Artificial Intelligence and Lethal Autonomous Weapons Systems

Major Mark Smith

Canada

CANADIAN FORCES COLLEGE – COLLÈGE DES FORCES CANADIENNES

JCSP 46 DL – PCEMI 46 AD
2019 – 2021

SOLO FLIGHT

ARTIFICIAL INTELLIGENCE AND LETHAL AUTONOMOUS WEAPONS SYSTEMS

By Major Mark Smith

ARTIFICIAL INTELLIGENCE AND LETHAL AUTONOMOUS WEAPONS SYSTEMS

<u>Introduction</u>

Currently western debate regarding Artificial Intelligence (AI) and its linkages to Lethal Autonomous Weapons Systems (LAWS), remains at loggerheads between those who perceive that AI weapons systems are *mala en se*, they are evil and constitute a moral wrong, and those who believe they must be developed in order to maintain a competitive advantage with our potential adversaries such as Russia and China. [1] There have been calls for outright bans and laws preventing the developing and use of LAWS as they are too destabilizing and could permit battles to be fought in larger scales and faster temporally than humans can comprehend. [2] AI systems are systems that maybe used to make life or death decisions and include LAWS, systems that can kill targets without human intervention, and Decisions Support Systems (DSS), systems that can assist with complex decision making. [3] AI systems can help commanders, staffs, and partners share information and make decisions in operational environments that are volatile, uncertain, complex, and ambiguous (VUCA). [4]

AI cannot act as a moral agent due to technological limitations since it cannot replicate the learning functions that a human brain possesses. [5] Machine learning is the process of analysing training data and the subsequent learning and detection of patterns that can

[1] Pfaff, Anthony. *The Ethics of Acquiring Disruptive Technologies Artificial Intelligence, Autonomous Weapons, and Decision Support Systems.* (Prism 8, no.3), 130.

[2] Jenkins, Cameron. *AI Innovators Take Pledge Against Autonomous Killer Weapons.* Last accessed 27 May 21. Google AI Chiefs, Elon Musk Sign Pledge Against Autonomous Weapons : NPR, 1-2.

[3] Pfaff, Anthony. *The Ethics of Acquiring Disruptive Technologies Artificial Intelligence, Autonomous Weapons, and Decision Support Systems.* (Prism 8, no.3), 130.

[4] Ploumis, Micahil. *Mission Command and Philosophy for the 21ˢᵗ Century.*(Comparative Strategy, 39:2. 2020), 210.

[5] Pfaff, Anthony. *The Ethics of Acquiring Disruptive Technologies Artificial Intelligence, Autonomous Weapons, and Decision Support Systems.* (Prism 8, no.3), 130.

then be used in decision-making. [6] This learning does contain errors, biases, and the end results may need additional analysis on potential impacts to humans, hence the need for ethical considerations to be applied to protect human dignity. [7] The application and study of AI and ethics is a subfield of applied ethics and it is concerned with the ethical issues surrounding the development, deployment, and use of AI. [8]

Despite the need and demand for ethical principles applied to AI, and to AI and LAWS, ethical standards do not always influence behaviour as evidenced by Googles cancellation of Project Maven due to backlash from its employees due to its perception of being biased and weaponized AI. [9] Additionally, questions have been raised on which applicable and interested entities and which set of ethics should be applied. [10] From a military perspective, the development and use of AI is now at a state that required command engagement as it is military commanders who can overcome gaps in expertise and provide meaningful, ethical and lawful direction on AI's use. [11]

This paper will discuss briefly how machines learn, on going developmental principles, the applicability of International Humanitarian Law, and ethics surrounding the use of AI and LAWS to kill humans. The debate of using AI and LAWS has been on going for decade but limits of technology have prevent this from happening however it can be expected that AI

[6] Latonero, Mark. *Governing Artificial Intelligence: Upholding Human Rights and Dignity.* (Data and Society. Annenberg School for Communication, University of Southern California, 10 October 2018). 8.

[7] *Ibid, 8.*

[8] Salles, Arleen, Evers, Kathinka & Farisco, Michele. *Anthropomorphism in AI.* (AJOB Neuroscience, 11:2. 2020). 89.

[9] Hao, Karen. *Establishing an AI Code of Ethics Will be Harder Than People Think.* (MIT Technology Review. 21 October 2018).

