De Lege Ferenda (2021) Vol IV, Issue ii, 1–20

Artificial Intelligence and International Law: Towards a New Accountability Framework

Nirmalya Chaudhuri*

Abstract

Artificial intelligence, despite its revolutionary potential, brings forth pressing questions of accountability when its actions result in transnational consequences. Due to the difficulty in ascertaining AI's exact functioning, coupled with the astronomical pace of technological development, it is imperative to determine whether the existing international legal regime is suitable for tackling the problem of accountability. By discussing various modes of accountability under international law, it is argued in this article that none of the legally established mechanisms can satisfactorily ensure accountability for actions of AI entities. For instance, both doctrinal and pragmatic concerns preclude the fixing of accountability on the AI entity, or holding the manufacturer of the entity accountable under international law. Similarly, due to the (often) unpredictable nature of functioning of AI, individual criminal responsibility will not be sufficient to cover all kinds of cases. While State responsibility may sound attractive, rebutting the defence of *force majeure* will often prove to be insurmountable. In this article, it is argued that absolute State liability can be a possible solution, which will entail holding the State accountable for transnational consequences caused by the actions of AI entities used by the State, its citizens, and corporate nationals. The defence of *force majeure* would also be precluded under the proposed accountability regime. In order to disincentivise hacking of AI software by foreign States or those acting under their control, accountability will be shifted onto the shoulders of a third State when it is shown that the latter was responsible for the consequences arising out of the

* Law student at the West Bengal National University of Juridical Sciences, India. The author is grateful to the anonymous reviewers for their useful suggestions and comments. All errors remain the author's own. The author can be reached at nirmalyac08@gmail.com.

actions of the AI entity. The proposed accountability regime, which closely mirrors that governing outer space activities, can go a long way in international regulation of AI without hindering technological progress.

Keywords: artificial intelligence, international law, accountability, State responsibility, absolute liability

I. Introduction

Once a topic confined to the pages of science fiction novels, artificial intelligence (AI) today plays a significant role in our daily lives. However, its transnational impact, though considerable, seems to be a neglected topic as far as the international legal order is concerned. In fact, it would not be an exaggeration to claim that international law may require significant changes in order to keep pace with the astronomical speed at which AI technology is evolving around the world.[1]

Defining the ambit of AI poses serious challenges.[2] However, it can generally be stated that the spectrum of AI broadly comprises 'weak' or 'narrow' AI with simpler algorithms and lesser computational ability, and 'strong' AI with enhanced computing and autonomy.[3] At the far end of the spectrum, lies artificial general intelligence (AGI), which has not been achieved at this stage, and which would include systems that are able to self-evolve and possess more abilities than it enjoyed at the time when it was programmed.[4] The International Committee of the Red Cross (ICRC) has defined autonomous systems as those that are capable of receiving information from the environment, processing it, and taking appropriate action without human aid or intervention.[5]

Despite its revolutionary potential, AI can be subject to inadvertent failure or deliberate misuse, with its effects reaching far beyond national borders. Scholars have pointed out how AI can be used for illegal surveillance through facial recognition,

1 See Matthijs M Maas, 'International Law Does Not Compute: Artificial Intelligence and the Development, Displacement or Destruction of the Global Legal Order' (2019) 20(1) Melbourne Journal of International Law 29.

2 See Rex Martinez, 'Artificial Intelligence: Distinguishing between Types & Definitions' (2019) 19(3) Nevada Law Journal 1015.

3 Michael Guihot, Anne F Matthew and Nicholas P Suzor, 'Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence' (2017) 20(2) Vanderbilt Journal of Entertainment and Technology 385, 395-396.

4 ibid.

5 International Committee of the Red Cross, 'Autonomy, artificial intelligence and robotics: Technical aspects of human control' (*ICRC*, August 2019), 7 <https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control> accessed 16 May 2021.

or for interference in the electoral process in democracies.[6] Similarly, AI profiling can possibly violate human rights of migrants and asylum seekers, by imputing negative labels to racial or ethnic minorities, through the use of facial recognition technology.[7] Moreover, the decisions made by AI systems often do not perfectly follow a cause-and-effect relationship, leading to unpredictable results.[8] Due to the 'black-box' nature of functioning of AI and the accompanying uncertainty, it has been contended that unregulated AI can potentially lead to serious violations of transnational law.[9] Technologically advanced 'neural networks', that are meant to operate like the human brain, can learn from its external environment in complex ways that cannot be predicted by humans at the time of the initial programming.[10] Such concerns are not unfounded, since the unpredictable consequences flowing from automated decision-making has is already happening.[11] For instance, the use of AI in financial markets has led to sudden flash crashes, owing their origin to uncertain decision-making and interaction with other AI algorithms.[12] In addition to such drawbacks, studies have shown that AI systems are not immune from errors and systemic bias.[13]

In the face of such serious dangers, it is imperative to evolve a legal mechanism by which accountability can be fixed for the actions of AI, especially when such consequences transcend national borders. Accountability, its practical manifestation in the form of granting access to effective remedies, and their enforcement in cases of violation, constitute the cornerstones of international

6 Axel Walz and Kay Firth-Butterfield, 'Implementing Ethics into Artificial Intelligence: A Contribution, from a Legal Perspective, to the Development of an AI Governance Regime' (2019) 18(1) Duke Law and Technology Review 176, 194.

7 Ana Beduschi, 'The Big Data of International Migration: Opportunities and Challenges for States Under International Human Rights Law' (2018) 49(3) Georgetown Journal of International Law 981, 1010-1011.

8 Ryan Abbott and Alex Sarch, 'Punishing Artificial Intelligence: Legal Fiction or Science Fiction' (2019) 53(1) UC Davis Law Review 323, 331.

