanisms which could facilitate compliance monitoring for AI arms control regimes, including anti-tamper techniques such as cryptographically hashed software, as well as the use of hardware-affixed sensors that measure the Van Eck radiation emitted by computing hardware when (AI) code is run, aberrations in which could indicate non-compliant actions.[104]

This highlights the potentially important role that new technologies can play in improving the ability of states and other parties to make verifiable claims about the properties, capabilities, or inputs of new technologies. Technology can therefore not just incrementally contribute to the deterrent or compliant pull of treaties; it can also, in fact, shift the possibility frontier for which arms control or non-proliferation treaties are politically possible in the first place.

Moreover, it is important to note how since the days of the Cold War, the rise of the digital economy has increasingly complemented 'national technical means' with a host of distributed sensors. As a result, over the past decades, the International Atomic Energy Agency (IAEA) has begun to draw on open-source information and commercial satellite imagery in order to provide early warning in nuclear safeguards and to counter nuclear proliferation;[105] commercial satellites have also been proposed to monitor the

---

[99] ibid.

[100] ibid.

[101] For recent developments, see for instance Areg Danagoulian, 'Verification of Arms Control Treaties with Resonance Phenomena' (2020) 30 Nuclear Physics News 25.

[102] Matthijs M Maas, 'How Viable Is International Arms Control for Military Artificial Intelligence? Three Lessons from Nuclear Weapons' (2019) 40 Contemporary Security Policy 285; Waqar Zaidi and Allan Dafoe, 'International Control of Powerful Technology: Lessons from the Baruch Plan' (Center for the Governance of AI, Future of Humanity Institute 2021) <https://www.fhi.ox.ac.uk/wp-content/uploads/2021/03/International-Control-of-Powerful-Technology-Lessons-from-the-Baruch-Plan-Zaidi-Dafoe-2021.pdf>. But see also the critique of current framing- and forum choices strategies used in the pursuit of global regulation of autonomous weapons, in: Elvira Rosert and Frank Sauer, 'How (Not) to Stop the Killer Robots: A Comparative Analysis of Humanitarian Disarmament Campaign Strategies' (2020) 0 Contemporary Security Policy 1.

[103] See also Erica D Borghard and Shawn W Lonergan, 'Why Are There No Cyber Arms Control Agreements?' (*Council on Foreign Relations*, 16 January 2018) <https://www.cfr.org/blog/why-are-there-no-cyber-arms-control-agreements> accessed 22 January 2018; Michael J Glennon, 'The Dark Future of International Cybersecurity Regulation' (2013) 6 Journal of National Security Law & Policy 563.

[104] Matthew Mittelsteadt, 'AI Verification: Mechanisms to Ensure AI Arms Control Compliance' (Center for Security and Emerging Technology 2021) 18–23 <https://live-cset-georgetown.pantheonsite.io/research/ai-verification/> accessed 18 February 2021.

[105] Giacomo GM Cojazzi and others, 'Collection and Analysis of Open Source News for Information Awareness and Early Warning in Nuclear Safeguards' (2013) 50 ESARDA Bulletin 94; David Albright, Sarah Burkhard and Allison Lach, 'Commercial Satellite Imagery Analysis for Countering Nuclear Proliferation' (2018) 46 Annual Review of Earth and Planetary Sciences 99.

implementation of the Kyoto Protocol's Global Emissions Trading Program.[106] Moreover, as noted by Livingston & Risse, the emergence and ubiquity of twenty-first century digital technologies, including mobile phones, commercial high-resolution imaging satellites and social media has already begun to enable near-constant surveillance not just by states or 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.[107]

Such open-source technologies have begun to play a key role in monitoring various war crimes, as well as human rights violations.[108] In at least one case, this has led to an indictment under international criminal law: in 2017, digital sleuths identified Libyan execution sites, by triangulating geographical features found in propaganda videos that had been posted to social media, leading to an International Criminal Court warrant for the arrest of Mahmoud Mustafa Busayf Al-Werfalli, a Libyan warlord.[109] While such analysis is predominantly done by humans at present, it involves distinct tasks that could be automated or enhanced by AI tools. Indeed, one recent experiment involved the development and testing of a machine learning algorithm, trained on synthetic data, to go through footage of airstrikes in order to help investigators identify UK-manufactured cluster munitions that are illegally used in conflict; in tests, this system sped up analysis 100-fold, while reducing risk of trauma for investigators.[110] In another project, run by Amnesty International and the Citizen Evidence Lab, researchers deployed machine learning to support the large-scale analysis of satellite data, in order to detect the destruction of human settlements in Sudan's Darfur region.[111] There are various other ways machine learning tools could contribute to the detection or investigation of various

---

[106] Allison F Gardner, 'Environmental Monitoring's Undiscovered Country: Developing a Satellite Remote Monitoring System to Implement the Kyoto Protocol's Global Emissions-Trading Program' (2000) 9 New York University Environmental Law Journal 152.

[107] Livingston and Risse (n 54) 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' (Carr Center for Human Rights Policy 2018) <https://carrcenter.hks.harvard.edu/files/cchr/files/ccdp_2018_003_humanrightsdocumentation.pdf> accessed 26 August 2020. On the 'multi-use' of remote sensing satellites, see also Nathan Edward Clark, 'Blurred Lines: Multi-Use Dynamics for Satellite Remote Sensing' (2019) 10 Journal of International Humanitarian Legal Studies 171.