[10] *Ibid.*

[11] Losito, Marc and Anderson, John. *The Department of Defense's Looming AI Winter.* (War on The Rocks. 10 May 2021).

systems will become smart enough to make life or death decisions without a human in, or on, the decision loop. Western adversaries such as China and Russia are continuing the development of AI and LAWS, western countries run the risk of losing a competitive advantage if we do not explore and develop AI and LAWS. Initial United Nations reporting indicates that an AI and LAWS system, the *Kargu 2* was used by Turkish forces in Libya since January 2020. [12] The thesis of this paper is: Why should the CAF understand and adopt AI technologies with respect to LAWS using legal and ethical frameworks or it risks loosing its standing with peer nations and risks being put at a competitive disadvantage with adversaries. The use of AI and LAWS is at a point where it must move from the philosophical and technical realms into military command spheres and this topic now requires military commanders attention and understand so as to keep moving into the future.

How Do Machines Learn?

Autonomy can be described as a spectrum that has three main features: the type of task being performed, the human-machine relationship, and the level of decision making sophistication. [13] An AI system can be more autonomous by increasing its autonomy along this spectrum but as it becomes more autonomous it may become less predictable because how the AI system determines the end-state of its task is not yet fully understood. [14] AI and LAWS are autonomous not because of intelligence but because of freedom, and how much freedom humans allow them to have; intelligence can be added to the system but it is trusting the system and

[12] United Nations Security Council. *Final Report of the Panel of Experts on Libya Established Pursuant to Security Council Resolution 1973 (2011).* 11 March 2021. Last accessed 2 June 2021. https://undocs.org/S/2021/229

[13] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War. (*W.W. Norton and Company. New York | London. 2018), 57.

[14] *Ibid, 62.*

giving it freedom is what makes it more and more autonomous. [15] Commanders need to understand and trust AI to make correct decisions before they employ it to its full capabilities along the spectrum. [16]

Large data sets, also know as big data, fast processing speeds, algorithms and deep neural networks (DNN) can be combined to train machines to learn and this learning can further lead to decision support and decision making. [17] Deep Learning is a technique that uses neural networks to train algorithms to conduct tasks. [18] DNN uses a limited amount of simplified, homogenous artificial neurons that lack consciousness. [19] (The creation of DNN is inspired by the neural networks of the human brain and functions in a basic, similar manner but DNN does not yet have the same capabilities as a human brain's neural network; AGI may bridge the gap between DNN and a human brain. [20]

Neural networks work based on connections within a network and the connections from input data that train the connections to learn. [21] DNN take this a step further and have multiple hidden layers between connections and adding more connections allows DNN to learn more and handle increasing complex tasks. [22] DNN are not difficult to acquire, some can be purchased

[15] *Ibid, 90.*

[16] *Ibid, 141.*

[17] Scharre, Paul. *Highlighting Artificial Intelligence. (*Strategic Studies Quarterly. 11, no 4. Winter 2017).

[18] Horowitz, Michael. *Artificial Intelligence, International Competition and the Balance of Power.* (Texas National Security Review. 1, no. 3. May 2018), 4.

[19] Salles, Arleen, Evers, Kathinka & Farisco, Michele. *Anthropomorphism in AI.* (AJOB Neuroscience, 11:2. 2020), 98.

[20] *Ibid, 20.*

[21] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* (W.W. Norton and Company. New York | London. 2018), 147.

[22] *Ibid, 20.*

from open source civilian companies such as Tensor Flow. [23]

AI systems can learn through trial and error which is a departure from learning by studying previously existing data and then determining predictions based on that data. [24] This type of creative and strategic learning is called reinforcement learning and it works by an AI trying different solutions and then adopting the solutions that work better than others. [25] With enough trial and error, AI can adopt and reinforce solutions to problems and is applicable when decisions need to made in sequence and when an action, not a prediction, is required; it is currently used in the financial sector to buy and sell stock. [26]

Currently, AI is at a point too early to determine if they are evil but not at a point that makes them fully permissible to kill humans, militaries must continue to develop and explore AI's use and functionality to better understand and become comfortable with their limitations and use. [27] AI and LAWS cannot yet be held accountable for errors due to lack of full understanding of how they work. [28] Failure to continue AI development will see western militaries fall behind countries that are developing its use, such as China and Russia, and lose our competitive advantage. [29] Debates on banning the use of AI and LAWS exist but these