9 ICRC Report (n 5) 10-11.

10 Ashley Deeks, 'The Judicial Demand for Explainable Artificial Intelligence' (2019) 119(7) Columbia Law Review 1829, 1832-1833.

11 See Matthew O Wagner, 'You Can't Sue a Robot: Are Existing Tort Theories Ready for Artificial Intelligence?' (2018) 1(4) RAIL: The Journal of Robotics, Artificial Intelligence and Law 231, 231.

12 Yavar Bathaee, 'The Artificial Intelligence Black Box and the Failure of Intent and Causation' (2018) 31(2) Harvard Journal of Law and Technology 889, 924.

13 See Sonia K Katyal, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66(1) UCLA Law Review 54.

law.[14] Due to the unpredictability and complexity associated with AI decision-making, it becomes extremely difficult to pinpoint blame on any single actor, leading to further complications at the remedial stage. While this conclusion can be reached for any AI system in general, the argument of lack of accountability has been extensively raised in the case of autonomous weapon systems (AWS),[15] resulting in a clarion call to ban such armaments.[16] Christof Heyns argues that the absence of a mechanism to ensure accountability in matters of life and death is itself a violation of the right to life and human dignity, and this accountability vacuum created by AWS can be a legitimate ground for banning such weapons.[17]

Keeping similar concerns in mind, the European Union decided to incorporate certain limited safeguards in the General Data Protection Regulation (GDPR).[18] For example, Article 22(1) provides for the right not to be legally affected by decisions or profiling made solely by an automated system. The provision for compensation in cases where the GDPR is violated can be seen as providing for a remedy, thus satisfying the need for accountability.[19] More importantly, it has been contended that the GDPR grants a right of explanation to determine the process

[14] International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 art 2(3); Basic Principles and Guidelines on the Right to a Remedy and Reparation for Victims of Gross Violations of International Human Rights Law and Serious Violations of International Humanitarian Law, UNGA Res. 60/147 (21 March 2006) UN Doc A/RES/60/147.

[15] Human Rights Watch & Harvard Law School International Human Rights Clinic, 'Mind the Gap: The Lack of Accountability for Killer Robots' (9 April 2015) <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots> accessed 16 May 2021. Note that this report does not discuss State responsibility.

[16] Human Rights Watch & Harvard Law School International Human Rights Clinic, 'Losing Humanity: The Case Against Killer Robots' (19 November 2012) <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots> accessed 16 May 2021.

[17] Christof Heyns, 'Human Rights and the use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement' (2016) 38(2) Human Rights Quarterly 350, 373. However, certain authors have pointed out that potential gaps in fixing accountability cannot be the sole reason to ban AWS. See Charles J Dunlap, Jr, 'Accountability and Autonomous Weapons: Much Ado About Nothing' (2016) 30(1) Temple International and Comparative Law Journal 63, 66.

[18] Council Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.

[19] ibid art 82(1).

adopted by the automated system in arriving at its conclusions, as well as providing safeguards against discrimination.[20]

In the case of AWS, it has been contended that "meaningful human control" should be a prerequisite, so that a human being can take the final call on whether to use force or not, rather than delegate such a function to a machine.[21] As Heyns argues, "meaningful human control"[22] can only be achieved when human beings have the sole ability to take decisions on not only how to use force; but also determine as to when, where, and against whom such force is to be used. Yet, as the ICRC has noted, the problem with this 'human-on-the-loop' approach is that the person charged with taking the final decision may not have full knowledge of the situation and act with automation bias. They may also simply want to shift accountability to the AI system for fear of making an erroneous decision.[23] Automation bias refers to the tendency of humans to act in accordance with machine-generated output, rather than searching for information that could refute the inference arrived at by the machine.[24] Moreover, studies have shown that in the face of sophisticated automation technology, persons charged with monitoring the functioning of the machine exhibit over-reliance on automation, and show reduction in skill levels along with decreased awareness.[25] This nullifies the purpose of the entire exercise. Therefore, such solutions are unlikely to help as far as accountability for the actions of AI are concerned.

In Part II of this article, the various modes of accountability under international law are discussed, in order to show that none of them are suitable for the satisfactory regulation of AI on the international plane. In Part III, a model of absolute State liability is put forth as a possible solution to this legal vacuum, which

[20] Brandon W Jackson, 'Artificial Intelligence and the Fog of Innovation: A Deep-Dive on Governance and the Liability of Autonomous Systems' (2019) 35(4) Santa Clara High Technology Law Journal 35, 44.

[21] Heyns (n 17) 375.

[22] ibid 375-376.

[23] ICRC Report (n 5) 9.

[24] Mary L Cummings, 'Automation and Accountability in Decision Support System Interface Design' (2006) 32(1) The Journal of Technology Studies 23, 25.

[25] ibid 24.

would involve holding the State accountable for transnational consequences of AI used by its own organs, citizens, and corporate nationals. Part IV concludes.

II. Inapplicability of Traditional Modes of Accountability Within the International Legal Framework

A. Responsibility of the AI system

The first possible contender for fixing accountability is the AI system itself. There has been considerable debate on the question of whether AI systems should be granted legal personality,[26] so that they can be held accountable for their actions. Certain scholars point out that throughout history, legal personality under municipal law has been extended to inanimate objects such as ships and idols. It has also been extended to entities which are not natural persons, such as corporations and governmental bodies.[27] On the other hand, it has been contended that the analogy between corporations and AI ignores the undeniable fact that corporations act through human agents; while the AI system, once programmed, does not require humans to act as agents for performing its tasks.[28] While this contention sounds attractive, achieving legal personality under municipal law and under international law are quite distinct, as discussed under Part II.B below using corporations as an example.