[108] Ella McPherson, Isabel Guenette Thornton and Matt Mahmoudi, 'Open Source Investigations and the Technology-Driven Knowledge Controversy in Human Rights Fact-Finding' in A Koenig, S Duberley and D Murray (eds), *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation and Accountability* (Oxford University Press 2019).

[109] Boutin (n 49).

[110] Nesta, 'Documenting Mass Human Rights Violations through Collective Intelligence' (*nesta*, 2020) <https://www.nesta.org.uk/feature/collective-intelligence-grants/documenting-mass-human-rights-violations-through-collective-intelligence/> accessed 29 June 2020; Karen Hao, 'Human Rights Activists Want to Use AI to Help Prove War Crimes in Court' [2020] *MIT Technology Review* <https://www.technologyreview.com/2020/06/25/1004466/ai-could-help-human-rights-activists-prove-war-crimes/> accessed 29 June 2020.

[111] Milena Marin, Freddie Kalaitzis and Buffy Price, 'Using Artificial Intelligence to Scale up Human Rights Research: A Case Study on Darfur' (*Amnesty International & Citizen Evidence Lab*, 6 July 2020) <https://citizenevidence.org/2020/07/06/using-artificial-intelligence-to-scale-up-human-rights-research-a-case-study-on-darfur/> accessed 26 August 2020; Julien Cornebise and others, 'Witnessing Atrocities: Quantifying Villages Destruction in Darfur with Crowdsourcing and Transfer Learning' (2018) <https://aiforsocialgood.github.io/2018/pdfs/track1/80_aisg_neurips2018.pdf> accessed 26 August 2020. Notably, they did not publicly share the source code of this project, to avoid it being used to target villages.

violations under international law,[112] or even the early detection and prevention of violence[113]—although this 'AI Turn' in global governance also has its critics.[114]

As such, from a more modest perspective, there may be various roles for machine learning approaches—if appropriately and rigorously vetted—in enforcing international law, and in facilitating the negotiation of agreements that are sensitively dependent on the ability to monitor compliance.[115]

C.   The replacement of international 'law'? Shifts in regulatory modality

Notably, even if AI were to be used to change the '*input*' of international law (e.g., the process of treaty negotiation or adjudication), or strengthen its enforcement, this would not change the nature of the '*output*' of those legal processes (that is, normative rules, to greater or lesser degree backed by sanction). 'Legal automation' would no doubt alter the texture of international law irrevocably—but it would presumably remain *law*. This would not change the core, normative 'regulatory modality' of international law—the instruments through which this system seeks to regulate and change the behaviour of its constituents (whether there are construed narrowly as states, or more broadly, as also including other stakeholders[116]).

However, even in a more modest usage, the deployment of AI systems in the service of international law, or in the measuring of global governance, could *shift the preferences, values, or tool choices* of global regulators and state actors alike in key ways.[117] As such, AI could drive or facilitate a shift towards novel 'modes' of achieving societally desired regulatory outcomes, which no longer rely to the same extent on normative laws. This discussion draws on Lawrence Lessig's theory of the various 'regulatory modalities' of law, social norms, markets, and architecture.[118] For instance, in the domestic law context, Lessig famously chronicled how the architecture of computer 'code' enabled US regulators to emphasize different 'modalities' in the production and enforcement of law on cyberspace. Specifically, it allowed them to reduce their reliance on normative 'law', and instead regulate tech companies in order to embed architectural constraints on online behaviour, in areas such as zoned speech or privacy protection.[119]

---

[112] Maas, 'International Law Does Not Compute' (n 21) 44–45. (reviewing several recent use cases of AI tools in predicting the activities of poachers, the 'Sentry' system providing civilians in the Syrian conflict through advance warning of incoming airstrikes, and in DNA sequencing for use in forensic investigations of war crimes).

[113] Eleonore Pauwels, 'Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention' (Global Center on Cooperative Security 2020) <https://www.globalcenter.org/wp-content/uploads/2020/10/GCCS_AIData_PB_H.pdf>.

[114] Maria Sapignoli, 'The Mismeasure of the Human: Big Data and the "AI Turn" in Global Governance' (2021) 37 Anthropology Today 4.

[115] Cummings and others (n 4) 3. ("Monitoring the outputs of sensors set up to verify compliance with, for instance, a nuclear, chemical or biological arms control treaty might well be a deadening job for human analysts – albeit one that would require significant specialist training and experience. By contrast, a machine learning system set up to do the same would never tire or grow bored of its task. And while it might (especially in the process of learning) flag a large number of 'false positives', by assigning human specialists to oversee and correct it the system's accuracy would quickly increase.").

[116] J Pauwelyn, RA Wessel and J Wouters, 'When Structures Become Shackles: Stagnation and Dynamics in International Lawmaking' (2014) 25 European Journal of International Law 733.

[117] Maas, 'International Law Does Not Compute' (n 21).

[118] Lawrence Lessig, 'The New Chicago School' (1998) 27 The Journal of Legal Studies 661; Lessig, 'The Law of the Horse' (n 32).

[119] Lessig, 'The Law of the Horse' (n 32) 503–505.

