AUTOMATED DECISION RESEARCH

Artificial intelligence and automated decisions: shared challenges in the civil and military spheres

SEPTEMBER 2022
REPORT

CONTENTS

Key messages 3

Introduction 4

Section one: the range of challenges 6

Section two: bias and discrimination 8

Section three: explainability 10

Section four: accountability 12

Conclusion 15

Endnotes 16

WWW.AUTOMATEDRESEARCH.ORG
INFO@AUTOMATEDRESEARCH.ORG
@AUTOMATEDDR

AUTOMATED DECISION RESEARCH

Artificial intelligence and automated decisions: shared challenges in the civil and military spheres

This report provides an initial sketch of civil responses to AI and automated decision-making, while contextualising these responses in relation to autonomy in weapons systems.

KEY MESSAGES

- The extent to which the serious risks and challenges presented by the use of AI and automated decision-making technologies are recognised by many states and international bodies in the civil space should be taken as validation for parallel and related concerns in the military space.
- Given the nature and scale of harms at stake in automated processing in the context of military targeting, and the difficulties in applying civilian oversight mechanisms in the military space, the challenges associated with autonomous weapons systems are particularly acute.
- The development of human rights-based responses to AI and automated decision-making in the civil space should require states to attend to the rights of those affected in the military space. Recognising the rights of people who would be affected is a fundamental starting point for the generation of necessary rules.
- The importance of accountability and responsibility for the use of autonomous weapons systems has implications for how certain rules need to be drawn in order to avoid an erosion of accountability, and to ensure the protection and fulfilment of international humanitarian law rules and fundamental human rights.

INTRODUCTION

For more than nine years, autonomous weapons systems have been the subject of international discussion in various fora, including in the UN Human Rights Council, the UN General Assembly First Committee on Disarmament and International Security (hereafter UN First Committee), and the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS). In these discussions, states, United Nations agencies, international organisations, and non-governmental organisations have highlighted the various and serious ethical, moral, humanitarian and legal implications of artificial intelligence (AI) and autonomous weapons systems. Despite a majority of states supporting negotiation of a legal instrument, the Sixth Review Conference of the CCW in December 2021 failed to agree a mandate to work towards any form of regulation.[1]

This paper provides an initial sketch of responses to AI and automated decision-making in wider society, while contextualising these responses in relation to autonomy in weapons systems. In compiling the report, forty states were identified which had publicly released specific policy documents or other strategies on the development and use of artificial intelligence domestically.[2] The report assesses numerous relevant national AI strategies and positions, EU-level reports and regulations, international guidelines, and other documents, in order to draw out core themes and concerns regarding the adoption and use of AI and automated decision-making technologies in the civil space. Written submissions and contributions of states to the GGE on LAWS and statements to the UN First Committee were also examined, together with recordings of GGE and First Committee sessions.

In our analysis of national policies, states in the Global North were notably more represented than states in the Global South. This is indicative of existing disparities - political, economic, social, and technological - in the ability and capacity to adopt AI technologies, and to adequately respond to this major shift in governance and automation. The UN News Centre, for example, notes that 'there is strong evidence to suggest that AI is playing a role in making the world more unequal, and is benefiting a small proportion of people.'[3] As UNESCO's Recommendation on the ethics of artificial intelligence recognises, 'AI technologies can deepen existing divides and inequalities in the world, within and between countries.'[4] The Recommendation further recognises that 'justice, trust and fairness must be upheld so that no country can be left behind, either by having fair access to AI technologies and enjoying their benefits or in the protection against their negative implications.'[5] More concerted efforts are required to share technological advances freely and recognise contributions from the Global South, while building awareness of the attendant risks and challenges.

The paper highlights four primary areas related to AI and automated decision-making which have drawn significant attention from states and various international bodies and which might be considered particularly relevant to the consideration of autonomy in weapons systems. We have organised our analysis around these four major themes, which are listed below:

1. The use of AI and algorithmic decision-making systems pose significant risks and challenges, particularly in the areas of peace and security; transparency and explainability; responsibility and accountability; and bias and discrimination.

2. Automated & algorithmic systems engender particular concerns regarding bias, discrimination, inequality and human dignity.

3. AI and automated decision-making systems must be transparent, explainable, and understandable.

4. Accountability and responsibility for automated decisions is essential, and people must have recourse to justice and the right to remedy.

AI and automated decision-making technologies are permeating society in multiple areas, even though serious ambiguity still exists regarding the processes at play, and despite widespread recognition of clear challenges regarding: the trustworthiness, predictability and reliability of these technologies; issues around transparency and explainability; responsibility and accountability for harmful decisions and the right to remedy; and in the critical effects on digital dehumanisation and the right to human dignity. Further, concerns have been raised by legal institutions, civil society organisations, and others on the role of such technologies in amplifying bias and discrimination. These concerns gain further prominence in the area of autonomous weapons systems, as the opaque nature of algorithmic processing and targeting pose unprecedented challenges to the meaningful implementation of human rights and respect for international humanitarian law. In the military terrain, where the risks from such technologies are arguably significantly higher, and avenues for oversight, accountability, responsibility, and remedy more complicated, the challenges seem particularly acute.