Further, it has been argued that since AI systems lack moral agency, it would be difficult to hold them responsible for even grave breaches of international law.[29] On this question however, Hallevy feels that modern AI with significant cognitive ability can be shown to possess both *mens rea* (the mental element) and the *actus reus* (the act or omission) required to commit a crime.[30] Even if that argument is accepted, practical difficulties would not allow for accountability to be fixed. For instance, even if the AI system is convicted, the logical course of punishing the AI would lead to absurd results. It is not too difficult to guess that robots cannot be imprisoned or even fined, as they would generally lack bank accounts, assets,

[26] See Lawrence B Solum, 'Legal Personhood for Artificial Intelligence' (1992) 70(4) North Carolina Law Review 1231.

[27] ibid 1239.

[28] Vikram R Bhargava and Manuel Velasquez, 'Is Corporate Responsibility Relevant to Artificial Intelligence Responsibility?' (2019) 17(Special Issue) Georgetown Journal of Law & Public Policy 829, 841.

[29] UN Human Rights Council, 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns' (9 April 2013) UN Doc A/HRC/23/47, 14, para 76.

[30] Gabriel Hallevy, 'The Criminal Liability of Artificial Intelligence Entities: From Science Fiction to Legal Social Control' (2010) 4(2) Akron Intellectual Property Journal 171, 187-188.

or cash.[31] This shortcoming would even make civil liability for AI systems only a remote theoretical possibility. Yet, Hallevy contends that the common modes of punishment can be applied to AI - deletion of software instead of capital punishment, restricting its freedom of action for a limited period instead of imprisonment, and compulsory use of the AI system for the benefit of the community instead of fines or community service.[32] While such novel propositions seem attractive, assigning responsibility to AI systems is still at the deliberative stage, and cannot be relied upon as a mechanism for fixing international accountability.

B. Corporate responsibility in international law: holding the corporation using AI accountable

One of the most appealing solutions is to hold the corporation using AI accountable, given the fact that private corporations would be a major user of advanced AI technology. However, there is a glaring lack of consensus regarding whether corporations can be considered as subjects of international law, which is still primarily State-centric in nature.[33] Apart from academic materials, the decisions of American courts on the interpretation of the Alien Tort Statute (ATS)[34] provide useful guidance on this point. However, on the question of whether corporations are liable for violation of international law, the decisions have been far from consistent.[35]

Recently, in *Jesner v. Arab Bank*,[36] the United States (US) Supreme Court refused to conclusively answer this question, and observed that it was doubtful whether corporations can indeed be held responsible under international law. On a broader scale, the issue remains unsettled and debatable.[37] As Julian Ku observes, international legal instruments generally desist from imposing direct liability on private actors. Instead, the respective states are given the onus of imposing obligations on private entities or individuals within their respective domestic

31 ibid 199.

32 ibid 196-199.

33 See Emeka Duruigbo, 'Corporate Accountability and Liability for International Human Rights Abuses: Recent Changes and Recurring Challenges' (2008) 6(2) Northwestern Journal of International Human Rights 222.

34 Alien Tort Statute, 28 U.S. Code § 1350.

35 See, *Kiobel v. Royal Dutch Petroleum Co.*, 621 F.3d 111, 149 (2nd Cir. 2010) asserting that the notion of corporate liability for violating customary international law is not universally accepted. But see, *In re South African Apartheid Litigation: Ntsebeza v. Ford Motor Company*, 15 F.Supp.3d 454, 464-465 (S.D.N.Y. 2014) stating that there is no basis to claim that corporate liability is not recognized under customary international law.

36 *Jesner v. Arab Bank Plc*, 138 S. Ct. 1386, 1402 (2018).

37 Julian G Ku, 'The Curious Case of Corporate Liability under the Alien Tort Statute: A Flawed System of Judicial Lawmaking' (2011) 51(2) Virginia Journal of International Law 353, 377.

jurisdictions.[38] This proposition logically follows from the fact that private actors like corporations do not have the authority to enter into binding treaties under the international legal order.[39] Ku further fortifies his argument by stating that after the Second World War, although individual responsibility was imposed on persons involved in the operations of I.G. Farben for crucially assisting the Nazis, the firm itself was never charged under international law.[40] In the present-day context, the Rome Statute expressly limits the jurisdiction of the International Criminal Court (ICC) to natural persons, thus excluding juristic persons like corporations from being held criminally liable under international law.[41]

However, the growing influence of multinational corporations (MNC) has led to scholars calling for corporations to be held liable for wrongful actions under international law, including through corporate criminal responsibility.[42] It has been contended that in order to preclude the situation in which corporate entities enjoy complete immunity, corporations should at least be recognized as 'participants' in the international legal regime for them to be held responsible for their actions.[43] Recognising corporations as 'participants' in international law signifies a realisation of the fact that they play a significant role in the formulation of the rules of international law, especially in certain areas like international investment law.[44] This shift has been made possible by giving corporations the right to participate in various fora, where the rules of international law are formulated and deliberated upon.[45] Simply put, accepting corporations as 'participants' in the international legal order signifies the undeniable reality that international law shapes and is, in turn, shaped by the actions of transnational corporations.

The test laid down in the *Reparation for Injuries* case has been widely accepted as laying down the criteria for determining whether an entity possesses international legal personality.[46] According to this test, an entity must possess rights and duties

[38] ibid 384.

[39] ibid.

[40] ibid 379.

[41] Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 90 (Rome Statute) art 25(1).

[42] Thompson Chengeta, 'Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law' (2016) 45(1) Denver Journal of International Law & Policy 1, 37-38.