Likewise, Roger Brownsword has argued that new technologies introduce a 'double disruption' to law.[120] In the first place, they affect the substance of legal rules. 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 management'.[121] Accordingly, Brownsword has chronicled how emerging digital technologies of social control may even shift the core 'regulatory attitude' of authorities, and effect a shift away from traditional 'legal coherentism', towards 'regulatory-instrumentalism' or even 'technocracy'.[122] If a similar shift occurred in the context of global governance, this could also drive value shifts. Speculatively, one can imagine the availability of more powerful AI monitoring capabilities would reduce the need for consensus (so long as there was consensus about trusting the monitoring capabilities).

More directly, Deeks has suggested how the proliferation of certain AI technologies in state departments may also shift how states seek to resolve international disputes. It could lead to more unilateral or strategic uses of this technology. For instance, she notes how "[i]f a state could predict in advance which way a tribunal is likely to resolve a particular dispute, it would allow the state to decide whether to pursue the case or to choose an alternative, such as settling it through diplomatic negotiations or dropping the matter entirely."[123]

However, there are also more collaborative cases of such 'replacement'. For instance, there may be situations where algorithms can be used to speed up the negotiation process between states, either as 'negotiation support systems' that can propose potential divisions of interests, or by helping locate third-party proposals.[124] However, in such cases, rather than support traditional channels of international law (e.g. arbitration bodies or international courts) through the 'automation' of their operations, AI tools would contribute to the 'replacement' of such tools as dominant instruments of international conflict resolution or collaboration.

IV.   AI's effects on global regime complexity

Having discussed the sources and vectors of potential AI-driven governance disruption (and especially automation), it is important to discuss not just the direct effects of such legal displacement in international law, but also the indirect effects on the structure and texture of the international legal system.

International legal scholars have for many years reckoned with various new trends and shifts in the global institutional and legal landscape. These include patterns of *institutional proliferation*,[125] the ongoing *fragmentation* of international law resulting in complex inter-regime impacts and externalities,[126] and growing patterns of *contested*

---

[120] Brownsword, 'Law and Technology: Two Modes of Disruption, Three Legal Mind-Sets, and the Big Picture of Regulatory Responsibilities' (n 33) 6–15.
[121] See also Brownsword, 'In the Year 2061' (n 17); Brownsword, 'Technological Management and the Rule of Law' (n 22).
[122] Brownsword, 'Law Disrupted, Law Re-Imagined, Law Re-Invented' (n 48).
[123] Deeks (n 18) 628.
[124] ibid 633–637.
[125] Kal Raustiala, 'Institutional Proliferation and the International Legal Order' in Jeffrey L Dunoff and Mark A Editors Pollack (eds), *Interdisciplinary Perspectives on International Law and International Relations: The State of the Art* (Cambridge University Press 2012).
[126] Martti Koskenniemi and Study Group of the International Law Commission, 'Fragmentation of International Law: Difficulties Arising from the Diversification and Expansion of International Law' (United Nations - General Assembly 2006) A/CN.4/L.682 <http://legal.un.org/ilc/documentation/english/a_cn4_l682.pdf>; Frank Biermann and others, 'The Fragmentation of Global Governance Architectures: A Framework for Analysis' (2009) 9 Global Environmental Politics 14. But see also the counterarguments in Rakhyun E Kim, 'Is Global Governance Fragmented, Polycentric, or Complex? The State of the Art of the Network Approach' [2019] International

multilateralism.[127] Others have identified trends of *legal stagnation*, and have argued that global governance is increasingly marked by a shift towards informality in many issue areas such as international environmental governance or cyberspace.[128]

Let us focus on the 'fragmentation' of international law. Contemporary regime complex theory has explored many drivers of regime complex evolution and development.[129] However, while scholars have reckoned with the effects of exogenous political or institutional shocks to either the trajectory of a regime complex,[130] or to the continued viability of the international liberal order as a whole,[131] there appears to be relatively little examination of the effects of technology-driven change on regime complexity. This is unfortunate because, as the lens of governance disruption has shown, the use of various technologies can in some cases exert considerable effects on the norms, processes, or political scaffolding of the international legal order.

As such, it may be productive to explore the implications of AI-driven governance disruption lens for the processes of regime complex evolution. Such examination is certainly more speculative, and it should be re-emphasised that the following sections are not meant as strong predictions, but as conditional scenarios that anticipate the implications of various types of AI-driven governance disruption on regime complexity.

For the sake of argument, this following assumes no- or only very modest further capability developments in the available AI capabilities,[132] although it does weakly assume the gradual continuing dissemination or application of existing AI capabilities into more areas and domains, in order to anticipate the implications of distinct patterns of dissemination and usage on regimes.

As such, it is argued that we can distinguish between 10 different roles or effects of AI systems on the international legal order. By requiring constant substantive legal development, AI systems can reinforce pre-existing governance trends; it can serve as (1) legal '*canary* in the coal mine', highlighting the need for greater harmonization or cross-regime dialogue. However, in already-fragmented settings, it can serve as (2) *tough knot* or (3) *generator of regime fault lines*, potentially increasing fragmentation. Under even modest scenarios of legal displacement (whether automation or replacement), AI systems can serve variably as a (4) *shield*, (5) *patch*, (6) *cure*, or (7) *accelerator* of the fragmentation of international law. Finally, AI tools may serve as (8) *differential enabler*; (9) *driver of*

Studies Review <https://academic.oup.com/isr/advance-article/doi/10.1093/isr/viz052/5571549> accessed 16 February 2020; Mette Eilstrup-Sangiovanni, ''Ordering' Global Governance Complexes: The Evolution of the Governance Complex for International Civil Aviation'
<https://www.repository.cam.ac.uk/handle/1810/313926> accessed 13 December 2020.