[23] Tensor Flow. Last accessed 2 June 2021. https://www.tensorflow.org/

[24] Hume, Kathryn and Taylor, Matthew E. *Why AI That Teaches Itself to Achieve a Goal Is the Next Big Thing.* Last accessed 30 May 2. April 21, 2021. https://hbr.org/2021/04/why-ai-that-teaches-itself-to-achieve-a-goal-is-the-next-big-thing

[25] *Ibid,.*

[26] *Ibid,.*

[27] Pfaff, Anthony. *The Ethics of Acquiring Disruptive Technologies Artificial Intelligence, Autonomous Weapons, and Decision Support Systems.* (Prism 8, no.3), 135.

[28] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* (W.W. Norton and Company. New York | London. 2018), 414.

[29] Pfaff, Anthony. *The Ethics of Acquiring Disruptive Technologies Artificial Intelligence, Autonomous Weapons, and Decision Support Systems.* (Prism 8, no.3), 135.

discussions limit and create a premature judgement of future technologies. [30]

Issues and problems with data can lead to AI not learning correctly and then make poor or incorrect decisions based on the bad data and they also require extensive and well formatted data sets. [31] This is similar to a human brain learning improperly due to poor or incorrect learning environments. Errors in machine learning can be categorized in three ways: disparate treatment is a decision making process error that has outcomes change due to the change in a specific value while all other values remain the same. [32] A disparate impact is a decision making error when a disproportionally large group of positive outcomes to specific groups and thirdly, a decision making process has disparate mistreatment when its accuracy rate is different for different specific groups. [33]

Present day AI systems are not able to process large amounts of data stemming from complex environments but in the future it maybe possible that AI systems will be able to process large amounts of disparate data. [34] AI can be conceptually categorized in the following manner: Artificial Narrow Intelligence (ANI) which currently exists, Artificial General Intelligence (AGI) which does not currently exists but would be similar to human levels of intelligence, and Artificial Super Intelligence (ASI) which does not exist but would be beyond human levels of intelligence. [35] To date, no AI systems is capable enough to autonomously apply the

[30] Price, Mathew, Walker, Stephen, and Wiley, Will. *The Machine Beneath. Implications of Artificial Intelligence in Strategic Decision Making.* (Prism 7, No 4. 8 November 2018), 97.

[31] Daniels, Owen, and Williams, Brian. *Day Zero for Military AI.* (War on the Rocks. 28 January 2020).

[32] Zafar, Muhammad Bilal; Valera, Isabel; Gomez-Rodriguez, Manuel; Gummadi, Krishna P. *Fairness Constraints: A Flexible Approach for Fair Classification.* (Journal of Machine Learning Research 20, 2019), 2.

[33] *Ibid,.* 2.

[34] Price, Mathew, Walker, Stephen, and Wiley, Will. *The Machine Beneath. Implications of Artificial Intelligence in Strategic Decision Making.* (Prism 7, No 4. 8 November 2018), 96.

[35] *Ibid,.* 96.

Laws of Armed Conflict (LOAC) principles of military necessity, distinction, and proportionality. [36] It is estimate that functional AGI systems are 20-30 years away and it is also assessed that AGI and ASI systems would be difficult to explain and document due to the complexity of their deep learning reasoning systems. [37] AGI could be used to solve problems across a range of cognitive tasks and could also factor nuance, ambiguity, and uncertainty. [38] The transition timeline from AGI to ASI could be minutes to decades but researchers do not yet know for certain. [39]

Developmental Principles.

The United States Department of Defense (DOD) in 2018 is exploring ways to build a framework for AI and one component is five ethic principles: (a) Responsible: humans should remain responsible for AI development, deployment, use and outcomes; (b) Equitable: avoid unintended bias that would cause harm; (c) Traceable: experts maintain understanding fo the technologies; (d) Reliable: AI systems are defined, tested and quality assurance is maintained; and (e) Governable: humans or automated deactivation can occur if systems demonstrate unintended behaviours. [40] The Government of Canada is in the process of developing guiding principles for AI, broadly applicable to all federal departments. [41] These principles are in line

[36] *Ibid,. 96.*

[37] Galliot, Jai and Scholz, Jason. *Artificial Intelligence in Weapons: The Moral Imperative for Minimally- Just Autonomy.* (Air Force Journal of Indo-Pacific Affairs. 1, no 2. Winter 2018), 61.

