

ADVANCING INTELLIGENCE AND GLOBAL SOCIETY: INTERNATIONAL LAW'S ROLE IN GOVERNING THE ADVANCE OF ARTIFICIAL INTELLIGENCE

*Lesley Nash*¹

INTRODUCTION

Advancing technology changes the fabric of global society, from electricity to the rise of social media, yet law has always struggled to keep pace with such technological advances,² and this problem has accelerated with the increased pace of technological change in the twentieth and twenty-first centuries.³ Although modern society is faced with a multitude of issues springing from new technologies, law and governance structures have lagged⁴ while the technologies themselves have advanced.⁵ This discrepancy between technology and law is particularly glaring in the artificial intelligence (“AI”) field. Though the solution to creating “true” artificial general intelligence is still elusive,⁶ “weak” forms of the technology are altering the fabric of society, from social media⁷ to the orchestration of war.⁸ These technological advances remake everyday existence, yet global regulatory functions are not sufficiently robust to oversee these changes.

The fact that international law has *not yet* exhibited meaningful regulatory control over artificial intelligence technology does not mean that it *cannot*.

¹ J.D. expected 2020, University of Kentucky College of Law; M.A. 2017, University of Kentucky Patterson School of Diplomacy.

² See, e.g., *Olmstead v. United States*, 277 U.S. 438 (1928), *overruled* by *Katz v. United States*, 389 U.S. 347 (1967), and *Berger v. New York*, 388 U.S. 41 (1967) (holding that warrantless wiretapping by law enforcement did not violate the fourth or fifth amendment); *Katz v. United States*, 389 U.S. 347 (1967) (tracing the evolution of Fourth Amendment protections against “unreasonable searches and seizures” as they relate to electronic wiretaps), *discussed* by Nicandro Iannacci, *Katz v. United States: The Fourth Amendment adopts to new technology*, NAT’L. CONST. CTR. (Dec. 18, 2018), <https://constitutioncenter.org/blog/katz-v-united-states-the-fourth-amendment-adapts-to-new-technology> [<https://perma.cc/7VKB-5H3Y>].

³ Vivek Wadhwa, *Laws and Ethics Can’t Keep Pace with Technology*, MIT TECH. REV. (Apr. 15, 2014), <https://www.technologyreview.com/s/526401/laws-and-ethics-cant-keep-pace-with-technology/> [<https://perma.cc/K9Q5-NHRF>] (“These regulatory gaps exist because laws have not kept up with advances in technology. The gaps are getting wider as technology advances...”).

⁴ *Id.* (“We haven’t come to grips with what is ethical, let alone what the laws should in be in relation to [such] technologies . . .”).

⁵ *Id.* (“Today, technology is on an exponential curve... changes of a magnitude that once took centuries now happen in decades, sometimes in years.”).

⁶ MARGARET A. BODEN, *ARTIFICIAL INTELLIGENCE: A VERY SHORT INTRODUCTION* 19 (2018).

⁷ John Ellett, *New AI-Based Tools Are Transforming Social Media Marketing*, FORBES (July 27, 2017, 6:00 AM), <https://www.forbes.com/sites/johnellett/2017/07/27/new-ai-based-tools-are-transforming-social-media-marketing/#162c713369a2> [<https://perma.cc/43FL-97MH>].

⁸ See PAUL SCHARRE, *ARMY OF NONE: AUTONOMOUS WEAPONS AND THE FUTURE OF WAR* 45 (2018) (“At least thirty nations currently employ supervised autonomous weapons systems of various types to defend ships, vehicles, and bases from attack.”).

International law offers a structure of governance over issues that are too broad for unilateral state regulation, or that implicate international interests.⁹ Although international law has often been denigrated as weak,¹⁰ it has great potential to offer solutions for global problems that are too large for states to tackle alone. This note seeks to unveil the benefits of using international law to approach the problems and potential of AI as well as suggest a possible method of doing so that can increase regulation and aid in the development and advancement of safe AI technology.

Part I will provide a succinct overview of the current state of artificial intelligence, including the various types of autonomous states of technologies. Section A will include a discussion of the definitions of relevant technologies as well as their modern uses. Section B will touch briefly on several examples of the legal and regulatory issues that have arisen from this technological paradigm. Section A of Part II will discuss the current state of international legal and regulatory structures, while Section B will consider how international law might provide regulation and oversight of this advancing technological sector. Section C of Part II will examine several national and international policies regulating artificial intelligence and what lessons can be drawn from existing structures. Section A of Part III will then draw from institutions and structures offered in Part II, Section B, as well as best practices considered in Section C. Here, I will argue that international law offers the best path forward to functional oversight, regulation, and promotion of advancing AI technologies, and will propose a framework for such an international regulatory structure. Section B of Part III will briefly answer questions related to why international law is not already in use. Finally, Part IV will conclude with a few remarks about both the potential and danger inherent in advancing AI technology and reiterate the call for international regulation and oversight.

I. ARTIFICIAL INTELLIGENCE IN THE MODERN WORLD

A. *Understanding Artificial Intelligence: Definitions and Current State of Technology*

AI is ubiquitous in popular culture,¹¹ but the reality of the technology is much different than popular imaginings. AI can be divided into two general categories: “weak” or “specific” AI and “strong” or “general” AI.¹² “Weak” or “specific” AI is an application or system with a specific function, in which the AI often “outperform[s] even the most expert humans.”¹³ “Strong” or “general” AI (often

⁹ Rüdiger Wolfram, *International Law*, MAX PLANCK ENCYCLOPEDIA OF PUB. INT’L L. ¶16 (last updated Nov. 2006), <http://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1424> [https://perma.cc/4ZWN-LLMA].

¹⁰ See, e.g., John Bolton, *Is There Really Law in International Affairs*, 10 TRANSNAT’L L. & CONTEMP. PROBS. 1, 28–30 (2000).