SECTION ONE: THE RANGE OF CHALLENGES

The use of AI and algorithmic decision-making systems pose significant risks and challenges, particularly in the areas of peace and security; transparency and explainability; responsibility and accountability; and bias and discrimination.

The serious risks and challenges presented by AI and autonomous weapons systems are well-established, having been the subject of discussion at UN-level for nine years. Many of these and other related concerns are also widely expressed in general civil responses to AI and automated decision-making.

During discussions of the CCW GGE on LAWS, the UN First Committee, and in other international fora, numerous states have raised specific concerns regarding the significant ethical, legal, humanitarian, and moral implications of artificial intelligence and autonomous weapons systems. These risks include, but are not limited to: the risk of digital dehumanisation[6] and the perpetuation of biases and discrimination in multiple and intersecting areas, including racial and gender discrimination and stereotyping; unpredictability and unreliability of autonomous weapons systems; the lack of effective accountability and responsibility for decisions made by automated systems, including those which implicate the targeting of people and the use of force; automation bias, and a lack of human judgment and understanding in decision-making; the undermining of state compliance with international humanitarian law, international human rights law and international criminal law; the lowering of the threshold for the use of force; force multiplication and arms races; the risk of asymmetric warfare, and conflict escalation and reprisal; and serious challenges to international and regional peace, security, and stability.[7]

In the non-military sphere, the risks and challenges accompanying the use of artificial intelligence and algorithmic systems are widely recognised. UNESCO's Recommendation on the ethics of artificial intelligence notes that AI technologies 'raise fundamental ethical concerns.'[8] In September 2021, a report from the UN High Commissioner for Human Rights, Michelle Bachelet, called for a moratorium 'on the sale and use of AI systems that carry a high risk for the enjoyment of human rights, unless and until adequate safeguards to protect human rights are in place',[9] with the Commissioner observing that AI technologies can 'have negative, even catastrophic, effects if deployed without sufficient regard to their impact on human rights'.[10]

Similarly, recognising that associated risks and harms exist across various areas and that such risks must be adequately addressed is a central component of many states' AI strategies and policies:

- UNESCO's Recommendation on the ethics of AI observes that AI has a potential impact on, including but not limited to, 'human dignity, human rights and fundamental freedoms, gender equality, democracy, social, economic, political and cultural processes, scientific and engineering practises, animal welfare, and the environment and ecosystems'.[11]
- The African Commission on Human and People's Rights, in its 2021 Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa, recognises that 'new and emerging technologies such as Artificial Intelligence (AI), robotics and other new emerging technologies present both opportunities and perils for the promotion and protection of human and people's rights in Africa', and notes 'the challenges and concerns that are posed by autonomous systems that are not under meaningful human control.'[12]

- The EU's Independent High-Level Expert Group on AI's Ethics Guidelines for Trustworthy AI notes that AI systems 'pose certain risks and may have a negative impact, including impacts which may be difficult to anticipate, identify or measure,' and states that 'AI systems need to be human-centric, resting on a commitment to their use in the service of humanity and the common good, with the goal of improving human welfare and freedom.'[13]
- The OECD Recommendation of the Council on Artificial Intelligence recognises that AI 'has pervasive, far-reaching and global implications that are transforming societies, economic sectors and the world of work, and are likely to increasingly do so in the future', and that 'these transformations may have disparate effects within, and between societies and economies, notably regarding economic shifts, competition, transitions in the labour market, inequalities, and implications for democracy and human rights, privacy and data protection, and digital security.'[14]
- Germany's Artificial Intelligence Strategy of the German Federal Government notes that in order to effectively counteract risks, 'specific requirements need to be set for the development and use of AI systems and the framework for responsible development and use of these systems for the benefit of the common good needs to be aligned accordingly.'[15]
- Norway's National Strategy for Artificial Intelligence notes that 'Developing and using artificial intelligence can create challenges and raise many complex questions. This particularly applies to AI that builds on personal data.'[16]
- The final report of Finland's Artificial Intelligence Programme notes that 'the deployment of AI has raised issues and concerns relating to human rights and the realisation of democracy'.[17]
- Sweden's National Approach to Artificial Intelligence notes that 'Sweden needs to continue to develop efforts to prevent and manage the risks associated with AI', and states that 'the risks associated with AI are not only technical but also ethical'.[18]
- China's Ethical Norms for New Generation Artificial Intelligence states that activities in AI should 'abide by the common values of humankind, respect human rights and the fundamental interests of humankind', and 'effectively protect the legitimate rights and interests of all relevant stakeholders'.[19] Meanwhile, China's New Generation Artificial Intelligence Development Plan notes that 'AI is a disruptive technology with widespread influence' that will have 'far-reaching effects on the management of government, economic security, and social stability, as well as global governance.' As such, 'we must attach great importance to the potential safety risks and challenges'.[20]
- The UK's National AI Strategy states that 'we have a responsibility to not only look at the extreme risks that could be made real with AGI (Artificial General Intelligence), but also to consider the dual-use threats we are already faced with today', and further notes that 'there is growing awareness in industry and by citizens of the potential risks and harms associated with AI technologies.' [21]