[127] Julia C Morse and Robert O Keohane, 'Contested Multilateralism' (2014) 9 The Review of International Organizations 385; Amitav Acharya, 'The Future of Global Governance: Fragmentation May Be Inevitable and Creative Global Forum' [2016] Global Governance 453; Michael Zürn, 'Contested Global Governance' (2018) 9 Global Policy 138.

[128] Pauwelyn, Wessel and Wouters (n 117). On cyberspace governance, see also Joseph S Nye, 'The Regime Complex for Managing Global Cyber Activities' (Global Commission on Internet Governance 2014) 1 <https://dash.harvard.edu/bitstream/handle/1/12308565/Nye-GlobalCommission.pdf> accessed 3 September 2019.

[129] Karen J Alter and Kal Raustiala, 'The Rise of International Regime Complexity' (2018) 14 Annual Review of Law and Social Science 329; Laura Gómez-Mera, Jean-Frédéric Morin and Thijs Van De Graaf, 'Regime Complexes' in Frank Biermann and Rakhyun E Kim (eds), *Architectures of Earth System Governance: Institutional Complexity and Structural Transformation* (Cambridge University Press 2020).

[130] Hanzhi Yu and Lan Xue, 'Shaping the Evolution of Regime Complex: The Case of Multiactor Punctuated Equilibrium in Governing Human Genetic Data' (2019) 25 Global Governance: A Review of Multilateralism and International Organizations 645.

[131] See Karen J Alter, 'The Future of International Law' (2017) 101 iCourts Working Paper Series <https://papers.ssrn.com/abstract=3015177> accessed 11 June 2020. (examining the prospects for the global liberal order if the US turns away from its values).

[132] However, this may be an overtly pessimistic assumption that may well be proven wrong by continued progress in the coming years.

value shifts, or (10) *asymmetric weapon*—which could potentially contribute to patters of governance destruction or further erosion in the international legal order.

A.   Fragmentation: AI as a legal canary, tough knot & as generator of regime fault lines

At a structural level, how will AI's ability to generate frequent legal uncertainty—and a resulting need for constant *development* or realignment in the substance, laws or norms of international law—affect the trajectory of a regime complex?

One key caveat here is that global governance obviously consists of a wider range of norms and governance instruments than solely those embedded in writ 'hard law'. Critically, many of these softer governance instruments are potentially more resilient to 'legal' disruption—because they are already phrased broadly, depend on discretion for implementation, or are more flexible and easy to review and update in the face of changing technological circumstances.[133] Nonetheless, even if this means that such soft governance instruments are slightly better insulated from—or at least better able to adapt to—AI-driven 'legal' shocks, the conceptual and practical changes forced by the use of these systems still bear on other forms of governance, and may (or arguably should) warrant re-examination.

Taking this into account, the prospect of AI systems driving legal development could, paradoxically, exert two divergent effects on regime complexes. On the one hand, situations of clear inadequacy of existing legal instruments or concepts might serve as an opportunity to clarify long-present ambiguities in established global norms, and harmonize latent conflicts between existing regimes, by getting these out in the open, in the context of an urgent governance problem that requires addressing. In this way, AI systems—or particular legal problems or puzzles they flag—can highlight long-standing tacit tensions between existing norms and regimes, and in doing so can serve as a (1) '*canary in the coalmine*' of the integrity of the international legal system in the digital age.[134] More broadly, the broad-spectrum legal disruption that is generated by AI systems might illustrate the confusing and fragmenting effects of 'modernity', which, as noted, some regime systems theorists point to as a major factor in the trend towards fragmentation and regime complexity.[135] In this way, the precise effects of AI-driven governance development on the trajectory of a governance architecture will likely depend sensitively on the pre-existing configuration of that architecture. Indeed, it some ways, it could exacerbate pre-existing tendencies.

On the one hand, within an already-integrated regime complex, a centralised, authoritative international institution could in principle seize upon AI's legal disruption within one area in order to kick-start and facilitate a broader dialogue about long-overdue systemic revision or legal innovation in underlying rules or concepts. For instance, such an institution could seize upon specific problems posed by AI-enabled military surveillance platforms or 'lethality-enabling technologies',[136] to address the broader fact

---

[133] This is one reason why Crootof considers 'soft law' a potentially promising alternate avenue for the international regulation of certain new technologies, especially when compared to the propensity for hard-law treaties to be rendered obsolete by advancing technology, and the problems involved with then amending them. Crootof, 'Jurisprudential Space Junk' (n 23) 124–126.

[134] The terminology of 'canaries' is inspired by the more specific usage of 'artificial canaries' in: Carla Zoe Cremer and Jess Whittlestone, 'Artificial Canaries: Early Warning Signs for Anticipatory and Democratic Governance of AI' (2021) 6 International Journal of Interactive Multimedia and Artificial Intelligence 100.

[135] See also Jean-Frédéric Morin and others, 'How Informality Can Address Emerging Issues: Making the Most of the G7' (2019) 10 Global Policy 267.