¹¹ See, e.g., Michael Hogan & Greg Whitmore, *The top 20 artificial intelligence films- in pictures*, THE GUARDIAN (Jan. 8, 2015, 7:29 AM), <https://www.theguardian.com/culture/gallery/2015/jan/08/the-top-20-artificial-intelligence-films-in-pictures> [https://perma.cc/E74F-3TH5].

¹² Kathleen Walch, *Rethinking Weak vs String AI*, FORBES (Oct. 4, 2019, 6:30 AM), <https://www.forbes.com/sites/cognitiveworld/2019/10/04/rethinking-weak-vs-strong-ai/#7da76f676da3> [https://perma.cc/B7YL-YCHK].

¹³ BODEN, *supra* note 6, at 18.

referred to as artificial general intelligence, or AGI), on the other hand, is more akin to the AI of pop culture, where the program or system is not merely “specifically” gifted but rather achieves “human-level” performance across a spectrum of individual challenges that would allow the AI to “think.”¹⁴ Though AGI is not yet realized, researchers have made progress on several fronts related to general intelligence, including visual analysis, object recognition, and behavioral interactions.¹⁵ Specific intelligences, on the other hand, are common, operating as systems that are designed to follow a “special-purpose algorithm,” which may render the program an expert search engine¹⁶ or chess player,¹⁷ but incapable of harnessing human ‘common sense.’

A discussion of AI necessitates one of automation. Paul Scharre notes three degrees of autonomy that are helpful when discussing AI.¹⁸ First, semiautonomous operations are those in which “the machine performs a task and then waits for a human user to take an action before continuing,”¹⁹ or “human in the loop” processes.²⁰ Second, there are supervised autonomous operations in which, once in operation, “the machine can sense, decide, and act on its own, but a human observer can . . . intervene;” or “human on the loop” processes.²¹ Finally, there are fully autonomous operations, in which “systems sense, decide, and act entirely without human intervention;” or “human out of the loop” processes.²² Programs often move among these types of processes when completing a task and they can be conceived of as a continuum: as programs grow more sophisticated, they require less human intervention and oversight to complete tasks.²³

There is also a difference between automatic, automated, and autonomous intelligence in machines. Automatic programs are simple, highly predictable, and display no decision-making qualities.²⁴ Automated programs are more complex, rule-based systems that may consider a range of variables before acting.²⁵ Autonomous programs are sophisticated, goal-oriented, and may be considerably less predictable in their processes.²⁶ Like process levels, intelligence levels operate on a spectrum, with intelligence growing as a program moves down the continuum

¹⁴ See *id.* at 18–19; see NICK BOSTROM, SUPERINTELLIGENCE: PATHS, DANGERS, STRATEGIES 16 (2014) (discussing the role of data mining in the global financial market).

¹⁵ *Id.* at 14–16.

¹⁶ *Id.* at 16 (noting “the demarcation between artificial intelligence and software in general is not sharp . . . this brings us back to McCarthy’s dictum that when something works it is no longer called AI”).

¹⁷ *Id.* at 12–14. Deep Blue, a chess-playing AI, made news in 1997 when it beat Garry Kasparov, the world chess champion. Unlike Gary Kasparov, however, Deep Blue could not carry that intelligence to other areas, a clear example of a narrow or specific AI.

¹⁸ SCHARRE, *supra* note 8, at 28.

¹⁹ *Id.* at 29.

²⁰ *Id.*

²¹ *Id.*

²² *Id.* at 30.

²³ See *id.* (describing how a Roomba, for example, might move among different processes during completion of its task).

²⁴ *Id.*

²⁵ *Id.* at 31.

²⁶ *Id.* at 30–31.

from automatic to autonomous.²⁷ Autonomous programs do not “think;” if their processes are opaque it is because there is not a simple connection between input and output as there is in an automatic program; rather, autonomous, “goal-oriented” systems assimilate a wide variety of input and produce an output through a process that may be unintelligible to human observers.²⁸

B. AI Interactions with the Modern World: Influence, Benefits, and Dangers

Understanding AI as more than the humanoid robot or omnipotent mastermind enables a deeper understanding of the ways in which AI technology already interacts with and influences global society, as well as of the reasons why greater regulation and oversight is beneficial. AI operates across a multitude of sectors, influencing fields from social media to the global economy and everything in between. The following examples highlight the benefits and the dangers of continually advancing, and often under- or unregulated artificial intelligence technologies.

The first example occurred on May 6, 2010, when the Dow Jones industrial average careened wildly, losing nearly ten percent of its value in just under fifteen minutes and then, within a half hour, rebounded to its prior level.²⁹ Following investigations into what became known as the “Flash Crash,” it became clear that the crisis, which was described by traders as “horrifying,”³⁰ had been set off by a single trading algorithm programmed to sell off a specific type of contract.³¹ These contracts were in turn purchased by specifically programmed purchasing algorithms; the competing algorithms entered into a fast-paced trading race, in which the pace of trade triggered other algorithms to offload their contracts as well, interpreting the fast pace of trading as high liquidity.³² Although stability was soon restored, at the peak of the crisis, “a trillion dollars had been wiped off the market” and investors around the world were shaken.³³

The Flash Crash was not the result of a rogue algorithm or of a weak AI breaking away from programming. It was an example of a weak AI following its programming to the letter in spite of the catastrophic effects of doing so. The Flash Crash was caused by human programmers’ failure to understand the effects of their algorithm

²⁷ *Id.* at 30.