As is recognised by the majority of states, such risks are also present in the military space. As mentioned previously, these risks have been noted in the context of international discussions on autonomous weapons systems. For example, in a joint statement to the Sixth Review Conference of the CCW in December 2021, a group of thirteen states exhorted the international community 'to move expeditiously to address humanitarian, military, legal, ethical, moral and other concerns over the far-reaching implications of the development and use of these weapons.' [22] Similarly, the African Group acknowledged that autonomous weapons systems 'raise ethical, legal, moral and technical questions, as well as international peace and security-related questions which should be thoroughly deliberated and examined.'[23] Austria, in its statement to the 2021 UN First Committee, noted that advances in technology and AI extend to armed conflict with 'potentially unacceptable consequences'.[24] In the same forum, Pakistan observed that 'the development and potential deployment of autonomous weapons carries serious implications for international humanitarian law as well as regional and international peace and security,'[25] with new technologies affording 'new means of war and therefore heightened risks and threats.'[26]

The EU's written contribution to the GGE on LAWS in June 2021 noted that 'a normative and operational framework should recognise that risks can emerge during the development, testing and deployment of weapons systems based on emerging technologies in the area of LAWS and recommend risk mitigation measures.'[27]

The extent to which the serious risks and challenges of AI are recognised by many states and international bodies in the civil space should be taken as validation for parallel and related concerns in the military space.

SECTION TWO: BIAS AND DISCRIMINATION

Automated & algorithmic systems engender particular concerns regarding bias, discrimination, inequality and human dignity.

Of particular importance in discussions on risks and harms in the context of autonomous weapons systems are systems specifically designed or used to target humans. As a group of ten states observed in a joint contribution to the GGE on LAWS, 'these systems boil life-and-death determinations down to data points, thus dehumanising people and violating their right to dignity. They are also vulnerable to exacerbating social injustices due to algorithmic bias.'[28] Purporting to be able to distinguish between combatants and civilians, between active combatants and those hors de combat, or between civilians, civilian persons with disabilities, and civilians directly participating in hostilities on the basis of data acquired by sensors and processed and classified by algorithms raises serious legal, ethical and moral concerns. For example, this has been highlighted by the UN Special Rapporteur on the rights of persons with disabilities, Gerard Quinn, in his 2021 report, which states that 'the deployment and use of fully autonomous weapons systems, like other artificial intelligence systems, raises concerns as to the ability of weaponry directed by artificial intelligence to discriminate between combatants and non-combatants, and make the nuanced determination as to whether an assistive device qualifies a person with disabilities as a threat.'[29] As such, and as Brazil, Chile, and Mexico have observed, 'precisely because of their impact on the right to life and human dignity, the use of force, increasingly mediated through technology, must consider ethical considerations and their societal implications as the main parameters on which to confront these challenges.'[30]

In a non-military context, the possibility of digital dehumanisation, and of the further embedding of bias, discrimination and inequality through automated decision-making, for instance in relation to race and gender, among other factors, is well-known and widely acknowledged. The EU's Independent High-Level Expert Group on AI's Ethics Guidelines for Trustworthy AI, for example, states that 'Equal respect for the moral worth and dignity of all human beings must be ensured. This goes beyond non-discrimination, which tolerates the drawing of distinctions between dissimilar situations based on objective justifications', and notes that 'respect for human dignity entails that all people are treated with respect due to them as moral subjects, rather than merely as objects to be sifted, sorted, scored, herded, conditioned or manipulated.'[31] States also acknowledge the importance of respect for human dignity, non-discrimination, and equality, and of a 'humane' and 'human-centred' AI:

- The OECD principles for responsible stewardship of trustworthy AI stipulates that 'AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle. These include freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognised labour rights.'[32]