[43] UN Human Rights Council, 'Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie- Business and human rights: mapping international standards of responsibility and accountability for corporate acts' (9 February 2007) UN Doc A/HRC/4/035, para 20.

[44] Jose E Alvarez, 'Are Corporations Subjects of International Law?' (2011) 9(1) Santa Clara Journal of International Law 1, 9.

[45] ibid.

[46] Reparation for Injuries Suffered in the Service of the United Nations (Advisory Opinion) [1949] ICJ Rep 174.

on the international plane, and must also be legally equipped with the ability to enforce its rights by bringing claims that are enforceable in international law.[47] Responsibilities of corporate entities are certainly not unknown in international law.[48] Coupled with the fact that corporations often enjoy various rights and the power to enforce them under specialised regimes like international investment law, one can argue that transnational corporations should be granted international legal personality.[49] Yet, as has been discussed above, plausible arguments to the contrary also exist and have often been accepted in judicial decisions. Therefore, unless this hurdle of doctrinal uncertainty is cleared, it would be extremely risky to fix accountability for the actions of AI on the sole premise of corporate responsibility in international law.

Even if we assume that corporations can be legally sued for contravening international law, more obstacles arise due to the peculiar nature of AI technology. Due to the 'black box' nature of AI, it would be difficult to prove that high ranking officials of the corporation were involved in the breach caused by the actions of the AI entity. This is a prerequisite for fixing corporate criminal responsibility in many jurisdictions.[50] In the case of corporate civil responsibility, a heavy burden is placed upon the aggrieved party to file a civil suit before a foreign court.[51] More importantly, it is uncertain as to how product liability regulations would apply to AI entities,[52] since it would be difficult to prove that the manufacturer had foreseen the harm caused by a machine that can act autonomously, learn from its external environment, and function differently from how it was originally programmed.[53]

[47] ibid 179.

[48] See for example 'Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework' (2011) UN Doc HR/PUB/11/04, Principles 11-24.

[49] Karsten Nowrot, 'Transnational Corporations as Steering Subjects in International Economic Law: Two Competing Visions of the Future' (2011) 18(2) Indiana Journal of Global Legal Studies 803, 825-826.

[50] Geneva Academy of International Humanitarian Law and Human Rights, 'Autonomous Weapon Systems under International Law' (Academy Briefing No. 8, November 2014), 22 <https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf> accessed 16 May 2021.

[51] UN Human Rights Council (n 29) 15, para 79.

[52] ibid.

[53] Daniel N Hammond, 'Autonomous Weapons and the Problem of State Accountability' (2015) 15(2) Chicago Journal of International Law 652, 666-667.

In a nutshell, relying on corporate accountability is a risky venture, both due to doctrinal and practical shortcomings.

C. Individual criminal responsibility: holding the individual responsible for the functioning of AI accountable

The third potential target for fixing accountability is the individual who is responsible for the consequences arising out of the use of AI. In cases of armed conflict, individual responsibility or command responsibility can be seen as a potential avenue to fix accountability for the unlawful use of AWS technology. Concerns have been raised that AWS that can choose and attack targets at will without human intervention may breach the principles of proportionality and distinction, i.e. the obligation to distinguish between civilian and military targets and attack only the latter.[54] This would result in a serious violation of international humanitarian law (IHL)[55] and may even be regarded as war crimes.[56] It has, therefore, been contended that the person deploying the AWS and the commander authorizing or monitoring such conduct should be held criminally responsible.[57]

Under the Rome Statute, it must be proved that the person possessed the requisite intent and knowledge in committing the elements constituting the crime, for criminal liability to be invoked.[58] Similarly, under command responsibility, a superior commander is responsible for the actions of a subordinate, only if the commander knew that the subordinate was going to commit a violation of international law, and having known so, failed to make reasonable efforts to prevent the act, or punish for such a course of conduct.[59]

It is the fulfilment of these basic preconditions, coupled with the autonomy and unpredictability of AI systems, that make the imposition of criminal responsibility difficult. As Hammond rightly points out, a commanding officer, who had no role to play in the programming of the AWS, would not know how the machine would function in every conceivable situation, and whether it would violate the legal standards during armed conflict.[60] In such situations, proving

54 ibid 673-674.

55 Protocol Additional to the Geneva Convention of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I) (adopted 8 June 1977, entered into force 7 December 1979) 1125 UNTS 3 (Additional Protocol I) art 51(4).

56 ibid art 85. Note, however, that the acts have to be committed wilfully, in order to be classified as war crimes, as contemplated by Article 85.

57 Kelly Cass, 'Autonomous Weapons and Accountability: Seeking Solutions in the Law of War' (2015) 48(3) Loyola of Los Angeles Law Review 1017, 1066.

58 Rome Statute (n 41) art 30. Note, however, that the intent or knowledge does not have to be proved if such a requirement has been explicitly excluded by the Statute, as expressed by the wording of Art. 30.

59 Additional Protocol I (n 55) art 86(2).

60 Hammond (n 53) 664-665.

intention becomes a Herculean task.[61] In most cases, the prosecution can at most show that the commander or the deploying officer should have been more careful and undertaken due diligence measures to prevent harm. Yet, such a finding, even if proved, may be insufficient for conviction, since it is debatable whether mere negligence or recklessness satisfies the standard of the mental element required for committing international crimes.[62]

Undoubtedly, individual responsibility can be proved in those simpler cases where it can be shown that the manufacturer wilfully programmed the AI system in such a way that it would violate IHL, or the deploying officer used such technology with the intention to commit war crimes.[63] Realistically speaking, such simple fact scenarios are unlikely to materialise in practice.[64]