²⁸ *See id.* at 32; JAMES BARRAT, *OUR FINAL INVENTION: ARTIFICIAL INTELLIGENCE AND THE END OF THE HUMAN ERA* 113–14 (2013). The varying levels of complexity involved in autonomous systems has advanced in recent decades, with growing bodies of research on artificial neural networks (ANNs) and genetic algorithms, among others. While this Note does not go into depth on any of these processes, a deeper understanding of the technical aspects of AI research is helpful to those interested in more fully understanding the complexities of regulation and oversight. *See also* BARRAT at 74–75; BOSTROM, *supra* note 14, at 10–11; RAY KURZWEIL, *THE AGE OF SPIRITUAL MACHINES: WHEN COMPUTERS EXCEED HUMAN INTELLIGENCE* 81 (2000).

²⁹ Graham Bowley, *Lone \$4.1 Billion Sale Led to ‘Flash Crash’ in May*, N.Y. TIMES (Oct. 1, 2010), <https://www.nytimes.com/2010/10/02/business/02flash.html> [<https://perma.cc/2MJN-22YD>].

³⁰ SCHARRE, *supra* note 8, at 199.

³¹ *Id.* at 203.

³² BOSTROM, *supra* note 14, at 17.

³³ *Id.*

following its directive to its logical conclusion.³⁴ While the use of algorithmic programming granted benefits in the form of higher trade volume, the potential danger of unforeseen programming consequences clearly played out.

Another illustrative example is the infamous case of Stuxnet malware, which was created to compromise the Siemens machines controlling centrifuges in Iran’s Natanz nuclear facility.³⁵ Though the facility did not suffer catastrophic damage from Stuxnet, the attacks did reduce the lifetime of the centrifuges, as well as undermining confidence in the security of the Iranian facility.³⁶ In addition to these long-term deleterious effects of the program, Stuxnet also ushered in a new era of cyber warfare, hailed as the world’s “first military grade cyber weapon.”³⁷

Stuxnet heralded a sea change in malware; the virus was not contained in the Natanz facility but spread globally, likely transferred by laptops or USB drives infected with the virus.³⁸ Though Stuxnet was designed to attack a specific make of Siemens controller, its presence on the internet affords hackers and programmers with access to the virus’ blueprints an opportunity to dismantle, alter, and learn from the way Stuxnet operates.³⁹ Concern over such cyberattacks has only increased since this first major international incident occurred.⁴⁰

These cases point to a sector of technology and innovation that is advancing—or, perhaps, has advanced—past the point of legal and regulatory control.⁴¹ This note offers these instances as examples of just a few of the diverse situations in which advancing AI and automation technology would benefit from a system of oversight and regulation.

II. INTERNATIONAL LAW: PROMISES, FAILINGS, AND POTENTIAL

A. *Why International Law?*

The rise and increased visibility of the modern international legal system developed in the post-World War II and Cold War eras.⁴² International law is “the legal order ... meant to structure the interaction between entities participating in and shaping

³⁴ *Id.* at 21 (“Smart professionals might give an instruction to a program based on a sensible-seeming and normally sound assumption... this can produce catastrophic results when the program continues to act on the instruction... even in the unanticipated situation where the assumption turns out to be invalid.”).

³⁵ FRED KAPLAN, *DARK TERRITORY: THE SECRET HISTORY OF CYBER WAR 203–11* (2016).

³⁶ Ralph Langner, *Stuxnet’s Secret Twin*, Foreign Policy (Nov. 19, 2013, 5:26 PM), <https://foreignpolicy.com/2013/11/19/stuxnets-secret-twin/> [<https://perma.cc/B5Y4-KREF>].

³⁷ BARRAT, *supra* note 28, at 256.

³⁸ Langner, *supra* note 36.

³⁹ *Id.*

⁴⁰ See, e.g., Natasha Turk, *The next 9/11 will be a cyberattack, security expert warns*, CNBC June 1, 2018, 7:55 AM), <https://www.cnbc.com/2018/06/01/the-next-911-will-be-a-cyberattack-security-expert-warns.html> [<https://perma.cc/7C37-WGSK>].

⁴¹ See Ian Kerr and Katie Szilagy, *Asleep at the switch? How killer robots become a force multiplier of military necessity*, in *ROBOT LAW*, 354 (Ryan Calo, A. Froomkin, and Ian Kerr, eds., 2016) (arguing that, by failing to properly regulate, oversee, and guide the advancement of AI tech, in this case autonomous weapons, society essentially allows new technology to “determine its own use.”).

⁴² Oscar Schachter, *The UN Legal Order: An Overview*, *THE UNITED NATIONS AND INT’L. L.* 3 (Christopher Joyner, ed., 1997) available at <https://www.jstor.org/stable/2204020>.

international relations.”⁴³ Some scholars have argued that international law is not “law,” per se⁴⁴ given its lack of authority and enforcement structures,⁴⁵ but others have noted that “almost all nations observe almost all principles of international law ... almost all of the time.”⁴⁶ International law can help preserve peace and security, manage interstate social and economic disputes, and protect the interests of the international community as a whole.⁴⁷

AI is a problem—like the global arms race or climate change—that implicates all of global society.⁴⁸ Whether operating in financial markets, conflict situations, or social media and data-gathering, advancing AI crosses and will continue to cross national boundaries; as Erdélyi and Goldsmith suggest, purely national responses to this rising challenge may conflict and create more problems than they solve.⁴⁹ Furthermore, isolated national or corporate attempts to solve the emerging research and regulatory problems created by AI may be hasty, ignoring investments in safety to be first to reach a benchmark in machine intelligence.⁵⁰ Advancing AI creates an opportunity for international law to step into a gap that national law is not sufficient to fill.

The problem arises from the fact that international law is not law in the traditional sense of national law in which the sovereign creates the system of laws by which its citizens abide.⁵¹ In the international legal systems, the states engaging in the system are themselves sovereign.⁵² The pertinent questions then become: In what circumstances do states comply with international law and international

⁴³ Wolfrum, *supra* note 9.