- The African Commission's 2021 Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa calls on State Parties 'to ensure that the development and use of AI, robotics and other new and emerging technologies is compatible with the rights and duties in the African Charter and other regional and international human rights instruments, in order to uphold human dignity, privacy, equality, non-discrimination, inclusion, diversity, safety, fairness, transparency, accountability and economic development as underlying principles that guide the development and use of AI, robotics and other new and emerging technologies.'
- Chile's National Artificial Intelligence Policy emphasises that 'AI should not discriminate based on protected categories or be used to the detriment of any group. In particular, it is especially important that AI be developed with a gender and sexual diversity perspective, including historically relegated groups such as indigenous peoples, people with special abilities, or the most vulnerable sectors of our economy'.[33]
- Mexico's Artificial Intelligence Agenda in Brief states that 'The problems associated with biases and errors generated by autonomous systems are the Achilles heel of AI.'[34]
- Ireland's National Artificial Intelligence Strategy states that 'there is a risk that AI systems could lead to unfair discrimination and unequal treatment. The risk of discrimination can arise in many ways, for instance biased training data, biased design of algorithms, or biased use of AI systems.' It also states that 'AI-based systems have the potential to exacerbate existing structural inequities and marginalisation of vulnerable groups.'[35]
- Germany's Artificial Intelligence Strategy of the German Federal Government notes that 'when using AI, effective protection against discrimination, manipulation or any other misuse needs to be ensured.'[36]
- France's For a Meaningful Artificial Intelligence: Towards a French and European Strategy report notes that 'AI can reproduce bias and discrimination and is becoming increasingly present in our social and economic environments, so opening the black box is a key democratic issue.'[37]
- The Netherlands' Strategic Action Plan for Artificial Intelligence notes that 'low quality data can cause bias and imbalances in AI: bias when the data itself contains prejudices (because people have them too), bias through prejudice in the algorithms, and imbalances when data is not representative.' It also notes the risks of dehumanisation and the influence of AI on making choices.[38]
- Serbia's Strategy for the Development of Artificial Intelligence notes that 'the automation of decision-making using artificial intelligence, or of the analyses that influence decision-making, involves the risk of using criteria which are discriminatory by nature.'[39]
- Singapore's Model Artificial Intelligence Governance Framework includes a number of guiding principles for the use of AI, including that 'AI solutions should be human-centric. As AI is used to amplify human capabilities, the protection of the interests of human beings, including their well-being and safety, should be the primary consideration in the design, development and deployment of AI.'[40]
- The European Commission's proposed AI Act notes that 'the use of AI with its specific characteristics (e.g. opacity, complexity, dependency on data, autonomous behaviour) can adversely affect a number of fundamental rights enshrined in the EU Charter of Fundamental Rights' including the right to human dignity, non-discrimination, and equality between women and men.[41]

All of the above concerns find a corollary in the context of autonomous weapons systems and, as the UK has acknowledged, 'the normative landscape, encompassing obligations and principles, ethical and legal concerns is of central importance to the LAWS debate'.[42] Concerns in this area are frequently articulated by states. In terms of the further embedding of existing structural inequalities, for example, Palestine has drawn attention to the detrimental impact autonomous weapons

systems would have on the Global South, underlining that 'the populations of the Global South will bear the brunt of new weapons' as these populations have historically 'been the most affected and harmed by the emergence of new technologies and weapons systems'.[43] Canada has recognised, from discussions with 'Indigenous and civil society partners', concerns that autonomous systems could 'exacerbate existing power imbalances and biases.'[44] At the Sixth Review Conference of the CCW in December 2021, Sierra Leone called for a ban on sensor-based weapons systems which reduce humans to data points 'to be sorted and processed by algorithms', using 'proxy indicators such as weight, heat, movements, and specific biometric indicators such as skin colour as a basis for encoding patterns of sensor data as a representation of humans.'[45] A joint commentary by a group of nine states notes that the challenges of cognitive limitations of the system, epistemological limitations, and algorithmic bias need to be considered when 'designing, deploying and using' weapons systems based on emerging technologies, in order for such systems to be operated in conformity with IHL.[46]

Given the nature and scale of harms at stake in automated processing in the context of military targeting, and the difficulties in applying civilian oversight mechanisms in the military space, the challenges associated with autonomous weapons systems are particularly acute.

SECTION THREE: EXPLAINABILITY

AI and automated decision-making systems must be transparent, explainable, and understandable.

In the international discussions on autonomous weapons systems, transparency, explainability and understandability have been recognised as key considerations. These issues are well articulated in the civil sector, where they are recognised as essential for the ethical and legal use of automated decision-making technologies and in maintaining an appropriate link between those who employ such systems, and those impacted by them. As such, transparency, explainability and understandability are seen as important regarding rights to remedy for people that may have been exposed to harms.

The need for identifiable, transparent, explainable, and understandable AI and automated decision-making and processing is invoked in numerous national AI strategies, and other regulations, principles, documents and guidelines:

- UNESCO's Recommendation on the Ethics of AI describes the transparency and explainability of AI systems as 'often essential preconditions to ensure the respect, protection and promotion of human rights, fundamental freedoms and ethical principles'. The Recommendation notes that a 'lack of transparency could also undermine the possibility of effectively challenging decisions based on outcomes produced by AI systems and may thereby infringe the right to a fair trial and effective remedy', and states that 'in the case of AI applications that impact the end user in a way that is not temporary, easily reversible or otherwise low risk, it should be ensured that the meaningful explanation is provided with any decision that resulted in the action taken in order for the outcome to be considered transparent.' The Recommendation also expressly links transparency and explainability to adequate responsibility and accountability measures.[47]
- The OECD Recommendation of the Council on Artificial Intelligence states that 'AI Actors should commit to transparency and responsible disclosure regarding AI systems' and should provide 'meaningful information, appropriate to the context, and consistent with the state of art' to 'foster a general understanding of AI systems; to make stakeholders aware of their interactions with AI systems, including in the workplace; to enable those affected by an AI system to understand the outcome; and, to enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information on the factors, and the logic that served as the basis for the prediction, recommendation or decision.'[48]