[38] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* (W.W. Norton and Company. New York | London. 2018), 369.

[39] *Ibid,. 372.*

[40] Vergun, David. *Defense Innovation Board Recommend AI Ethical Guidelines.* Last accessed 26 May 2021. https://www.defense.gov/Explore/News/Article/Article/2006646/defense-innovation-board-recommends-ai-ethical-guidelines/

[41] Government of Canada. *Responsible used of Artificial Intelligence.* Last accessed 26 May 21. Responsible use of artificial intelligence (AI) - Canada.ca

with other countries and consist of: (a) understand and measure AIs impacts; (b) transparency when AI is used; (c) explain AI decision making; (d) be open with source code, data and information but protecting personal information and national security; and (e) provide training to users. [42] The CAF *Strong Secure Engaged* publication does not mention AI. The Treasury Board Secretariat has created the Directive on Automated Decision-Making which is the GOCs overarching policy on using AI to make or help make decisions in a manner that is in line with administrative law principles of transparency, accountability, legality and procedural fairness. [43] Much like financial accountability, and CAF AI initiatives would have to fall under the TBS's Directive on Automated Decision-Making.

AI is not limited to LAWS or DSS but can also be used to analyze Intelligence, Surveillance, and Reconnaissance (ISR) data, controlling sensor platforms and automated target recognition (ATR); functions that until now, were limited due technology. [44] AI use in supporting decision making will continue to evolve and be used, as AI systems become more intelligent, think faster than humans about complex tasks, and humans become more trusting and accepting of the AI systems themselves. [45] AI systems will be able to provide military commanders better understanding of complex environment across multiple domains at a rate faster than an adversary. [46]

[42] *Ibid,.*

[43] Government of Canada. Directive on Artificial Decision Making. Last accessed 26 May 21. Directive on Automated Decision-Making- Canada.ca (tbs-sct.gc.ca)

[44] Geist, Edward and Andrew J. Lohn. *How Might Artificial Intelligence Affect the Risk of Nuclear War?* (Santa Monica, CA: RAND Corporation, 2018), 9.

[45] *Ibid,.* 18.

[46] Price, Mathew, Walker, Stephen, and Wiley, Will. *The Machine Beneath. Implications of Artificial Intelligence in Strategic Decision Making.* (Prism 7, No 4. 8 November 2018), 96.

The progress of AI sees four broad themes: superintelligence, limited breakout, continuous incremental progress, and AI plateau. [47] Superintelligence is a state where machine intelligence significantly matches human intelligence but it is not anticipated to be fielded in the short term. [48] A limited breakout scenario would see AI become almost as intelligence as humans, through discontinuous jumps in progress, as long as software and hardware are capable of supporting it. [49] Continual incremental progress of AI would result due to constant advancement of software and hardware and is viewed as the most likely of the four themes to occur. [50] AI Plateau is not viewed as a likely scenario due to increasing development of AI but it is one that needs to be considered. [51]

International Humanitarian Law

A common theme amongst authors is using International Human Rights as a key component guiding the ethical use of AI, not only AI and LAWS to best protect and respect human dignity. [52] From a Canadian perspective, the Treasury Board's Algorithmic Impact Assessment does contain a provision to address individual liberty rights. [53] The International Bill of Rights ratified by approximately 170 countries state that human rights are indivisible,

---

[47] Geist, Edward and Andrew J. Lohn. *How Might Artificial Intelligence Affect the Risk of Nuclear War?* (Santa Monica, CA: RAND Corporation, 2018), 12.

[48] *Ibid,*. 13.

[49] *Ibid,*. 13.

[50] *Ibid,*. 13.

[51] *Ibid,*. 13.

[52] Latonero, Mark. *Governing Artificial Intelligence: Upholding Human Rights and Dignity.* (Data and Society. Annenberg School for Communication, University of Southern California, 10 October 2018). 5.

[53] *Ibid,*. 13.

interdependent, and inter-related. [54] Human rights are also further defined as having five themes: non-discriminatory, having equality, able to participate in politics, assure privacy, and have freedom of expression. [55] The principle of avoiding unnecessary suffering has also been suggested as another IHL to consider. [56] All of these features of human rights are difficult to define, complex, nuanced, open to subjectivity and interpretation. The same attributes that when coupled with the same attributes of AI, may introduce biases and errors.