⁴⁴ Bolton, *supra* note 10, at 48 (“International law is not law; it is a series of political and moral arrangements that stand or fall on their own merits, and anything else is simply theology and superstition masquerading as law.”).

⁴⁵ See Jana von Stein, *Compliance with International Law*, INT’L STUDIES ASS’N AND OXFORD U. PRESS (last updated Nov. 2017) <http://www.janavonstein.org/uploads/4/6/1/9/46194525/oxford-encyclopedia.pdf> [<https://perma.cc/L9A5-SLD5>].

⁴⁶ LOUIS HENKIN, *HOW NATIONS BEHAVE* 47 (2d. ed. 1979). For the discussion herein, see *supra* Part II Section B, at 6–8.

⁴⁷ *Id.*

⁴⁸ See, e.g., Olivia Erdélyi and Judy Goldsmith, *Regulating Artificial Intelligence: Proposal for a Global Solution*, ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE, 1, 2, 9 (2018), https://www.aies-conference.com/2018/contents/papers/main/AIES_2018_paper_13.pdf [<https://perma.cc/N6SG-GWA8>].

⁴⁹ *Id.* at 1-2.

⁵⁰ BOSTROM, *supra* note 14, at 249. This possibility is particularly concerning in two instances: first, in the case of lethal autonomous weapons; and second, in the case of AGI. Bostrom writes:

Consider a hypothetical AI arms race in which several teams compete to develop superintelligence. Each team decides how much to invest in safety—knowing that resources spent on developing safety precautions are resources not spent on developing the AI... there might be a risk-race to the bottom, driving each team to take only a minimum of precautions. *Id.* at 247.

⁵¹ Samantha Besson, *Sovereignty*, MAX PLANCK ENCYCLOPEDIA OF PUBLIC INTERNATIONAL LAW (Last updated April 2011), <http://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1472> [<https://perma.cc/Z3VW-4NQH>].

⁵² See, e.g., Shen, *infra* note 54 (discussing that states often follow international law, although they are not forced to do so, and can regularly choose not to do so).

obligations, and how can this general compliance be used to create an international structure of governance and oversight for advancing AI technology

B. Under What Circumstance Do States Adhere to International Law?

Academics and international law practitioners have long questioned why states seem to mostly follow international law. This law, which is composed not only of the formal treaties between states but also of more the general principles of customary international law,⁵³ has been able to function more or less effectively for centuries, despite its lack of total enforcement power over sovereign states. Though the debates behind why states recognize international law are intense and ongoing,⁵⁴ of more particular interest to this examination are the following questions: Under what circumstances do states follow international law, and how can this knowledge be applied to the creation of an international governance structure for advancing AI? Despite arguments that states are not obligated to follow international law,⁵⁵ there are more instances of states complying with international laws than not.⁵⁶ The commonality in many of these instances may be, rather than some sense of morality or complicated philosophical principle, the less benign and more realist idea of state self-interest.⁵⁷ States, though they regularly come together to work toward some common purpose, are individual actors that must shape their own policy considerations towards international issues.⁵⁸ The goal for international law is to provide regimes that states can follow that achieve international legal goals while also providing an appealing choice to state self-interest. Such a choice to follow international law can be seen in the disarmament treaties and in international cooperation on nuclear technology that began in the 1950s, and many of these arrangements continue to have a high level of state adherence today.⁵⁹

As previously noted, international law is most useful in circumstances in which one state alone is not capable of managing a problem, or when the interests of the

⁵³ Wolfrum, *supra* note 9.

⁵⁴ See, e.g., Harld Hongju Koh, *Why Do Nations Obey International Laws?*, 106 YALE L.J. 2599, 2602–03 (1997) (arguing that international laws create normative structures that are internalized and reproduced in domestic law, which helps create national understanding of sovereignty and a nation’s place in global society, forming a type of virtuous feedback cycle); see generally Jianming Shen, *The Basis of International Law: Why Nations Observe*, 17 Dickson Int’l L. 287 (1999) (discussing a wide range of theories of observance of international law, including naturalist theories, positivist theories, and other more modern political science theories such as power politics and peaceful coexistence.).

⁵⁵ Eric Posner, *Do States Have a Moral Obligation To Obey International Law?*, 55 STAN. L. REV. 1901, 1902, 1919 (2003) (arguing that states do not, in fact, have a moral obligation to follow international law, but may have prudential reasons for doing so).

⁵⁶ von Stein, *supra* note 45, at 20 (noting that “various mechanisms... can help to ensure that states keep their international promises much of the time”).

⁵⁷ *Id.* at 1918.

⁵⁸ Juliet Kaarbo, Jeffrey S. Lantis, Ryan K. Beasley, and Michael T. Snarr, *The Analysis of Foreign Policy in Comparative Perspective*, FOREIGN POLICY IN COMPARATIVE PERSPECTIVE: DOMESTIC AND INTERNATIONAL INFLUENCES ON STATE BEHAVIOR, 4 (2nd. ed., Ryan K. Beasley, Juliet Kaarbo, Jeffrey S. Lantis, and Michael T. Snarr, eds. 2013).