- Brazil's Artificial Intelligence Strategy highlights that 'for an algorithm to be "explainable" or "interpretable", it is desirable that the steps of the machine learning process that resulted in an inference are traceable and that the variables that weighed in the decision making can pass scrutiny', and further notes that 'transparency is also an important element of AI governance structures, either with regard to information regarding the interaction with AI systems (disclosure), or with regard to the idea of explainability of decisions made by autonomous systems.'[49]

- Germany's Artificial Intelligence Strategy of the German Federal Government argues that 'it is hugely important to continue to apply high standards for the introduction and use of AI and to ensure that AI is non-discriminatory and transparent, complies with existing rules governing the protection of personal data, information security requirements and upholds citizens' trust and confidence. Traceability and verifiability of decisions, as well as transparency, fairness and non-discrimination, safety and security and participation are key to creating trust in the use of AI in public administration.'[50]

- France's For a Meaningful Artificial Intelligence report states that 'there needs to be greater transparency and auditability concerning autonomous systems', and that a meaningful AI 'implies that AI should be explainable'.[51]

- India's Artificial Intelligence Task Force report notes that 'AI systems must be transparent, i.e. they must be known to humans as machines and their performance, including their learning, must be verifiable/auditable.'[52]

- Mexico's Artificial Intelligence Agenda states that non-technical standards for AI systems should 'include but not be limited to transparency and easily explainable AI systems'.[53]

- Ireland's National Artificial Intelligence Strategy states that 'transparency in the use of AI systems is critical for building public trust. The opaque nature of many AI algorithms may also obscure the reasoning behind AI-based decisions and can cause problems from the perspective of explainability and accountability. It is therefore important to ensure that information about how AI systems make consequential decisions is public and understandable.'[54]

- The Netherlands' Strategic Action Plan for Artificial Intelligence argues that AI should be developed in a way that is 'human-centric, trustworthy, transparent and explainable'. The strategic action plan further notes, that, where the use of profiling is accompanied by automated decision-making, 'the data subject should be made aware of that and useful information should be provided on the underlying logic, the importance and the expected consequences of such processing for the data subject.'[55]

- Luxembourg's AI strategy states that AI 'needs to be understandable, transparent and, ultimately, trustworthy.'[56]

- The EU's General Data Protection Regulation (GDPR) is relevant to this discussion, as it applies to the 'processing of data wholly or partly by automated means'. The GDPR stipulates in Article 5, in regards to the processing of personal data, that 'personal data should be processed lawfully, fairly, and in a transparent manner in relation to the data subject'; Article 12 sets out further rules regarding transparent information, communication and modalities for the exercise of the data subjects rights, stipulating that any information or communication under articles 13 and 14, 15-22 and 34 of the Regulation should be provided to the data subject in a 'concise, transparent, intelligible and easily accessible form, using clear and plain language, in particular for any information addressed specifically to a child.' Under Articles 14 and 15, data subjects should be informed of 'the existence of automated decision-making, including profiling', and at least in the cases where Articles 22.1 and 22.4 are implicated, 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.'[57]

- The G20's Guiding Principles on AI states that 'AI actors should commit to transparency and responsible disclosure regarding AI systems', including the provision of meaningful information 'to enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information on the factors, and the logic that served as the basis for the prediction, recommendation or decision.'[58]

- The EU's Independent High-Level Expert Group on AI's Ethics Guidelines for Trustworthy AI report states that 'processes need to be transparent, the capabilities and purpose of AI systems openly communicated, and decisions – to the extent possible – explainable to those directly and indirectly affected. Without such information, a decision cannot be duly contested.'[59]

In the military sphere, these issues are particularly salient given the importance of responsibility and accountability for ensuring human dignity and respect for international humanitarian law and international human rights law. For example, the United Kingdom has noted that human responsibility for autonomous systems rests on an appropriate level of understanding of these systems by military personnel.[60] France and Germany have argued that 'developers, commanders and operators must have a sufficient understanding' of how a weapons system operates, which enables 'the commanders and operators to predict (prospective focus) and explain (retrospective) the behavior of weapons systems.' [61] China has noted that states should commit to not developing or using autonomous weapons systems whose effects can't be sufficiently understood, predicted or explained.[62] At the Sixth Review Conference of the CCW, Austria warned that 'We are at the crossroads of crossing red lines of humans being killed by algorithms and systems that humans potentially do not even understand retrospectively.' [63]

However, in contrast to conversations in the civil sector, discussions regarding the transparency, explainability and understandability of autonomous weapons systems have tended to focus on these issues more around the interest of those deploying and operating such weapons, rather than in support of the rights of those who may be targeted or otherwise harmed and affected by them. The development of human rights-focussed responses to AI and automated decision-making in the civil space should impel states to attend to the rights of those affected as a fundamental starting point for the generation of necessary rules in the military space.