International Humanitarian Law (IHL), also known as the Law Of Armed Conflict (LOAC), places limits on armed conflict and it protects people who are not or no longer participants in conflicts. [57] The principles of proportionality, distinction, and precaution are applicable to the use of AI and LAWS/DSS. Although current AI technology is not able to meet the requirements of principles of proportionality, distinction, and precaution but may be able to with the ongoing development and research of AGI and ASI. IHL also requires that any potential new weapon, means, and method of warfare undergo legal review to determine if it is banned by IHL [58]; legal review would be require explain-ability and documentation. Compliance with IHL and LOAC will further strengthen the position that AI and LAWS is able to kill a human in a dignified manner. In 2008 the robotics researcher Ron Arkin determined that it was theoretically possible to create and "ethical governor" in AI that would prohibit it from conducting illegal and

---

[54] *Ibid,*. 13.

[55] *Ibid,*. 13.

[56] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* (W.W. Norton and Company. New York | London. 2018), 8.

[57] Sharkey, Noel. *Why Robots Should Not Be Delegated With The Decision to Kill.* (Connection Science, 29:2. 2017), 178.

[58] *Ibid,*. 180.

unethical acts and be programmed to comply with LOAC. [59]

<u>Ethics</u>

A decision making process and the ethical components in it can be summarized in four basic steps. The first step is to determine alternative choices available to the decision maker and the second steep is to then define relevant possible results to each alternative. [60] The third step further builds on the previous two by creating a probability estimate for each alternative choice and the last step is making a comparison to determine the value trade off for each choice; what is the value of the objective versus the costs. [61] The end result is a rational choice available to the decision maker and this process is one that AI can compute quickly. [62] Ethical considerations become most apparent at step four, determining a value trade off because it is as this point the variables of a human perspective come to into play. A human's judgement, which can be subjective, is used to make an assessment which may or may not result in the death of another human, depending on whether or not the trade off value is worthwhile.

Discussions on the use of AI and LAWS need to move past the technical realm and include legal and ethical considerations including international humanitarian law (IHL). The mutual reinforcement of ethical and technical needs will drive developments and guarantee that design and developments are based on ethical deliberations. [63] A key to this is that AI systems need to be explainable using non-technical terms that describe what the technology will do and

---

[59] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* (W.W. Norton and Company. New York | London. 2018), 448.

[60] Dewees, Brad, Umphres, Chris, and Tungjanuary, Maddy. *Machine Learning and Life and Death Decisions on the Battlefield.* (War on the Rocks. 11 January 2021).

[61]*Ibid.*

[62] *Ibid.*

[63] Daniels, Owen, and Williams, Brian. *Day Zero for Military AI.* (War on the Rocks. 28 January 2020).

its risks and it must also be documented via a paper trail to better ensure transparency. [64]

Consequentialist ethical debate, one that is can be summarized as follows: a ban would result in more harm and no ban would see that AI and LAWS is more precise and would be able to reduce casualties, both military and civilian. [65] A question still remains regarding empathy and if AI will be able to learn empathy, mercy and humanity. [66] A deontological ethical perspective, on that is rules based, can also be considered for AI, as some actions are evil no matter the outcome, they are *mal en se*. [67] Both perspectives have merits and can be used to justify as well as deny the use development and used of AI and LAWS.

Conclusion

The research, development and implementation of AI and LAWS will continue both amongst Canada's allies and our adversaries. The technical advances will continue at a rate that will see AI be able to learn human qualities and be able to understand and act on humanity and dignity towards other humans. Canada, and the CAF, must continue to explore and use AI and LAWS so as to remain a peer with our allies and to not loose our competitive advantage against our enemies. There will be a point in the future when the use of AI to kill humans will not be considered evil and the understanding and acceptance by military commanders needs to happen now in order to be prepared for the future use of AI and LAWS. AI and LAWS must be discussed rationally, in plain non- technical terms so that the public, policy makers and end users know what it is they are using and the lethal effects AI and LAWS will have on other humans. Additionally, AI and LAWS must comply with national and international laws, some of which may still have to be created and ethical considerations must also be debated prior to giving AI and LAWS the freedom to kill a human with autonomously.