⁵⁹ John Murphy, *Force and Arms*, THE UNITED NATIONS AND INTERNATIONAL LAW 122–29 (Christopher Joyner, ed., 1997).

international community as a whole are implicated.⁶⁰ Although adherence is not perfect, existential world crises have seen a majority of involved states come to the table and negotiate an international solution through the auspices of international law.⁶¹ Where advancing AI technology does not neatly align with existing international norms, it is necessary to create new structures of governance;⁶² which, like those governing arms proliferation, nuclear weapons, and climate change, advance an international policy goal and offer states benefits they would not be able to gain on their own. Jana von Stein notes that this type of mechanism, combining “the proverbial carrots and sticks; technical and financial assistance; [and] tying good behavior to a particular identity” can be quite effective in holding states to compliance with international law and institutions.⁶³ This research encourages a self-interested view of state compliance, one in which any new international law regime will need to offer states an incentive to comply with its norms.⁶⁴

Examples from other crises clearly show that the mere existence of an existential threat to international society is not necessarily enough to compel full compliance with international law and norms.⁶⁵ It is therefore necessary to make compliance with any international regulatory scheme more attractive to states. Increasing the benefits of compliance can be done in two parallel ways. First, create a system where compliance itself is valuable for states’ reputations.⁶⁶ Where states are seen as upholding their international obligations and complying with international laws and norms, other states may be more willing to enter into future agreements, grant more generous concessions in future negotiations, or cooperate on economic and regulatory projects.⁶⁷

Second, benefits to a state’s self-interest arise when compliance with international law grants the state some type of tangible gain. For example, although membership in the World Trade Organization (WTO) requires some concessions,

⁶⁰ See discussion *supra* II.A, at 6.

⁶¹ One particularly salient example in this case—though comparisons can be overdone—is the creation of the International Atomic Energy Agency in the wake of the Second World War, the bombing of Nagasaki and Hiroshima, and the understanding of what nuclear technology could do, both in terms of societal benefits and potential threats. The IAEA, founded in 1957, had 171 member states as of February 5, 2019. See INTERNATIONAL ATOMIC ENERGY AGENCY (IAEA), <https://www.iaea.org/about/governance/list-of-member-states> [<https://perma.cc/PM7U-5HWN>] (last visited Oct. 2, 2019); CERN and the Human Genome project also present good examples of international scientific collaboration, though without the immediacy that nuclear technology and now, arguably, AI technology present. See BOSTROM, *supra* note 14, at 253.

⁶² Grant Wilson, *Minimizing Global Catastrophic and Existential Risks from Emerging Technologies Through International Law*, 31 VA. ENVTL. L.J. 307, 349–350 (2013).

⁶³ See von Stein, *supra* note 45 (including an in-depth discussion of the elements of international normative structures that encourage compliance with international law).

⁶⁴ *Id.*

⁶⁵ See, e.g., Michael D. Shear, *Trump Will Withdraw U.S. From Paris Climate Agreement*, N.Y. Times (June 1, 2017), <https://www.nytimes.com/2017/06/01/climate/trump-paris-climate-agreement.html> [<https://perma.cc/5G67-4M95>].

⁶⁶ Andrew Guzman, *A Compliance-Based Theory of International Law*, 90 CALIF. L. REV. 1823, 1880 (2002).

⁶⁷ *Id.* at 1886–87.

member states are also able to access preferential trading with partners.⁶⁸ Creating a regulatory system that engenders state adherence around AI policy must keep three goals in mind: (1) provide a solution to a problem implicating the interests of the international community; (2) create a structure in which adherence to international norms creates a virtuous compliance cycle; and (3) incentivize states to comply with the governance structure's policies through tangible gains given to member states.

C. Regulatory Efforts, Recommendations, and Their Messages

Despite the lack of an overarching regulatory structure, there have been various state, multi-national, and non-governmental attempts to introduce coherence and regulatory oversight to AI research and use. NGOs, expert agencies, and even the United Nations have urged greater oversight of AI advances, while several states have also released plans for the advancement of AI technology. Multiple NGOs and other non-state actors have spoken out in favor of increasing regulation and oversight of AI research and use. Some include calls for increased regulation, while others offer paths forward or designs to emulate.⁶⁹ These are not merely specialist organizations, but rather some of the most well-known and integrated NGOs: in 2015, the United Nations Interregional Crime and Justice Research Institute (UNICRI) launched its Centre for Artificial Intelligence and Robotics, which was created to “educate and inform stakeholders . . . [and] progress discussion on robotics and artificial intelligence governance.”⁷⁰

Law and AI experts have called for international regulation and oversight of AI technology and use. One highly relevant proposal is creation of a new international organization to encourage policy discussion and eventual regulation of AI-related matters, which, though beginning as a voluntary advising body, could gain enforcement and oversight powers.⁷¹ One such body, the Center for the Governance of AI, is active at the international level, speaking to non-governmental research groups as well as national governments about the possible dangers and benefits of AI, as well as of policy paths forward to minimize the risks and establish a structure of development and governance for AI technology.⁷²

Much has also been proposed regarding the regulation of autonomous weapons; sensible, given their immense potential harm the increasing use of semi-

⁶⁸ WORLD TRADE ORGANIZATION, *Principles of the trading system*, https://www.wto.org/english/thewto_e/whatis_e/tif_e/fact2_e.htm [<https://perma.cc/T2FP-WDJ7>] (last visited Jan. 14, 2019).

⁶⁹ See, e.g., BODEN, *supra* note 6, at 147–49 (discussing NGO and expert calls for increased oversight).

⁷⁰ UNITED NATIONS INTERREGIONAL CRIME AND JUSTICE RESEARCH INSTITUTE, UNICRI Centre for Artificial Intelligence and Robots, http://www.unicri.it/in_focus/on/UNICRI_Centre_Artificial_Robotics [<https://perma.cc/6JHC-EGM3>] (last visited Jan. 13, 2019).

⁷¹ Erdélyi & Goldsmith, *supra* note 48, at 3.