SECTION FOUR: ACCOUNTABILITY

Accountability and responsibility for automated decisions is essential, and people must have recourse to justice and the right to remedy.

Autonomous weapons systems significantly implicate accountability and responsibility for violations of international humanitarian law, international human rights law, and international criminal law, and the right to remedy and reparation. These issues have been widely recognised in UN-level discussions.

In the civil sphere, the challenges presented to accountability, responsibility, remedy and reparation by AI and autonomous decision-making systems are accepted as an important concern. The need for regulation, for effective and responsive oversight mechanisms, and for accountability in the use of AI and automated decision-making systems is acknowledged as crucial, as is the right of those harmed by automated decision-making to have recourse to justice and remedy for harms:

- UNESCO's Recommendation on the ethics of AI states that 'Member States should ensure that it is always possible to attribute ethical and legal responsibility for any stage of the life cycle of AI systems, as well as in cases of remedy related to AI systems, to physical persons or to existing legal entities. Human oversight refers thus not only to individual human oversight, but to inclusive public oversight, as appropriate.' [64]
- The European Parliament's Committee on Legal Affairs Report on artificial intelligence: questions of interpretation and application of international law in so far as the EU is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice reiterates that 'autonomous decision-making should not absolve humans from responsibility, and that people must always have ultimate responsibility for decision-making processes

so that the human responsible for the decision can be identified',[65] while the Committee on Civil Liberties, Justice and Home Affairs Report on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters states that all AI solutions for law enforcement and the judiciary need to fully respect the principles of human dignity, and the right to an effective remedy, and stresses that 'the use of AI applications must be prohibited when incompatible with fundamental rights.'[66]

- France's For A Meaningful Artificial Intelligence report argues that it is vital that 'responsibility can be attributed to a human being via a predetermined procedure', and states that it is 'essential that legislation and ethics control the performance of AI systems.'[67]

- The G20's Guiding Principles on AI notes that 'AI actors should ensure traceability, including in relation to data-sets, processes and decisions made during the AI system lifecycle, to enable analysis of the AI system's outcomes and responses to inquiry, appropriate to the context and consistent with the state of the art,' and further notes that 'AI actors should be accountable for the proper functioning of AI systems and for the respect of the above principles, based on their roles, the context, and consistent with the state of the art'.[68]

- Portugal's AI strategy, AI Portugal 2030, specifically notes that in the area of automated decision-making, 'AI will completely change the paradigm of human-machine interface and the decision processes.' It further states that 'AI systems will make important and critical decisions autonomously. Society will demand transparency and auditability', and notes that 'the legal framework will have to be adjusted to determine liability in conflicts with the involvement of AI decision-making.'[69]

- The Netherlands' Strategic Action Plan for Artificial Intelligence notes that 'AI must be developed and applied within appropriate ethical and legal frameworks', and that it must then 'be possible to verify how parties have applied AI, especially when it comes to high-impact applications.'[70]

- Australia's Artificial Intelligence Ethics Framework states that 'AI systems that have a significant impact on an individual's rights should be accountable to external review' , and that 'when an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.'[71]

- Norway's National Strategy for Artificial Intelligence states that 'individuals or legal persons must have an opportunity to gain insight into how a decision that affects them was made', and stresses that 'the requirement of accountability complements the other requirements, and entails the introduction of mechanisms to ensure accountability for solutions built on AI and for their outcomes, both before and after the solutions are implemented. All AI systems must be auditable'. Further, the Strategy states that 'if unjust adverse impacts occur in a solution built on AI, mechanisms should be in place to ensure that such impacts can be reported.'[72]

- The EU's GDPR sets out numerous rights and obligations related to the automated processing of personal data, and the processing of personal data in general. These include, for example, the right to rectification of inaccurate personal data (Article 16); the right to restriction of processing where the accuracy of the personal data is contested by the data subject (Article 18); the right to object to the processing of personal data (in defined circumstances which are further described in the relevant article; Article 21); and, under Article 22, the 'right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.' Further, Chapter VIII of the Regulation sets out remedies, liability and penalties, and stipulates that data subjects have: the right to lodge a complaint with a supervisory authority (Article 77); the right to an effective judicial remedy against a supervisory authority (Article 78); the right to an effective judicial remedy against a controller or processor (Article 79); and, under Article 82, any person who has suffered material or non-material damage as a result of an infringement of the Regulation has the right to 'receive compensation from the controller or processor for the damage suffered.'[73]

Corresponding concerns exist in the military space. Numerous states have made clear the importance of accountability and responsibility. Switzerland has observed that states and humans must not escape international responsibility by a process of 'delegating' certain tasks to autonomous weapons systems, stressing that 'states and humans remain responsible for violations of IHL.'[74] The Non-Aligned Movement (NAM), during the Sixth Review Conference of the CCW, emphasised that there is an urgent need to pursue a legally-binding instrument which takes into account 'human responsibility as well as accountability for developing, deploying, and using any emerging weapons systems in the framework of the CCW and in accordance with applicable international law.'[75] France and Germany have stated that 'maintaining human responsibility and accountability at all times in all circumstances and across the entire lifecycle of the algorithm' is key.[76] The United Kingdom has stated that 'human accountability for military outcomes remains fundamental to effective and responsible-decision-making', and that the use of autonomous weapons systems 'does not, and cannot, negate state responsibility and individual accountability under IHL.'[77]