[136] Arthur Holland Michel, 'The Killer Algorithms Nobody's Talking About' (*Foreign Policy*, 20 January 2020) <https://foreignpolicy.com/2020/01/20/ai-autonomous-weapons-artificial-intelligence-the-killer-algorithms-nobodys-talking-about/> accessed 21 January 2020. For a recent report that does seek to explore

that the distinction between 'war' and 'peacetime'—a cornerstone of IHL—has, in practice already become blurred by new technologies.[137] Such legal development would not necessarily imply an abandonment of these well-proven frameworks; but it would reckon with the 'transversal' effects of trends in technology across different fields of (international) law, in order to yield more integrated bodies of law. Of course, even if centralised international institutions could in principle carry out such regulatory innovation and integration, it is not a given that they would also do so.[138]

On the other hand, starting from an already-fragmented institutional context—and across fragmented application domains—AI technology's propensity to generate situations of legal disruption (requiring development) may well exacerbate the forces of (further) fragmentation, as these systems become (2) *tough knots*. Given the extremely diverse set of perspectives through which one can approach AI technology, it seems unlikely that a fragmented regime complex would reliably be able to organically converge towards more integrated or harmonised policy responses. Similar AI architectures will be used across widely different contexts; in each of those issue areas, certain self-similar features of AI systems (such as algorithmic unpredictability or opacity, or the susceptibility to adversarial input) can be refracted into seemingly-distinct local problems.

If many distinct institutions and organisations are forced to grapple, in parallel and relative isolation, with certain local questions (such as the variable meaning of 'Meaningful Human Control' for LAWS, in healthcare, and in other contexts)) of underlying conceptual or legal disruption, it is likely that we may see distinct institutions and regimes resolve and decide these cases in different, and potentially contradictory ways.[139] In that way, the proliferation of AI systems could serve as a (3) *generator for regime fault-lines,* triggering a scattering of individual regimes' norms and policies, in ways that drive regime complexity, and that open up considerable scope for conflicts in terms of regime norms, operations, or impact.

B.   Automation: AI as shield, patch, cure or accelerator of fragmentation

As one subset of governance disruption, 'displacement' also highlights ways in which the integration or incorporation of AI tools into the practices of global governance can affect regime complexity or the coherence of international law. In brief, some AI tools, especially if they were accessible (or distributed) to many actors (whether states or non-state observer agencies or NGOs), could play some role in mitigating at least some trends towards regime complexity. They could variably serve as *shield* to the negative consequences of complexity; as *patch* to halt the further fragmentation of regimes, or as *cure* to avert or mitigate latent conflicts and aid in harmonisation of regimes. Yet, left unregulated, the free-for-all use of AI systems could also prove an *accelerator* of processes of legal and governance fragmentation.

---

the question of arms control, not just for LAWS, but for broader uses of AI in military decision-making, see Giacomo Persi Paoli and others, 'Modernizing Arms Control: Exploring Responses to the Use of AI in Military Decision-Making' (UNIDIR 2020) <https://unidir.org/publication/modernizing-arms-control>.

[137] Denise Garcia, 'Future Arms, Technologies, and International Law: Preventive Security Governance' (2016) 1 European Journal of International Security 94; Braden Allenby, 'Are New Technologies Undermining the Laws of War?' (2014) 70 Bulletin of the Atomic Scientists 21 (highlighting drones and cyberwarfare). See also Crootof, 'Regulating New Weapons Technology' (n 58) 10. And see generally Rosa Brooks, 'War Everywhere: Rights, National Security Law, and the Law of Armed Conflict in the Age of Terror' (2004) 153 University of Pennsylvania Law Review 675.

[138] Indeed, the question of whether, or under what conditions, international organizations would be capable of this, remains an open and interesting one.

[139] Crootof and Ard (n 29).

To start, (4) various actors might use AI tools to *shield* themselves from the negative operational consequences of a fragmented and distributed regime complex. For instance, the use of these tools in support of many 'routine' diplomatic tasks could free up the limited diplomatic resources or staff available to many smaller states, enabling them to participate more fully in various international fora, 'levelling' the playing field relative to the large and well-staffed foreign ministries of larger states.[140] More modestly, through translation and text (e.g. news) summarisation services, such AI tools could facilitate the ability of even less powerful actors to navigate dense regimes complexes, counteracting the democratic deficit identified by some concerned scholars.[141]

More actively, actors could also AI tools as a (5) *patch* to halt or pre-empt the further fragmentation of various regimes. For instance, Ashley Deeks notes how various 'text-as-data-tools' might be used to pre-emptively identify treaty conflicts (whether accidental or strategically engineered). She suggests that "[a] state could create a tool that allows it to compare proposed treaty language (either while the negotiations are underway, or before the state has ratified the treaty) to all other treaties to which the state is a party to detect similar language or very similar topics."[142] This could pre-empt or deter strategic efforts by states to create certain treaty conflicts in order to set the agenda for future negotiations,[143] because any negotiation partners would now be able to spot when such a strategy is being attempted.[144]

In principle, such AI tools could not only help spot and patch imminent conflicts produced by specific treaties, but could also be deployed as (6) a modest *cure* to more general pre-existing patterns of fragmentation in international law. For instance, Deeks suggests that states could use such tools even in the absence of specific negotiations, simply to map any tensions amongst their existing treaty commitments, "in order to identify gaps or other potential conflicts before they produce a real world problem."[145] Accordingly, other actors or international organisations could use such tools to chart, to some degree, major fault lines or gaps amongst international law's fragmented regimes. This would of course hardly be a panacea: questions of sufficient text data aside, it should be kept in mind that not all (or even most) practical regime conflicts emerge from inconsistent norm commitments recorded in the legal text.[146] Indeed, as noted by Morin and others, "[b]latant legal conflicts [amongst regimes] remain rare and a certain degree of normative ambiguity preserves the unity of the international legal system."[147] Instead, many regime conflicts may manifest at the level of impacts or institutional policies,[148] as a result of real-world inter-institutional interaction, and such negative externalities are not latent in the texts, and therefore not discoverable with such tools.