⁷² UNIVERSITY OF OXFORD FUTURE OF HUMANITY INSTITUTE, Centre for the Governance of AI, <https://www.fhi.ox.ac.uk/GovAI/> [<https://perma.cc/MM8G-3V4Y>] (last visited Jan. 13, 2019). Although outside the scope of this paper, FHI and the Centre for the Governance of AI have a wealth of research on desired policy outcomes of governance structures, as well as more technical information such as forecasts on future AI capabilities, malicious use, and machine learning advances, which can be accessed at <https://www.fhi.ox.ac.uk/publications/> [<https://perma.cc/9MEU-SCUB>] (last visited Jan. 13, 2019).

autonomous⁷³ and autonomous weapons⁷⁴ in the field. In 2012, the Campaign to Stop Killer Robots was founded.⁷⁵ This campaign, organized to stop the use of fully autonomous lethal weapons and maintain human control over the use of force, is supported by nearly 120 national, regional, and international NGOs.⁷⁶ Others have called for a more “vibrant, measured, and mature discussion of the relevant legal issues,”⁷⁷ arguing that the law of armed conflict will be shaped by the use of such autonomous weapons⁷⁸ and that a ban of such systems would ignore the military practicalities and political complexities that are already tied into states developing autonomous weapons systems.⁷⁹

Several national and supranational actors have also made steps forward in AI regulation and oversight. In June 2018, the European Union (“EU”) named 52 experts to its High Level Expert Group on Artificial Intelligence, which aims to produce policy recommendations on social, political, economic, and ethical issues related to AI, as well as balance economic competitiveness concerns tied to transparency, data-protection, and fairness.⁸⁰ In December 2018, the group published its draft AI Ethics Guidelines, which aims to “maximize the benefits of AI while minimising its risks” by “ensuring an ‘ethical purpose’ . . . and [being] technically robust.”⁸¹

Over twenty-five states have announced their AI strategies or have published plans for future strategies, including the US, Russia, China, and India.⁸² Many plans focus on maintaining a competitive edge in the emerging AI market, although several also consider the ethical and safety elements of advancing AI.⁸³ One strategy notable in its attention to safe progress is the US Department of Defense’s (DoD) attitude towards the development of autonomous weapons systems, which might be extrapolated to encompass advanced AI research.⁸⁴

⁷³ See SCHARRE, *supra* note 8, at 103 (“As of June 2017, sixteen countries possessed armed drones. . .”).

⁷⁴ See *id.* at 47–48 (discussing the Israeli Harpy drone, which is fully autonomous, requiring no human approval of its targets. It has been sold to China, India, and Turkey, among others).

⁷⁵ CAMPAIGN TO STOP KILLER ROBOTS, *About Us*, <https://www.stopkillerrobots.org/about/> [<https://perma.cc/VUD2-YXCM>] (last visited Oct. 6, 2019).

⁷⁶ *Id.*

⁷⁷ Michael N. Schmitt and Jeffrey S. Thurnher, “*Out of the Loop*”: *Autonomous Weapon Systems and the Law of Armed Conflict*, 4 HARV. NAT’L SEC. J. 231, 233 (2013);

⁷⁸ *Id.* at 233–34.

⁷⁹ *Id.* at 280–81.

⁸⁰ EUROPEAN COMMISSION, *High-Level Expert Group on Artificial Intelligence*, <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence> [<https://perma.cc/EF5J-6ZMY>] (last visited Jan. 13, 2019).

⁸¹ EU HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *Draft Ethics Guidelines for Trustworthy AI* (Dec. 18, 2018), <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai> [<https://perma.cc/822H-LR7V>].

⁸² Tim Dutton, *An Overview of National AI Strategies* (June 28, 2018), <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd> [<https://perma.cc/YN22-GGSE>].

⁸³ *Id.*

⁸⁴ SCHARRE, *supra* note 8, at 89.

The DoD’s Directive 3000.09 (“Autonomy in Weapons Systems”) creates three classes of weapons systems that are given a “green light” for development and use.⁸⁵ For proposed systems that would use autonomy or intelligence outside of these categories, the system issues a “yellow light,” requiring review before any further development of the technology, and then a second review before field use of the system.⁸⁶ Although this policy is created specifically for autonomous and intelligent weapons systems, its stated goal of “[minimizing] the probability and consequences of failures in autonomous [systems]”⁸⁷ is one that can easily be transferred to AI, creating a system of checks and review that would allow greater investment in safety in and control over the advancement of AI.⁸⁸

Although none of these proposals are truly international in scope, many of them offer strong elements that could be incorporated into an international regulatory regime, including the American policy discussed above,⁸⁹ or the Centre for the Governance of AI’s proposed research and development guidelines.⁹⁰ Drawing from best practices of states and NGOs would be beneficial to the proposed regulatory body, and would allow it to begin with a strong foundation.

III. GLOBAL GOVERNANCE OF AI: OVERSIGHT, REGULATION, AND PROMOTION

A. *The Regulatory Promise and Potential of International Law*

International law has most relevance where national law is not sufficient to protect the interests of global society; the advancement of AI presents an opportunity for greater robustness of international regulatory structures.⁹¹ The growth of AI technology calls for a response from international society. International law and institutions, calling upon both states’ tendencies to comply with international law when doing so is seen as virtuous⁹² and upon states’ individual self-interest, may be able to create a regulatory regime that is attractive enough to compel adherence from a majority of state players.⁹³ Such a regime would not seek to halt research on and development of AI, but to pursue such research and development safely and intentionally.⁹⁴

The most traditional method of international cooperation is, as recommended in multiple other publications,⁹⁵ the creation of an international treaty. This proposal

⁸⁵ *Id.* These three classes are “semiautonomous weapons, such as homing munitions...defensive supervised autonomous weapons, such as the ship-based Aegis weapon system...and non-lethal, non-kinetic autonomous weapons, such as electronic warfare.”

⁸⁶ *Id.*

⁸⁷ *Id.* at 90.

⁸⁸ See BOSTROM, *supra* note 14, at 206.

⁸⁹ SCHARRE, *supra* note 8, at 89.

⁹⁰ See Centre for the Governance of AI, *supra* note 72.

⁹¹ See, e.g., Wolfram, *supra* note 9 (considering areas falling under the governance of international law such as the high seas, climate issues, and international economic issues).