[64] *Ibid.*

[65] Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* (W.W. Norton and Company. New York | London. 2018), 432.

[66] *Ibid,. 433.*

[67] *Ibid,. 452.*

BIBLIOGRAPHY

Daniels, Owen, and Williams, Brian. "Day Zero for Military AI." *War on the Rocks*. 28 January 2020.

Dewees, Brad, Umphres, Chris, and Tungjanuary, Maddy. "Machine Learning and Life and Death Decisions on the Battlefield." *War on the Rocks*. 11 January 2021.

Galliot, Jai and Scholz, Jason. "Artificial Intelligence in Weapons: The Moral Imperative for Minimally- Just Autonomy." *Air Force Journal of Indo-Pacific Affairs 1,* no 2. Winter 2018.

Geist, Edward and Andrew J. Lohn. *How Might Artificial Intelligence Affect the Risk of Nuclear War?* Santa Monica, CA: RAND Corporation, 2018.

Government of Canada. *Responsible used of Artificial Intelligence*. Last accessed 26 May 21. Responsible use of artificial intelligence (AI) - Canada.ca

Government of Canada. Directive on Artificial Decision Making. Last accessed 26 May 21. Directive on Automated Decision-Making- Canada.ca (tbs-sct.gc.ca)

Hao, Karen. "Establishing an AI Code of Ethics Will be Harder Than People Think." *MIT Technology Review*. 21 October 2018.

Horowitz, Michael. "Artificial Intelligence, International Competition and the Balance of Power." *Texas National Security Review 1*, no. 3. May 2018.

Hume, Kathryn and Taylor, Matthew E. *Why AI That Teaches Itself to Achieve a Goal Is the Next Big Thing.* Last accessed 30 May 2. April 21, 2021. https://hbr.org/2021/04/why-ai-that-teaches-itself-to-achieve-a-goal-is-the-next-big-thing

Jenkins, Cameron. *AI Innovators Take Pledge Against Autonomous Killer Weapons*. Last accessed 27 May 21. Google AI Chiefs, Elon Musk Sign Pledge Against Autonomous Weapons : NPR

Latonero, Mark. *Governing Artificial Intelligence: Upholding Human Rights and Dignity.* Data and Society. Annenberg School for Communication, University of Southern California. 10 October 2018.

Losito, Marc and Anderson, John. *The Department of Defense's Looming AI Winter.* War on The Rocks. 10 May 2021.

Pfaff, Anthony. *The Ethics of Acquiring Disruptive Technologies Artificial Intelligence, Autonomous Weapons, and Decision Support Systems*. Prism 8, no.3.

Ploumis, Micahil. *Mission Command and Philosophy for the 21st Century.* Comparative Strategy, 39:2. 2020.

Price, Mathew, Walker, Stephen, and Wiley, Will. *The Machine Beneath. Implications of Artificial Intelligence in Strategic Decision Making.* Prism 7, No 4. 8 November 2018.

Salles, Arleen, Evers, Kathinka & Farisco, Michele. *Anthropomorphism in AI.* AJOB Neuroscience, 11:2. 2020.

Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* W.W. Norton and Company. New York | London. 2018.

Scharre, Paul. *Highlighting Artificial Intelligence.* Strategic Studies Quarterly. 11, no 4. Winter 2017.

Sharkey, Noel. *Why Robots Should Not Be Delegated With The Decision to Kill.* Connection Science, 29:2. 2017.

Tensor Flow. Last accessed 2 June 2021. https://www.tensorflow.org/

United Nations Security Council. *Final Report of the Panel of Experts on Libya Established Pursuant to Security Council Resolution 1973 (2011).* 11 March 2021. Last accessed 2 June 2021. https://undocs.org/S/2021/229

Vergun, David. *Defense Innovation Board Recommend AI Ethical Guidelines.* Last accessed 26 May 2021. https://www.defense.gov/Explore/News/Article/Article/2006646/defense-innovation-board-recommends-ai-ethical-guidelines/

Zafar, Muhammad Bilal; Valera, Isabel; Gomez-Rodriguez, Manuel; Gummadi, Krishna P. *Fairness Constraints: A Flexible Approach for Fair Classification.* Journal of Machine Learning Research 20, 2019.