However, states have also pointed to the difficulties in ensuring effective responsibility and accountability in the use of autonomous weapons systems, and have drawn parallels with already existing issues of accountability and responsibility in the use of armed drones.[78] In the context of armed drone use, the former UN Special Rapporteur on extrajudicial, summary or arbitrary executions, Agnés Callamard, said that drone operations are 'characterized by violations of the international obligation to investigate and to punish, where applicable, those responsible for violations of international humanitarian or human rights law. There is little public disclosure, as targeted killings by drones are wrapped in secrecy extending to the investigation of civilian deaths.'[79] Furthermore, the late Christof Heyns, also a former UN Special Rapporteur on extrajudicial, summary, or arbitrary executions, noted in relation to autonomous weapons systems that 'a lack of accountability for violations of the right to life is in itself a violation of that right, and if AWS create an accountability vacuum where mistakes occur, then the right to life is infringed.'[80]

It is noticeable that many of the orientations to accountability and responsibility in the civil space invoke or assert the rights of those affected by automated decision-making, including profiling and processing, and the potential for public bodies to apply scrutiny and oversight – an approach that is generally problematic in the military context. While there is a rhetorical recognition of the importance of responsibility and accountability, it is not clear how this acknowledgement can be squared with the harms likely to arise from the use of autonomous weapons systems. This has implications for how certain rules need to be drawn in order to avoid an erosion of accountability and ensure the protection and fulfilment of international humanitarian law rules and fundamental human rights.`

CONCLUSION

This paper aimed to draw out some of the primary concerns around AI and automated decision-making in the civil domain by examining states' national AI strategies and policies, existing and proposed regulations and reports at the EU level, and international ethical guidelines and recommendations on AI. As is noted in the introduction to this report, the disparity in national strategy and policy outputs on AI between the Global North and the Global South is representative of wider disparities in the access to and adoption of such technologies globally.

The paper highlighted four main areas of concern, all of which find a corollary in the military sphere (as evidenced by discussions on autonomous weapons systems in various international fora): that AI and automated decision-making present serious risks and challenges which must be addressed; that these technologies present particular risks for human dignity, inequality, digital dehumanisation, bias, and discrimination; that there is a need for systems to be transparent, explainable, and understandable; and that accountability and responsibility for the decisions of automated and autonomous systems must always be attributable to humans, with a right to remedy for those affected. We are not arguing here that these themes are exhaustive, but that they appear to be central areas of concern in general responses to AI and automated decision-making. Nor are we arguing that existing and proposed regulation of AI and automated decision-making in the civil sphere are sufficient to adequately address these concerns.

The breadth of responses to AI and automated decision-making systems in the civil sphere reinforces the necessity for responses in the military space. As is shown in this paper, there is widespread recognition that AI and automated decision-making systems present numerous risks and challenges – ethically, legally, and morally - and that specific responses to these risks and challenges are required. In the military space, the risks and harms arising from such systems are severe, involving possible death and injury to individuals, and the deprivation of fundamental human rights. At the same time, many of the mechanisms to mitigate risks or harms in the civil space are less readily applicable in the military space, in so far as they draw upon, for example, themes of transparency, explainability, stakeholder engagement, third-party oversight, etc. – all of which are traditionally problematic in the military context.

AUTOMATED DECISION RESEARCH

ABOUT AUTOMATED DECISION RESEARCH:

Automated Decision Research is the monitoring and research team of Stop Killer Robots - tracking state support for a legally binding instrument on autonomous weapons systems and conducting research and analysis on responses to autonomy and automated decision-making in warfare and wider society.

ENDNOTES:

1. See, for example, Stop Killer Robots (2021), 'Historic opportunity to regulate killer robots fails as a handful of states block the majority'. Available at: https://www.stopkillerrobots.org/news/historic-opportunity-to-regulate-killer-robots-fails-as-a-handful-of-states-block-the-majority/

2. Relevant state AI strategies and other documents on AI were identified through the OECD AI Policy Observatory, the Artificial Intelligence Index Report 2021, and government websites. OECD AI Policy Observatory available at: https://oecd.ai/en/; Artificial Intelligence Index Report 2021 available at: https://hai.stanford.edu/research/ai-index-2021

3. UN News (2020), 'Bias, racism and lies: facing up to the unwanted consequences of AI'. Available at: https://news.un.org/en/story/2020/12/1080192

4. UNESCO Recommendation on the ethics of artificial intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation

5. UNESCO Recommendation on the ethics of artificial intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation

6. See, for example, Stop Killer Robots (2021), 'Digital dehumanisation – when machines decide, not people'. Available at: https://www.stopkillerrobots.org/news/digital-dehumanisation-when-machines-decide-not-people/