⁹² See von Stein, *supra* note 45.

⁹³ See, e.g., IAEA, *supra* note 61, and the 170 member states of IAEA.

⁹⁴ See BOSTROM, *supra* note 14, at 206.

⁹⁵ See, e.g., Erdélyi and Goldsmith, *supra* note 48; Wilson, *supra* note 62, at 349–50.

offers a more incentivized approach: the creation of an international body of collaborating scientists, researchers, and experts in the field— both civilian and governmental— whose research and collaborative efforts are available only to parties to the treaty. A similar body has been used in response to a broad range of “global catastrophic risks” or “existential risks,”⁹⁶ proposing a regulatory body controlled by a group of experts to govern member states of previously created treaties.⁹⁷ This body of experts should include experts from civil society as well as government representatives, to promote transparency in regulation and oversight.⁹⁸ Leveraging the potential international pitfalls of unregulated AI, the treaty body could create a regime in which a state’s refusal to sign and ratify the new AI treaty and become part of the regulatory institution is seen as damaging to its reputation.⁹⁹ Encouraging consideration of the “global catastrophic risks” that might occur should state refuse to comply could also be a motivator.

The combination of these recommendations is the creation of a new international treaty body, overseen by the United Nations (possibly drawing from UNICRI, which already has subject-matter expertise on AI)¹⁰⁰ paired with an expert body or advisory panel serving the members of the new treaty. While the proposed treaty would provide general guidelines for member states on research and development of advancing AI, the expert body could provide case-by-case recommendations on new research and controversial development proposals. The body could also develop best practices and contribute to important advances through collaborative research.¹⁰¹ Ideally, the panel would also represent the cutting edge of AI research and development, with ideas shared freely among the body’s members.

While states may be less immediately open to joining, many AI experts have already expressed concern about the direction and speed of research, calling for guidance and even delay of certain strands of AI research as well as for more focus on developing AI safely and ethically¹⁰² and would likely be open to joining a body of this sort. One way to make this body more attractive is to encourage the membership and active participation of expert groups such as the American Association for AI and the Machine Intelligence Research Institute (“MIRI”)¹⁰³ and individual experts such as Max Tegmark¹⁰⁴ and Nick Bostrom,¹⁰⁵ all of whom have

⁹⁶ Wilson, *supra* note 62, at 308–11 (discussing the risks created by nanotechnology, AI, bioengineering, and the Large Hadron Collider).

⁹⁷ *Id.* at 355–56.

⁹⁸ *Id.* at 356–57.

⁹⁹ See von Stein, *supra* note 45, at 7–9 (discussing the role of reputation in creating state compliance).

¹⁰⁰ See UNICRI, *supra* note 70.

¹⁰¹ See BOSTROM, *supra* note 14, at 249–50 (discussing the benefits of collaboration, including “the sharing of ideas.”).

¹⁰² BODEN, *supra* note 6, at 147; Ian Semple, *Thousands of Leading AI Researchers Sign Pledge Against Killer Robots*, THE GUARDIAN (July 18, 2018), <https://www.theguardian.com/science/2018/jul/18/thousands-of-scientists-pledge-not-to-help-build-killer-ai-robots> [<https://perma.cc/4PPL-FKFB>].

¹⁰³ See BODEN, *supra* note 6, at 148–49.

¹⁰⁴ Max Tegmark, FUTURE OF LIFE INSTITUTE, <https://futureoflife.org/author/max/> [<https://perma.cc/3BGT-DJM4>] (last visited Mar. 15, 2019).

¹⁰⁵ Nick Bostrom, NICK BOSTROM, <https://nickbostrom.com/> [<https://perma.cc/Y248-ZG5F>] (last visited Mar. 16, 2019).

expressed concerns.¹⁰⁶ Participation by these experts in the proposed panel could further incentivize states to join, in order to gain access to their research and collaborative technological development.

States are more likely to adhere to international law when doing so promotes some international interest and offers incentives to states' self-interest. By offering an answer to the international challenges posed by the expansion of AI across all sectors, including financial, social, and military, the proposed treaty and body of experts would protect the interests of the international community. Further, by providing access to an international, collaborative body of experts that not only provides best practices recommendations and oversights but also to shared information, pooled resources, and joint research, the recommended treaty would offer states and other organizations tangible incentives to both join and adhere to the proposed convention.

A treaty and expert regulatory body could also help control AI advances in the future. While this discussion has focused mainly on weak AI, autonomy, and the possibility of creating AGI in the near to middle-term future, many experts are more concerned about the advances that might follow; namely, superintelligence,¹⁰⁷ which is "any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest."¹⁰⁸ Creating a regulatory body in the present will ensure there are safeguards in place in the event that AI technology reaches such heights, possibly preventing the disastrous consequences that might result.¹⁰⁹ These technological advances have not yet arrived, but they are on the horizon,¹¹⁰ and establishing an international oversight body early on could prevent more wicked challenges down the road.

B. If International Law Is the Answer, Why Is It Not Currently in Use?

If the potential gains from the international regulation and cooperation on advancing AI are so immense, why hasn't an international solution yet been accepted? There are two arguments, the first of which is principled and second of which is more pragmatic. First, international law lacks the capacity to properly regulate and oversee a field as rapidly advancing as AI.¹¹¹ Second, many powerful

¹⁰⁶ BODEN, *supra* note 6, at 147–48 (noting that there have been multiple expert conferences discussing AI safety as well as a number of open letters condemning use of, for example, autonomous weapons in war).

¹⁰⁷ See, e.g., BOSTROM, *supra* note 14, at 259–60; BARRAT, *supra* note 28, at 152–53; BODEN, *supra* note 6, at 131.

¹⁰⁸ BOSTROM, *supra* note 14, at 22 (internal footnote omitted).