---

[140] Katharina E Höne, 'Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy' (DiploFoundation 2019) <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> accessed 14 March 2019.

[141] Daniel W Drezner, 'The Power and Peril of International Regime Complexity' (2009) 7 Perspectives on Politics 65; Eyal Benvenisti and George W Downs, 'The Empire's New Clothes: Political Economy and the Fragmentation of International Law' (2007) 60 Stanford Law Review 595.

[142] Deeks (n 18) 616.

[143] Surabhi Ranganathan, *Strategically Created Treaty Conflicts and the Politics of International Law* (Cambridge University Press 2014) <https://www.cambridge.org/core/books/strategically-created-treaty-conflicts-and-the-politics-of-international-law/55EACC81A929FC19216E3380D2E9DF69> accessed 18 June 2020.

[144] Deeks (n 18) 616. ("[a] state that is concerned that its negotiating partner may be attempting this move could use web scraping and topic analysis on its negotiating partner's existing treaties to assess whether the partner is trying to play this game").

[145] ibid.

[146] Christian Kreuder-Sonnen and Michael Zürn, 'After Fragmentation: Norm Collisions, Interface Conflicts, and Conflict Management' (2020) 9 Global Constitutionalism 241.

[147] Morin and others (n 136) 2–3.

[148] Thomas Gehring and Sebastian Oberthür, 'The Causal Mechanisms of Interaction between International Institutions' (2009) 15 European Journal of International Relations 125.

26

However, while these uses seem to connote beneficial (or at least, integrative) outcomes of AI displacement on regime complexes, there are many scenarios under which AI tools could also serve as (7) an *accelerator* of regime fragmentation and complexity. After all, as has been discussed, many AI tools could also be used to challenge or subvert international legal processes, in ways that could speed up contestation. Even if AI technology finds productive uptake by many states to facilitate negotiations, this can have unanticipated side effects. After all, as Crootof has noted, speeding up the development of new international legal obligations—by any means—may certainly be necessary to address urgent problems, or help States avoid or resolve disputes; however it also "facilitates legal fragmentation [because] [s]peeding up the development of international legal obligations has expanded both the activities they regulate and opportunities for conflict among them."[149]

As always, of course, the deployment of these technologies is likely to have unanticipated downstream normative, functional, and legal effects. Eyal Benvenisti has argued that the increasing data-processing and sharing-capabilities of modern digital technologies, and the growth of modalities of governance that rely on decision-making by machines (rather than two-way exchange of stakeholders) as their input, are creating a foundational challenge to long-established principles in global administrative law, that "the more communication, the better", and that information exchange necessarily promotes institutional accountability.[150] All this is not a reason to avoid such tools, but should induce some caution about ensuring appropriate scaffolding or guidelines regarding their global use.

C.   Contestation: AI as differential enabler, driver of value shifts, or asymmetric weapon

Amongst leading state developers, perceptions of an emerging technology's very high strategic stakes—whether or not these views are justified—may inhibit willingness to even come to the negotiating table in good faith.[151] Moreover, AI's inherent definitional complexity and effects across many fields, might lend themselves naturally to strategies that seek to obstruct international regulatory action by dragging out debate at international fora. The dual-use nature of many 'inputs' of AI capability—computing hardware; training data; talent—might furthermore make international bans of certain applications difficult to monitor or enforce.[152]

All of this could result in notable AI issues remaining unsolved, in spite of them featuring (or being raised) on the international agenda. The clear inability of the existing governance arrangement to address these problems would drive processes of erosion in their legitimacy. More broadly, if AI tools can serve as a (8) *differential enabler* that is better suited to the interests of illiberal than liberal actors, or to powerful over the powerless actors, this can further exacerbate power inequalities, eroding the legal fiction of the sovereign equality of states,[153] and with it the legitimacy of the international liberal order. Moreover, for certain powerful actors, it might drive preference changes—desires to pursue other avenues of governance. In many other cases, this could drive 'voice and

---

[149] Crootof, 'Regulating New Weapons Technology' (n 58) 7.

[150] Eyal Benvenisti, 'Upholding Democracy Amid the Challenges of New Technology: What Role for the Law of Global Governance?' (2018) 29 European Journal of International Law 9.

[151] Picker (n 19).

[152] Although see the discussion of using various institutional, software or hardware mechanisms for arms control verification, Infra at III(C).

