NAME: TABITHA MBUTHIA

ORGANIZATION: Chemicals and Waste Youth Platform-MGCY

POSITION: Human Rights Focal Point

DATE: 24TH FEBRUARY, 2023

TOPIC: Peace and Security: Given the risks of AI and emerging technologies, how can the global governance system address these risks without stifling potential benefits? The World Bank identifies vertical inequalities as a major variable to conflict, what would be a viable proposition in conflict prevention to address this issue? How could we increase transparency in international security governance and architecture? Other specific proposals and innovations on ways to advance the youth, and wider peace and security agenda(s) through global governance reform?

Table of Contents

1.0 Introduction 3
2.0 Benefits of using AI 3
3.0 Risks Involved in Artificial Intelligence 4
4.0 Solutions to AI related Issues in the Security Sector 4
4.1 Engaging Stakeholders Effectively 4
4.2 Broadening Existing Platforms of Multilateral Engagement 5
4.3 Craft Suitable Regulations 5
4.4 Enhance Transparency, Oversight and Accountability. 5
4.5 The Application of Existing Laws 6
4.6 The adoption of a new treaty to govern AI in Security. 6
4.7 Identify a UN focal point on cyber issues 7
5.0 Conclusion 8

TWO SIDES OF A COIN; A CASE STUDY OF ARTIFICIAL INTELLIGENTS' BENEFITS, RISKS AND THE SOLUTIONS THEREIN.

1.0 Introduction

There have been massive technological advances in the past century trickling down to today.[1] This is with the inclusion of Artificial Technology (AI) whose definition dates back to 1955 as "making a machine behave in ways that would be called intelligent if a human being were so behaving."[2] It is simply the simulation of human intelligence processed by machines.

AI is diverse and encompasses subdisciplines such as natural language processing, machine inference, statistical machine learning and robotics.[3] It has been a major concern that AI will lead to superintelligence that will either achieve or surpass human intelligence.[4] Regardless, it is expected to bring about great benefits to almost all spheres of life.

This paper shall therefore seek to do a study on the risks inherent in the AI technology in the way they are designed and used as well as their converse benefits. This is all in relation to how the global governance system is supposed to address these issues as far as peace and security is concerned.

2.0 Benefits of using AI

AI has proven to be beneficial in the security sector ranging from national to international influence. Such benefits include but are not limited to: AI providing opportunities to collect data about crime and conflict and reduce the gap between warning and response. For example, crisis mapping, social media mapping, and crowdsourcing tools can help generate data on conflict indicators. The data generated from these tools can help identify patterns associated with conflict and peace in order to better inform conflict prevention efforts, or to monitor violations of cease-fires or human rights.

[1] Camino Kvanagh,"New Tech, New Threats, and New Governance Challenges: An opportunity to Craft Smarter responses?"(2019).
[2] John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude E. Shannon, "Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," *AI Magazine* 27, no. 4 (2006).
[3] Rodney Brooks, "The Origins of Artificial Intelligence," FoR&AI, April 27, 2018.
[4] Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence*, (New York: Knopf, 2017).

3.0 Risks Involved in Artificial Intelligence

The consequences of using AI is that it converges with other technologies such as biotech and nuclear domains with superior algorithmic discrimination, weak transparency and accountability especially in the decision-making processes. Moreover, there is a lot of limitation in conceptualizing ethical problems and investment in safety research and protocols.[5]

Due to the fact that AI derives its information from large quantities of collected, stored and processed data, there is also a lot of concern over data protection, privacy, transparency and accountability. It is also difficult to constrain their development and regulate their use due to the dual-use nature of AI. This means that they can be used to serve malicious purposes and enhance social and economical developments rendering efforts to manage them more complex.[6] To that end, many countries are deriving their military power form AI.[7]

They are relatively easy to access and use thus making them inherently vulnerable to exploitation and disruption. The general unpredictability surrounding AI algorithms and their susceptibility to bias, theft, and manipulation are all expected to pose national security risks. Deepfakes, which are "realistic photo, audio, and video forgeries" that may be used for "information operations," are one particularly perplexing potential tool of manipulation.[8] They also provoke disruptions of legal and regulatory orders available.

4.0 Solutions to AI related Issues in the Security Sector

AI poses a threat to national and international security. However, there are certain solutions that can be applied by integrating stakeholders, governments, regional and international bodies. They are:

4.1 Engaging Stakeholders Effectively

The control processes for AI must include legislators, regulators, researchers, and civil actors. Yet, it might be difficult to integrate all of these groups' specific policy recommendations and put them into practice on the security front. This is particularly true when using the solutions to a

[5] Lassonde School of Engineering, "Can AI Help Feed the World? The Future of Food is Here," *Medium,* April 16, 2018.
[6] Camino Kavanagh and Paul Cornish, "Preventive Diplomacy, ICT and Inter-State Conflict: A Review of Current Practice with Observations," *Swiss Federal Department of Foreign Affairs*, (forthcoming 2019).
[7] Elsa Kania, "Great Power Competition and the AI Revolution: A Range of Risks to Military and Strategic Stability," *Lawfare*, September 19, 2017.
[8] Kelley M. Sayler, "Artificial Intelligence and National Security," Congressional Research Service, Updated January 2019.

variety of cross-border security issues. Determining which AI concerns can be solved domestically and which ones necessitate cross-border coordination and cooperation requires the necessary parties to clearly define domestic and international duties.

4.2 Broadening Existing Platforms of Multilateral Engagement

It is crucial to clarify how different actors (and not only states) can contribute to the functioning of multilateral mechanisms focusing on international law or political norms. These forums include the various United Nations (UN) groups of Governmental Experts (GGEs) such as the CCW and other working groups for International Security and Technology (including ICT and autonomous weapons). These initiatives are flexible enough to recognize when new international norms or standards (binding or non-binding ones) are required to manage AI-driven challenges and risks and constrain certain applications of these technologies or specific behaviours by states and other actors.