Moreover, holding the manufacturer criminally liable may be impractical under most circumstances. Besides the ability to learn from environmental stimuli and prior use,[65] most forms of AI technology have multiple uses, only some of which may result in violations of international law.[66] An argument on similar lines was accepted in the trial of persons connected with the affairs of I. G. Farben after the Second World War. The tribunal held that though it was shown that the company supplied the deadly Zyklon B gas to the Nazis, it could not be proved, in the absence of conclusive evidence, that the persons running the company could have known the purpose for which the gas was being used, namely, extermination of the victims in the concentration camps. According to the tribunal, the gas had a legitimate use as an insecticide, and it could reasonably be argued that the officials of the company felt that it would be used for that purpose.[67] Similarly, in the case of AI, it would be open for the manufacturer to contend that the use and functioning of the AI system as contemplated during its manufacture and programming would not have violated international law. It would be quite difficult to rebut this contention, given the fact that advanced AI systems can potentially 'learn' from its external environment, as discussed elsewhere in this article. Therefore, as in the

61 Rebecca Crootof, 'War Torts: Accountability for Autonomous Weapons' (2016) 164(6) University of Pennsylvania Law Review 1347, 1375-1376.

62 Carrie McDougall, 'Autonomous Weapon Systems and Accountability: Putting the Cart before the Horse' (2019) 20(1) Melbourne Journal of International Law 58, 67.

63 Crootof (n 61) 1376-1377.

64 Even in such situations, individual criminal responsibility does not preclude the possibility of State responsibility, and vice versa. See, Rome Statute (n 41) art 25(4); International Law Commission, 'Draft Articles on Responsibility of States for Internationally Wrongful Acts' (2001) 2(2) Yearbook of the International Law Commission 26, UN Doc A/CN.4/SER.A/2001/Add.1 (Part 2) (ILC Draft Articles) art 58.

65 Hammond (n 53) 666-667.

66 Chengeta (n 42) 40.

67 *United States v. Krauch ("The I. G. Farben case")*, 8 Trials of War Criminals before the Nuernberg Military Tribunals Under Control Council Law No. 10, 1169.

case of corporate accountability, individual criminal accountability cannot also be used to satisfactorily regulate AI on the international plane.

D. State responsibility: holding the state using AI, or allowing AI to be used, accountable

The last possible contender for fixing responsibility for the actions of AI systems, and possibly the most appropriate,[68] is the State. As has been rightly pointed out, the development and operation of autonomous systems involve the contribution of a large number of people, and it is difficult to pinpoint blame on a few isolated individuals to assign criminal responsibility.[69] In such group-centric activities, the proper course would be to hold the State responsible,[70] and make it liable to pay reparation.[71] This proposition is attractive, and is in conformity with the larger ideal of international law that States should be held accountable if they use, or allow to be used, its territory for infringing the rights of other States and their people.[72]

It is well settled that in order to hold the State responsible, the wrongful act or omission has to be attributed to it.[73] It is easy to guess that the major users of AI technology will be private actors, including corporations. Coupled with the autonomy of AI systems, the problem of attribution would often be insurmountable. A possible approach is to consider the AI system as an 'entity' under Articles 5 and 7 of the Draft Articles on State Responsibility, so that both foreseen and unforeseen conduct could be attributed to the State.[74] However, the problem with this approach is that the AI system must perform functions that fall within 'governmental authority', leading to a restriction on attribution.[75] In the *Nicaragua* case, the ICJ held that State responsibility would be incurred for the actions of non-State actors, only if it could be shown that the State enjoyed "effective control" [76] over them. In that case, the US was not held responsible for the

[68] In the context of AWS, see Hammond (n 53) 668-671.

[69] Charles P Trumbull IV, 'Autonomous Weapons: How Existing Law Can Regulate Future Weapons' (2020) 34(2) Emory International Law Review 533, 592.

[70] ibid.

[71] *Case concerning the Factory at Chorzow (Germany v. Poland)* (Merits) PCIJ Rep Series A No 9.

[72] *Corfu Channel (United Kingdom v. Albania)* (Merits) [1949] ICJ Rep 4, 22; *Trail Smelter case (United States v. Canada)* (1941) 3 RIAA 1905, 1965.

[73] ILC Draft Articles (n 64) art 2.

[74] Christopher M Ford, 'Autonomous Weapons and International Law' (2017) 69(2) South Carolina Law Review 413, 476.

[75] ILC Draft Articles (n 64) 43, Commentary to art 5.

[76] *Military and Paramilitary Activities in and Against Nicaragua* (*Nicaragua v. United States*) (Merits) [1986] ICJ Rep 14, 64-65, para 115.

actions of the Contra rebels, even though it had trained and funded the Contras, supplied them with weapons and even substantially taken part in the selection and attacking of targets. The Court observed that though the Contras were highly dependent on the US, the State was not responsible since it was not proved that the US directed and enforced the commission of illegal acts by the Contras.[77]

If such a fact scenario is juxtaposed with the use of modern AI technology, it can well be argued that private corporations are not under the "effective control" of the State, in that they are free to undertake activities and take decisions without the State directing them to do so. Even if AI technology is used by the State (without being regarded as organs of the State),[78] it could be legitimately argued that since the final decisions on whether and how to act in a given situation rests with the AI entity, the State cannot be said to exercise "effective control" over it. Therefore, it has been contended that the "effective control" test provides an avenue for the State to violate its international obligations by letting private entities commit unlawful acts, and plead lack of "effective control" to avoid responsibility.[79]

In the field of human rights,[80] it has been recognized that the obligation of States extends not only to respecting human rights, but also to ensure that private actors within its territory or jurisdiction do not violate human rights.[81] It is precisely due to this reason that obligations are imposed upon States to ensure that corporations do not commit acts that are in violation of international human

[77] ibid. But see, *Prosecutor v. Tadic* (Judgment) ICTY IT-94-1-A (15 July 1999) 56, para 131 holding that a lesser standard of "overall control" would be sufficient in the case of organized military groups, which would include not only funding and training, but also assisting in committing illegal acts. The International Court of Justice (ICJ), however, reverted back to the test laid down under Art. 8 of the ILC Draft Articles (n 64) art 8. See, *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro)* (Merits) [2007] ICJ Rep 43, 210, para 406.