[153] Deeks (n 18) 580.

representation' goals. Either of these has generally been linked to processes of institutional proliferation and regime complexity.[154]

Simultaneously, AI's use as tool could offer specific capabilities that result in legal decline because it serves as (9) a *driver of value shifts*, altering the endogenous character of that order. This could be because AI systems might offer strategic capabilities which shift interests; they chip away at the rationales for certain powerful states to engage fully in, or comply with, international law regimes. While they would be unlikely to completely erode support, AI applications in areas such as surveillance could arguably reduce states' (perceived) dependence on multilateral security regimes to ensure their security from terrorist threats. Such capabilities could increase actor's willingness to 'defect', erode their willingness to support multilateralism, and could as such result in *preference changes* and the resulting increase in the prominence of forum-shopping strategies. That could produce a harmful 'non-regime-state' around many topics of AI.

In extreme cases, AI systems could enable new strategies—such as scalable computational propaganda—by which actors could directly challenge the legitimacy of international law or its component regimes.[155] This potential use as an (10) *asymmetric weapon* might enable the direct contestation and erosion of the liberal international order, more easily than its reinvigoration. AI has not created these particular problems, they it make them worse.

Given the above, we have discussed how the use of AI can shape both the normative and conceptual coherence of the regime complex (through driving *development* or *displacement*), as well as by shifting the incentives, values, or behaviour of various actors within that regime complex (potentially driving *destruction*). From a regime complex perspective, then, the key takeaway is that the governance disruptive impacts of AI may compound (1) trends towards regime fragmentation as a result of increased legitimacy problems or preference shifts; and (2) the problematic interactions of regimes (by increasing the number of contact surfaces).

To be sure, the above discussion of trajectories remains speculative, given the early state of AI governance at present. All this is not to make strong predictions, but rather to identify a range of ways in which AI applications, through governance disruption, could intersect with processes of regime complexity. In aggregate, these current trends and drivers might loosely suggest that, barring major interventions or external shocks, AI governance may remain fragmented for the time being.[156] However, this is certainly not a foregone conclusion, and it will be key to monitor developments both in the AI regime complex, as well as developments in AI's effects on and in international law, in order to identify in greater detail the overall arc of governance development.

V.   Responses to AI-driven disruption: paths and principles

Given these above challenges, how might the international legal order respond? In the first place, international lawyers and scholars of global governance should beware the unrestricted automation of international law, and actively scrutinize the trend towards 'high-tech international law'.[157]

---

[154] Alter and Raustiala (n 130).

[155] As discussed in Maas, 'International Law Does Not Compute' (n 21) 55.

[156] Peter Cihon, Matthijs M Maas and Luke Kemp, 'Should Artificial Intelligence Governance Be Centralised?: Design Lessons from History', *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (ACM 2020) 228 <http://dl.acm.org/doi/10.1145/3375627.3375857> accessed 12 February 2020; Thorsten Jelinek, Wendell Wallach and Danil Kerimi, 'Coordinating Committee for the Governance of Artificial Intelligence' (G20 - Policy Brief Taskforce 5 (Multilateralism) 2020) 2, 4 <https://www.g20-insights.org/policy_briefs/coordinating-committee-for-the-governance-of-artificial-intelligence/> accessed 8 July 2020.

[157] Deeks (n 18).

In general, they should be wary of the unregulated dissemination and incorporation of (AI) tools in processes of international lawmaking or adjudication. This is because, under plausible development and deployment conditions, these tools may remain largely restricted to major states (at least in the first instance), such that unrestricted usage is likely to further exacerbate power imbalances amongst leading AI states and smaller states, potentially feeding dissatisfaction and contestation.[158] Even in more unambiguously cooperative uses in support of conflict resolution, the use of such AI tools to speed up law-creation could accelerate overall processes of fragmentation, and therefore should be carefully monitored.

In particular, this lens might urge scrutiny of AI applications which could, directly or indirectly, shift the 'regulatory modality' away from international conflict resolution through legal norms or processes such as arbitration. The availability of certain AI tools such as population sentiment analysis, computational propaganda, or lie-detection systems might lead some states to seek to increasingly resolve certain foreign policy conflicts through bilateral state channels—or even through unilateral technological intervention—and only to submit for arbitration those issues where they predict they will win.[159] Such shifts could well erode the broader standing or relevance of the global legal order.[160] As such, they should be cautiously monitored and held accountable.

At the same time, while a general degree of caution towards international legal automation may be warranted, it should also be remembered that AI-driven displacement offers many meaningful opportunities to increase the efficacy of the global governance architecture, its resilience to future technological disruption as well as other shocks, and even potentially its legitimacy.

This speaks in favour of accelerating the development and—subject to rigorous and careful assessment and constraints—dissemination of specifically vetted AI tools that could help improve structural conditions for international cooperation. There are diverse applications of AI that could in principle help shift the international 'cooperation-conflict balance'. These could include AI monitoring capabilities that help provide greater assurance of the detection of state noncompliance, and in so doing help resolve situations of pervasive governance gridlock.[161]

Alternatively, technological and institutional interventions could increase the ability of various parties to make verifiable claims about their own AI capabilities, providing stronger foundations for AI-focused arms control regimes.[162] Moreover, various technological interventions (including but not limited to AI tools) could support cooperation and coordination more generally, such as systems that could enable states to reach agreement on issues more quickly.[163] Finally, from the perspective of regime complexity, relatively simple AI language models could allow the advance identification of emergent or imminent norm or treaty conflicts, whether accidental or strategically engineered.[164] Over time, they could also enable a greater mapping and monitoring of fragmentation both in the AI regime complex,[165] as well as in the broader governance

---

[158] ibid 644–646.