4.3 Craft Suitable Regulations

Twenty Eight nations have endorsed the proposal for a prohibition on fully autonomous weapons since 2017.[9] Eighty-five states "publicly clarified their views on deadly autonomous weapons systems in a multilateral arena" during the meeting of the CCW GGE's first session in 2018, with some promoting the creation of a legally enforceable agreement on fully autonomous weapons.[10]

Moreover, new methods of policy and regulation are required. For instance, developments in AI are already influencing complex conversations about the focus and justification of pertinent policies and the best kind of regulation (precautionary, preventive, reactive, or a combination of all three). Choosing between strict regulation, soft policy initiatives (such as principles, certification processes, and labeling systems) that fall short of binding legislation, and the present trend of self-governance measures by private enterprises raises a number of additional questions. Uncoordinated national norms and policies are likely to be ineffectual given the cross-border implications of the technologies in use and their increasing convergence.

4.4 Enhance Transparency, Oversight and Accountability.

More money will need to be spent on accountability, supervision, and transparency systems as a result of changing policy and regulatory measures. It will be necessary to decide whether

[9] Ibid.
[10] Campaign to Stop Killer Robots, “Country Views on Killer Robots,” April 13 2018.

national regulatory monitoring organizations should be public, private, or comprised of both public and private sectors. Additionally, it should be ensured that tech-related businesses and organizations consent to increased inspection.

Tech companies should, for instance, intensify internal oversight and external reporting of their self-regulatory efforts, give appropriately shielded, publicly funded researchers secure access to their data, and, most importantly, make sure that accountability extends to every link in the supply chain and that both the direct and indirect costs (such as those associated with labor and the environment) of the relevant technologies are understood.[11] A

n equally crucial tool for giving some of these processes legitimacy is agreement on what is morally acceptable in terms of industry funding and participation in monitoring bodies (such as ethics councils and advisory boards). Such examination would also aid in identifying any gaps that still exist and would involve more actors in determining whether and how a specific technology or its application should be controlled.

4.5 The Application of Existing Laws

There is need to create greater clarity and consensus on how to apply the existing laws to determine the frameworks that govern offensive state action by usage of AI and through new forms of warfare. The UN should take the lead in this issue.

An example is the GGE on Developments in the Field of Information and Telecommunications in the Context of International Security, established under the auspices of the UN General Assembly.[12] it unanimously concluded that international law, particularly the UN Charter, is applicable in cyberspace.[13]

4.6 The adoption of a new treaty to govern AI in Security.

Although there is existence of laws, the issue is that they may not be sufficient to deal with AI security threats. Both states and scholars have proposed a new treaty to address these issues. However, any attempt to create new cybersecurity laws will require policymakers to address three major underlying issues. First, they will have to consider which actors to address. Most

[11] Yochai Benkler, "Don't Let Industry Write the Rules for AI," *Nature*, May 2019; Vol. 569 (7755)
[12] , A/RES/57/53, A/RES/62/17, A/RES/65/41, A/RES/68/243.
[13] Group of Governmental Experts, *Report on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN Doc. A/68/98, June 24, 2013

existing laws focus on private actors without distinguishing between their motives, but it may be best for a different set of rules to apply when cyberattacks originate from a state. There is also a question of whether to distinguish between attacks by cybercriminals and attacks by cyberterrorists.[14]

Second, if policymakers put in place different rules for different actors, they must be able to attribute each act to determine which set of rules applies. Attributing cyberattacks is difficult, however, and simply determining an attack's source may not be enough to determine who is responsible. If governments are too careful to attribute, this could undermine attempts to hold those violating laws accountable.[15]

Third, policymakers must address the relationship between cybersecurity and human rights. In the Cybercrime Convention, for example, activists fear that grouping together crimes merely committed on the Internet and those for which the Internet is central opens the door to content controls. This highlights questions about the extent to which a new cybersecurity treaty would be able to safeguard human rights around the world. Existing guidance on human rights in the digital age developed within the UN system would likely have to be included as part of any such treaty.[16]

4.7 Identify a UN focal point on cyber issues

With ongoing efforts in regional bodies such as NATO, the Organisation for Economic Cooperation and Development (OECD), the Asia-Pacific Economic Cooperation (APEC), the Organization for Security and Cooperation in Europe (OSCE), the Organization of American States (OAS), and the Council of Europe, there is a risk that collective regional approaches to questions of sovereignty and jurisdiction will harden the stances of member states in negotiations at the UN. The appointment of a clear focal point within the UN system for particularly pressing discussions might help avoid such a situation. This focal point could also function as a test case for the establishment of other focal points as the Internet governance system evolves and more issues come up for discussion in the UN.

4.8 Make the UN the depository and safe-keeper of big data

[14] Anja Kovacs, "Addressing India's Global Cybersecurity Concerns: Norm Development, Regulatory Challenges, Alternative Approaches," Internet Democracy Project, August 18, 2015.
[15] Kovacs, "Addressing India's Global Cybersecurity Concern."
[16] Ibid.

The UN could help gather, collect, and store data, especially from regions where the infrastructure is not safe or sufficient. Member states could give this mandate to the UN, which would have to create and implement safeguards for the data.

5.0 Conclusion

All in all, it is important to recognize the essential role that Artificial Intelligence plays in the world and its evolving nature. It poses a major risk to the security sector both nationally and internationally. However, there are certain steps that can be taken to mitigate these risks that should be abided by. Therefore, AI will be used to its full potential and in a positive manner.