¹⁰⁹ See *id.*, at 95–99.

¹¹⁰ See *id.*, *supra* note 14, at 22–29 for an in-depth discussion of expert opinions on when human-intelligence level AI will be achieved.

¹¹¹ See, e.g., von Stein, *supra* note 45, at 21 (noting that not all states follow almost all of their agreements almost all of the time, and offering a discussion of the complex nature of state compliance); see also Wolfram, *supra* note 9, at 5, 14 (noting there is no enforcement mechanism in international law). Although these authors do not support these arguments, they do make note of them as critiques raised against international law.

states are simply disinterested in international regulation and oversight of advancing AI technology.¹¹²

First, some scholars argue that international law cannot create binding legal requirements.¹¹³ Without an overarching authority or enforcement mechanism, international law would lack the ability to enforce any new AI regime it attempted to impose, and thus would not be the preferred method of regulation. This argument can be answered by considering that international law, though lacking traditional enforcement power, does have other means, such as international interest, incentives to states, and reputational value to encourage compliance.¹¹⁴

Second, international regulation might not be in the best interest of all states. AI is a “dual-use” technology,¹¹⁵ and though few are opposed to the advancement of peaceful uses of AI,¹¹⁶ there has been opposition to advancing military uses.¹¹⁷ Many states, however, have already invested heavily in AI’s military potential¹¹⁸ and prefer a regime governed by national regulation. This “race dynamic,” where actors refuse to cooperate out of fear that they will not achieve a new technology first,¹¹⁹ emerges in recent research on public feelings about AI: in a January 2019 poll, more American respondents answered that they believed advancing AI could do more harm than good, yet there was uncertainty as to who, if anyone, should control that advancement.¹²⁰ There is a related concern that if other states are developing unsavory advances for AI, your state should as well,¹²¹ regardless of any regulatory structure.¹²²

This second set of arguments, however, merely repeats several underlying reasons for regulating AI in the first place, and can be answered by the promise inherent in an international regulatory structure that, through a series of incentives and reputational elements,¹²³ can gain a level of adherence high enough to undermine

¹¹² George Lucas, Jr., *Legal and Ethical Precepts Governing Emerging Military Technologies: Research and Use*, 2013 UTAH L. REV. 1271, 1275 (2013) (noting that international “regulatory statutes would prove unacceptable to, and unenforceable against, many of the relevant parties”).

¹¹³ Posner, *supra* note 55, at 1905.

¹¹⁴ See Part II Section B, *supra* page 10–13.

¹¹⁵ BARRAT, *supra* note 28, at 155.

¹¹⁶ See, e.g., BOSTROM, *supra* note 14, at 15–16 (discussing several current peaceful uses of AI, including increasing the speed and capacity of internet searches and voice and facial recognition).

¹¹⁷ See, e.g., HUMAN RIGHTS WATCH, *supra* note 81, at 12 (calling for an end to the use and prevention of future development and use of increasingly automated and autonomous drones in warfare).

¹¹⁸ See SCHARRE, *supra* note 8, at 102–03 (noting the number of states possessing and using armed drones). Consider also the case of Israel, which has developed the fully autonomous Harpy drone and sold this drone to, among others, China, India, and Turkey, creating both a military and financial incentive for Israel to avoid increased regulation of autonomous weapons. See *id.* at 45–48.

¹¹⁹ BOSTROM, *supra* note 14, at 246–49.

¹²⁰ Karen Hao, *Americans want to regulate AI but don’t trust anyone to do it*, MIT TECH. REV. (Jan. 10, 2019), <https://www.technologyreview.com/s/612734/americans-want-to-regulate-ai-but-dont-trust-anyone-to-do-it/> [<https://perma.cc/ZCX4-PWXZ>].

¹²¹ See, e.g., SCHARRE, *supra* note 8, at 117–19 (discussing the beginning of what may become an autonomous arms race).

¹²² *Id.* at 330 (“The main rationale for building fully autonomous weapons seems to be the assumption that others might do so”).

¹²³ See Part III Section A, *supra* page 11–14.

bad actors. Despite its flaws, international law still offers the best opportunity for true oversight and guidance of advancing AI.

IV. CONCLUSION

At the conclusion of his book *Superintelligence*, Nick Bostrom writes:

Before the prospect of an intelligence explosion, we humans are like small children playing with a bomb. Such is the mismatch between the power of our plaything and the immaturity of our conduct... A sensible thing to do would be to put it down gently, quickly back out of the room, and contact the nearest adult. Yet . . . some little idiot is bound to press the ignite button just to see what happens. Nor can we attain safety by running away . . . nor is there a grownup in sight.¹²⁴

Human society has held the nuclear bomb in its hands for well over fifty years,¹²⁵ and this new bomb is no different. Although there are dangers, we are equipped to handle them, provided regulatory oversight is imposed now rather than after the ignite button has been pressed. International governance offers an answer to the looming promises and pitfalls of advancing AI. The proposed regime could provide guidance and safety while also promoting a collaborative spirit that could see AI technology advance slightly more swiftly and much more safely.¹²⁶ An international body focused on safe development and use of AI would promote international welfare, search out solutions that work best, not first,¹²⁷ and ensure that global society benefits from the promise of AI rather than suffers from the dangers.

¹²⁴ BOSTROM, *supra* note 14, at 259.

¹²⁵ See INTERNATIONAL ATOMIC ENERGY AGENCY (IAEA), https://www.iaea.org/sites/default/files/16/08/iaea_safeguards_introduutory_leaflet.pdf [<https://perma.cc/PLH8-NDBA>] (last visited Jan. 13, 2019).

¹²⁶ See BOSTROM, *supra* note 14, at 306–07.

¹²⁷ BARRAT, *supra* note 28, at 266 (“Like natural selection, we choose solutions that work first, not best.”).