[159] ibid 628–629.

[160] However this could exacerbate the fragmentation of international law, adding a (vertical) dimension of fragmentation to the existing inter-regime conflicts.

[161] These could dissolve or at least reduce the 'transparency-security' trade-off, and facilitate 'closed' arms control deals, as discussed above.

[162] Miles Brundage and others, 'Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims' [2020] arXiv:2004.07213 [cs] 67–69 <http://arxiv.org/abs/2004.07213> accessed 16 April 2020.

[163] Deeks (n 18) 647–648.

[164] ibid 616.

[165] For instance, others and myself have elsewhere called for the use of various Natural Language Processing-tools to improve the monitoring of trends of conflict, coordination or catalyst in the AI regime complex. Cihon, Maas and Kemp (n 157) 233.

architecture, supporting processes of convergence. It would be valuable to offer such AI applications free or low-cost to many parties, and teach them how to deploy and use such systems.[166] Such applications could both increase the inclusion of many actors in the fora of international governance,[167] as well as help in the management of governance fragmentation more broadly.

Indeed, Dafoe and colleagues have sketched an open research agenda into how AI tools could potentially strengthen cooperation across many contexts; they identify a wide range of AI-supported mechanism design interventions that could improve critical 'cooperative capabilities', such as parties' *understanding* of the world, their *communication* abilities, their capacity to make credible *commitments*, or the *institutions* (such as norms or regimes) that shape and structure their decision environment.[168] However; the authors also caution that AI-supported cooperation systems could have potential downsides, noting that "(1) Cooperative competence itself can cause harms, such as by harming those who are excluded from the cooperating set and by undermining pro-social forms of competition (i.e., collusion). (2) Advances in cooperative capabilities may, as a byproduct, improve coercive capabilities (e.g., deception). […] (3) Successful cooperation often depends on coercion (e.g., pro-social punishment) and competition (e.g., rivalry as an impetus for improvement)".[169]

As such, AI's benefits will not manifest by default; and realizing these goods while avoiding these risks will require active and critical engagement by not just international lawyers, but also a range of actors across the global community at large.

VI.   Conclusion

This paper has sought to explore, at an initial level of detail, how we can analyse the direct and indirect effects of AI technologies on the viability, form, or functioning of the international liberal legal order.

It first provided a functional definition of AI, and a rationale for why we should expect this technology to have wide-ranging impacts on international law. It then explored the pedigree of these questions in existing scholarship on 'law, regulation and technology', and specifically the study of legal automation (LawTech), and studies of the interrelation between technological change and legal systems more broadly (TechLaw).

The paper then introduced a 'Governance Disruption' framework for exploring the international legal impacts of AI. This taxonomy distinguishes situations where AI applications create a need for international legal *development;* and situations where they can contribute to the potential *destruction* of specific regimes or the erosion of the international legal order. The paper then focused in specifically on the vectors of legal *displacement* (the potential effects of the 'automation' of international law), discussing the viability, form, and implications of three potential pathways of displacement: automation of rule creation; automation of monitoring & enforcement; or the 'replacement' of international law. The paper then sketched the *effects* of these trends on the integration or fragmentation of international law, distinguishing 10 different roles that AI systems may play. By requiring constant substantive legal development, AI systems can reinforce

---

[166] Deeks (n 18) 650–652.

[167] Höne (n 141).

[168] Allan Dafoe and others, 'Open Problems in Cooperative AI' [2020] arXiv:2012.08630 [cs] <http://arxiv.org/abs/2012.08630> accessed 5 January 2021. On the problems and challenges of AI systems creating structural shifts, see also Remco Zwetsloot and Allan Dafoe, 'Thinking About Risks From AI: Accidents, Misuse and Structure' (*Lawfare*, 11 February 2019) <https://www.lawfareblog.com/thinking-about-risks-ai-accidents-misuse-and-structure> accessed 12 February 2019; Agnes Schim van der Loeff and others, 'AI Ethics for Systemic Issues: A Structural Approach' (2019) <http://arxiv.org/abs/1911.03216> accessed 13 January 2020.

[169] Dafoe and others (n 169) 30.

pre-existing governance trends; it can serve as (1) legal '*canary* in the coal mine', highlighting the need for greater harmonization or cross-regime dialogue. However, in already-fragmented settings, it can serve as (2) *tough knot* or (3) *generator of regime fault lines*, potentially increasing fragmentation. Under even modest scenarios of legal displacement (whether automation or replacement), AI systems can serve variably as a (4) *shield*, (5) *patch*, (6) *cure*, or (7) *accelerator* of the fragmentation of international law. Finally, AI tools may serve as (8) *differential enabler*; (9) *driver of value* shifts, or (10) *asymmetric weapon*—which could potentially contribute to patters of governance destruction or further erosion in the international legal order. It was argued that an examination of possible trajectories for global governance regimes will have to increasingly calculate in the effects of AI-driven governance disruption on regime integrity: while some AI capabilities could see use in countering regime fragmentation or treaty conflicts, some AI capabilities could also serve as generators of new legal-conceptual fault lines, or spur increased patterns of contestation

In the face of these effects, I argued that international legal scholars have an important role to play in critically surveying this new technology and its effects on the international legal system.