[78] The actions of organs of the State can be directly attributed to the State. See, ILC Draft Articles (n 64) art 4; *Difference Relating to Immunity from Legal Process of a Special Rapporteur of the Commission on Human Rights* (Advisory Opinion) [1999] ICJ Rep 62, 87, para 62 stating that this principle has attained the status of customary international law.

[79] Mark Gibney, Katarina Tomasevski, and Jens Vedsted-Hansen, 'Transnational State Responsibility for Violations of Human Rights' (1999) 12 Harvard Human Rights Journal 267, 287-288.

[80] For an assessment of the impact of the use of AI by MNCs on basic human rights, see generally Emilie C Schwarz, 'Human vs. Machine: A Framework of Responsibilities and Duties of Transnational Corporations for Respecting Human Rights in the Use of Artificial Intelligence' (2019) 58(1) Columbia Journal of Transnational Law 232.

[81] UN Human Rights Committee, 'General Comment No. 31: The nature of the general legal obligations imposed on State Parties to the Covenant' (26 May 2004) UN Doc CCPR/C/21/Rev.1/Add.13 (General Comment 31) para 8; UN Human Rights Committee, 'General Comment No. 16: Article 17 (Right to Privacy), The Right to Respect of Privacy, Family, Home and Correspondence, and Protection of Honour and Reputation' (8 April 1988) UN Doc HRI/GEN/1/Rev.9 (Vol. I) paras 8-9.

rights law, instead of placing such a duty upon the corporations directly.[82] In this regard, States will incur international responsibility for the acts of private entities, if they fail to exercise due diligence in preventing the violation, and punishing the perpetrators.[83]

The logical corollary of this argument is that States would not be responsible for the actions of private individuals or entities, if it undertook due diligence efforts, irrespective of whether or not the violation took place.[84] In situations involving advanced technology, such as cyber activities and AI, directly attributing such actions to the State is challenging considering the complex mechanism of functioning that is involved.[85] Moreover, the problem with the due diligence approach in cases such as cyberattacks is that States can plead that though they employed the best possible means considering their resource constraints, the technology was too sophisticated for them to avert the damage caused.[86] In cases of AI, with its inherent uncertainty, States may argue that an isolated incident that could not be reasonably foreseen despite best efforts should not be cited as a ground to make them internationally responsible.

Even if the first hurdle of attribution is crossed, States would still be free to plead *force majeure* for those actions of the AI system that could not be anticipated,

[82] David Weissbrodt, 'Human Rights Standards Concerning Transnational Corporations and Other Business Entities' (2014) 23(2) Minnesota Journal of International Law 135, 154-156; UN Sub-Commission on the Promotion and Protection of Human Rights, 'Economic, Social and Cultural Rights: Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises with Regard to Human Rights' (26 August 2003) UN Doc E/CN.4/Sub.2/2003/12/Rev.2, para 1.

[83] *Velasquez Rodriguez v. Honduras*, Inter-American Court of Human Rights Series C, No. 4 (29 July 1988), 30-31, paras 172-174; *Social and Economic Rights Action Centre and Centre for Economic and Social Rights v. Nigeria*, African Commission on Human and Peoples' Rights, Comm. No. 155/96, 30th Ordinary Session (13-27 October 2001), paras 57-58 holding Nigeria responsible for the violation of rights of the Ogoni people by private oil companies and the State machinery; *Lopez Ostra v. Spain* [1994] ECHR 46 holding Spain responsible for not taking sufficient steps to prevent interference with the petitioner's right to respect for the home due to the activities of a private company; *Guerra v. Italy* [1998] ECHR 7 holding Italy responsible for not informing the petitioners of the risks involved due to the emission of toxic smoke by a factory owned by a private corporation.

[84] N L J T Horbach, 'The Confusion About State Responsibility and International Liability' (1991) 4(1) Leiden Journal of International Law 47, 57; Danwood Mzikenge Chirwa, 'The Doctrine of State Responsibility as a Potential Means of Holding Private Actors Accountable for Human Rights' (2004) 5(1) Melbourne Journal of International Law 1, 14-15.

[85] Peter Margulies, 'Sovereignty and Cyber Attacks: Technology's Challenge to the Law of State Responsibility' (2013) 14(2) Melbourne Journal of International Law 496, 502-504; Jovan Kurbalija, 'State Responsibility in Digital Space' (2016) 26(2) Swiss Review of International and European Law 307, 325.

[86] Ian Yuying Liu, 'State Responsibility and Cyberattacks: Defining Due Diligence Obligations' (2017) 4(2) Indonesian Journal of International and Comparative Law 191, 254.

in an attempt to evade responsibility.[87] Due to the uncertain nature of functioning of AI, it is possible that certain actions may be totally unforeseen and unexpected, leading to a potential argument of *force majeure*.[88] In the *Rainbow Warrior* arbitration, the tribunal held that the circumstance of *force majeure* would apply only when the prevailing situation makes the performance of the obligation an "absolute and material impossibility", and not merely because it would pose a heavy burden upon the State to discharge its duty.[89] In effect, the conduct is not borne out of the free will of the State, making it involuntary.[90] In the case of AI, situations may arise where its conduct could lead to total loss of control, and would make the performance of obligations utterly impossible.

However, *force majeure* does not apply if the State was responsible for the occurrence of the situation, or if the State had assumed the risk of the said situation arising.[91] The State cannot be held responsible if it, unknowingly and in good faith, contributed to the situation that arose; the situation must have been *caused* by its actions,[92] or through its neglect.[93] A possible example of *force majeure* would be the unlawful entry of aircraft, which got deflected due to atmospheric conditions, into the airspace of another State.[94] Such an example can be equated with autonomous AI systems, since its inherent unpredictability can be attributed to changes in its external environment and its complicated decision-making process, rather than being directed by the State. Moreover, even though the State can be said to have contributed to the situation by using or deploying the AI system in the first place, it would be difficult to prove causation in cases where the AI entity takes decisions and acts on them without any human intervention at any stage.

While it has been argued that force majeure should not apply when States neglect to regulate activities that can potentially cause harm,[95] it is almost impossible to predetermine the response of advanced AI entities to a stimulus before they are used, making regulation difficult.[96] The argument on assumption of risk may also be futile, since it mainly encompasses circumstances where the State had unequivocally agreed in advance to undertake the risk, or not plead force majeure, possibly through

87 Chengeta (n 42) 49.
88 ILC Draft Articles (n 64) art 23(1).
89 The Rainbow Warrior Affair (New Zealand v. France), (1990) 20 RIAA 215, 253, para 77.
90 ILC Draft Articles (n 64) 76, Commentary to Art. 23.
91 ibid Article 23(2).
92 ibid 78, Commentary to Article 23.
93 ibid 76-77.
94 ibid 77.
95 Myanna Dellinger, 'Rethinking Force Majeure in Public International Law' (2017) 37(2) Pace Law Review 455, 490.
96 Schwarz (n 80) 277.

its actions or international agreements.[97] Therefore, the existing legal standards governing State responsibility are insufficient for ensuring accountability for the actions of AI systems.

III. Absolute State Liability as a Possible Solution to the Problem of Accountability

The previous part shows the importance of creating a new regime of responsibility for fixing account of actions of AI. Some authors have argued in favour of a strict liability regime of State responsibility, especially in the field of AWS, by citing the inherently unpredictable and dangerous consequences flowing from the actions of AI.[98] The International Law Commission (ILC) too has tried to end the debate revolving around the need to prove fault in invoking State responsibility, by basing the nature of responsibility upon the content of the primary obligation involved.[99] Moreover, intention or the lack of it on the part of the State is irrelevant, except where it is a specific prerequisite for breach of the international obligation in question.[100]

What is being proposed in this paper is that States should be held responsible under an absolute liability regime, for the transnational consequences flowing from the use of AI entities by its organs, citizens, and corporate nationals. In effect, it would closely resemble the framework outlined in the Outer Space Treaty, 1967, according to which States are held responsible for "national activities" in space, including those carried out by private corporations and entities.[101] Such a step would preclude the necessity of using debatable legal propositions such as corporate responsibility or the legal personality of AI, and avoid the difficulty in attribution and countering the contention of *force majeure*. This mechanism recognizes that the major users of AI would be corporations and private actors.[102] Therefore, whenever corporations violate international law, they may be proceeded against by the State under their own legal framework, since corporate liability is well settled

97 ILC Draft Articles (n 64) 78, Commentary to Article 23.

98 Yannick Zerbe, 'Autonomous Weapons Systems and International Law: Aspects of International Humanitarian Law, Individual Accountability and State Responsibility' (2019) 29(4) Swiss Review of International and European Law 581, 604-605 (arguing for a high standard of State responsibility, bordering on absolute liability, in regulation of AWS, similar to the legal regimes that govern nuclear activities and outer space); Crootof (n 55) 1394.

99 ILC Draft Articles (n 64) 34-35, Commentary to Article 2.

100 ibid 36.

101 Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies (adopted 19 December 1966, entered into force 10 October 1967) 610 UNTS 205 art VI.

102 Note, however, that in the case of AWS, the State would generally be the sole user.

under municipal law. In order to finance reparations that the State would be liable to pay owing to the activities of private corporations or individuals, States would be completely free to devise their own methods, including similar absolute liability regimes under national law[103] for private entities using AI technology.

The proposed mechanism differs slightly from the "strict liability" frameworks found in existing literature on the topic. One strand of the existing literature deals with strict liability as understood under tort law, and therefore, AWS are sought to be regulated through the mechanism of "war torts".[104] The second strand proposes a more radical model where *force majeure* and other circumstances precluding wrongfulness are not permitted to be raised as possible defences, but which generally focusses on the international humanitarian law (IHL) aspects of AWS regulation.[105] The problem with the former is that strict liability under tort law recognises defences such as act of God,[106] which can possibly backfire due to the unpredictable nature of functioning of AI. Similarly, the latter being focussed only on the IHL aspects of AWS, cannot be satisfactory employed to regulate the use of AI for peaceful purposes and that too, by private actors acting independently of the State. Therefore, the proposed model in this article would not allow defences like *force majeure* or act of God to be raised, while also ensuring that the State is held responsible for the acts of its citizens and corporate nationals acting independently. As a result, the proposed model is much more expansive as far as the extent of State liability is concerned.

States are normally responsible for activities occurring within their territory.[107] Yet, the proposed framework would naturally entail a certain degree of extraterritoriality, since MNCs operating abroad would be the most frequent users of AI. In this respect, pragmatic concerns dictate that the home State (where the MNC or its parent company is incorporated or registered) should shoulder the responsibility, instead of the host State (where the MNC or its subsidiaries operate) doing so. Host States, which are generally developing nations with resource constraints, would often be powerless to act against influential MNCs. Invoking responsibility of States which do not have the capacity to regulate powerful

[103] For example, in India, industries that are engaged in hazardous activities bear absolute liability for the consequences arising out of their activities. The traditional defences against strict liability such as act of God, or fault of the victim, are not available. See, *M. C. Mehta v. Union of India*, (1987) 1 SCC 395, 420-421, para 31.

[104] Crootof (n 61) 1394.

[105] Zerbe (n 98) 604.

[106] *Rylands v. Fletcher* (1868) LR 3 HL 330, 339-340.

[107] See, e.g., Convention on Nuclear Safety (adopted 17 June 1994, entered into force 24 October 1996) 1963 UNTS 293, Preamble (iii) the State exercising jurisdiction over a nuclear installation is responsible for ensuring nuclear safety.

corporations would be unfair.[108] Further, as McCorquodale and Simons argue, home States provide vital concessions to their corporate nationals in the form of loans, assistance through export credit agencies, and investment guarantees. Consequently, they should be held responsible if the actions of such corporations are such that they would have led to State responsibility if they had been committed by the home State directly.[109]

Moreover, although the parent company of an MNC and its foreign subsidiaries are legally distinct from each other, State practice has shown that home States of the parent company often regulate the activities of its foreign subsidiaries operating in other countries.[110] For instance, in certain cases where the subsidiary was economically dependent upon the parent company, the actions of the former had been attributed to the latter by treating both of them as a single economic unit in the field of competition law.[111] Recently, the Court of Appeals in France allowed charges of financing of terrorist outfits to be framed against the French company Lafarge for payment of money to the Islamic State by its Syrian subsidiary.[112] These cases show that 'piercing the corporate veil' to hold parent companies accountable for the misdeeds of its foreign subsidiaries is not unheard of.

Even under the Outer Space Treaty, it has been argued that the State where the corporation or the parent company (in case of MNC) is registered, should be the State that would be held responsible under Article VI of that treaty.[113]

[108] Chirwa (n 84) 26-28.

[109] Robert McCorquodale and Penelope Simons, 'Responsibility beyond Borders: State Responsibility for Extraterritorial Violations by Corporations of International Human Rights Law' (2007) 70(4) *Modern Law Review* 598, 613-614.

[110] ibid 616-617.

[111] Case 48/69, *Imperial Chemical Industries Ltd. v. Commission of the European Communities* (1972) ECR 619.

[112] Claire Tixeire, Cannelle Lavite and Marie-Laure Guislain, 'Holding Transnational Corporations Accountable for International Crimes in Syria: Update on the Developments in the Lafarge Case (Part I)' (*Opinio Juris*, 27 July 2020) <http://opiniojuris.org/2020/07/27/holding-transnational-corporations-accountable-for-international-crimes-in-syria-update-on-the-developments-in-the-lafarge-case-part-i/> accessed 18 July 2021.

[113] Kofi Henaku, 'Private Enterprises in Space Related Activities: Questions of Responsibility and Liability' (1990) 3(1) Leiden Journal of International Law 45, 51-52. For the subtle nuances governing the issue of State responsibility in this field, see generally Krystyna Wiewiorowska, 'Some Problems of State Responsibility in Outer Space Law' (1979) 7(1) Journal of Space Law 23.

Similarly, in the field of human rights, extraterritoriality is not unprecedented in international law.[114]

Under the proposed framework, although States would not be permitted to plead force majeure, they should be exempted from responsibility if they can prove that the damage caused by the AI system was the direct result of an act of agencies, private individuals, or corporations within the jurisdiction of another State[115]. In this case, the latter State should be held responsible. This exception is especially important in the field of AI systems, in order to provide a disincentive against hacking of AI software, and stopping culpable actors from claiming reparations for harm arising out of their own fault.

IV. Conclusion

Due to the autonomy and unpredictability of AI, traditional modes of accountability would fail miserably, highlighting the need for a novel approach that does not ignore the astronomical pace of technological development. In a primarily State-centric international legal order, the onus should fall upon States, with their massive regulatory and enforcement powers, to take responsibility for the perils that AI has to offer. In this respect, a model of absolute State liability on the international plane has been proposed in this paper, which is flexible enough to allow States to devise tailor-made strategies in the domestic sphere, in accordance with their unique national circumstances.

The question arises as to why States might be willing to bind themselves within such a restrictive liability framework. The answer is obvious: the only alternative to legal regulation of AI is a complete ban on such technology. Activists all over the world have called for banning AI technology in the field of facial recognition, AWS, and algorithmic vetting of asylum seekers, focusing primarily on the adverse impact such systems have on human rights. The voices will only grow louder unless an acceptable mode of accountability is arrived at. In the midst of such campaigns, States may wish to pay a small price in order to enjoy the benefits that AI can offer, both material and strategic.

The restrictive agreements relating to outer space form the basis for the proposed model in this paper. When those agreements were signed, outer space

[114] See for example United Nations Committee on Economic, Social and Cultural Rights (UNCESCR), 'General Comment No. 15: The Right to Water (Arts. 11 and 12 of the Covenant)' (20 January 2003) UN Doc E/C.12/2002/11, para 33 calling upon States to ensure that its own citizens and corporate nationals do not violate the right to water of people living in other nations.

[115] As regards liability for damage, a similar exception exists in the regime governing space law. See, Convention on the International Liability for Damage Caused by Space Objects (adopted 29 November 1971, entered into force 1 September 1972) 961 UNTS 187, art VI.

constituted the unknown, offering opportunities as well as dangers. Similarly, AI technologies today offer infinite advantages and potentially catastrophic consequences. Coupled with limited knowledge about their inner workings, this constitutes a suitable case for experimenting with absolute State liability, through a binding international legal instrument.