7. See, for example, discussions of states parties to the Convention on Certain Conventional Weapons during the meetings of the Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Recordings are available via UN Web TV

8. UNESCO Recommendation on the ethics of artificial intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation

9. UNHCR (2021), 'The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights', A/HRC/48/31. Available at: https://www.ohchr.org/EN/Issues/DigitalAge/Pages/cfi-digital-age.aspx

10. UN News (2021), 'Urgent action needed over artificial intelligence risks to human rights'. Available at: https://news.un.org/en/story/2021/09/1099972

11. UNESCO Recommendation on the ethics of artificial intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation

12. African Commission on Human and Peoples' Rights, Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa - ACHPR/Res. 473 (EXT.OS/ XXXI) 2021. Available at: https://www.achpr.org/sessions/resolutions?id=504

13. European Commission Independent High-Level Expert Group on Artificial Intelligence (2019), Ethics Guidelines for a Trustworthy AI. Available at: https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1

14. OECD Recommendation of the Council on Artificial Intelligence. Available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

15. Artificial Intelligence Strategy of the German Federal Government. Available at: https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf

16. Norwegian National Strategy for Artificial Intelligence. Available at: https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd-353b6ae8/en-gb/pdfs/ki-strategi_en.pdf

17. Leading the way into the age of artificial intelligence: Final report of Finland's Artificial Intelligence Programme 2019. Available at: https://tem.fi/en/publication?pubid=URN:ISBN:978-952-327-437-2

18. Swedish National approach to artificial intelligence. Available at: https://www.government.se/4a7451/contentassets/fe2ba005fb49433587574c513a-837fac/national-approach-to-artificial-intelligence.pdf

19. International Research Center for AI Ethics and Governance, The Ethical Norms for the New Generation Artificial Intelligence, China. Available at: https://ai-ethics-and-governance.institute/2021/09/27/the-ethical-norms-for-the-new-generation-artificial-intelligence-china/

20. New America Foundation Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017). Available at: https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/

21. United Kingdom National AI Strategy. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf

22. Statement by Argentina, Costa Rica, Ecuador, El Salvador, Guatemala, Palestine, Philippines, Kazakhstan, Nigeria, Panama, Peru, Sierra Leone, and Uruguay, Sixth Review Conference of the Convention on Certain Conventional Weapons. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/RevCon/statements/17Dec_G13.pdf

23. Statement by theAfrican Group, Sixth Review Conference of the Convention on Certain Conventional Weapons. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/RevCon/statements/14Dec_AfricanGroup.pdf

24. Statement by Austria, UN General Assembly First Committee, 06 October 2021. Available at: https://www.reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com21/statements/6Oct_Austria.pdf

25. Statement by Pakistan, UN General Assembly First Committee, 13 October 2021. Available at: https://media.un.org/en/asset/k12/k12houm5np

26. Statement by Pakistan, UN General Assembly First Committee, 18 October 2021. Available at: https://media.un.org/en/asset/k1c/k1cagsmnip

27. Written contribution of the EU to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/EU.pdf

28. Written contribution of Argentina, Costa Rica, Ecuador, El Salvador, Panama, Palestine, Peru, Philippines, Sierra Leone, and Uruguay to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/G10_sept.pdf

29. Report of the Special Rapporteur on the rights of persons with disabilities (2021), Artificial intelligence and the rights of persons with disabilities, A/HRC/49/52. Available at: https://www.ohchr.org/en/documents/thematic-reports/ahrc4952-artificial-intelligence-and-rights-persons-disabilities-report

30. Written contribution of Brazil, Chile and Mexico to the CCW Group of Governmental Experts on LAWS. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/Brazil-Chile-Mexico.pdf

31. European Commission Independent High-Level Expert Group on Artificial Intelligence (2019), Ethics Guidelines for a Trustworthy AI. Available at: https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1

32. OECD Recommendation of the Council on Artificial Intelligence. Available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

33. National Artificial Intelligence Policy of Chile. Available at: https://minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento_politica_ia_digital_.pdf

34. The Agenda in Brief - Artificial Intelligence in Mexico: A National Agenda. Available at: https://36dc704c-0d61-4da0-87fa-917581cbce16.filesusr.com/ugd/7be025_85f5cec6ea584d8a842d11ad401c0685.pdf

35. AI - Here for Good: National Artificial Intelligence Strategy for Ireland. Available at: https://www.gov.ie/en/publication/91f74-national-ai-strategy/

36. Artificial Intelligence Strategy of the German Federal Government. Available at: https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf

37. For a Meaningful Artificial Intelligence: Towards a French and European Strategy. Available at: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf

38. Dutch Strategic Action Plan for Artificial Intelligence. Available at: https://www.rijksoverheid.nl/ministeries/ministerie-van-economische-zaken-en-klimaat/documenten/beleidsnotas/2019/10/08/strategisch-actieplan-voor-artificiele-intelligentie

39. Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025. Available at: https://www.media.srbija.gov.rs/medsrp/dokumenti/strategy_artificial_intelligence.pdf.

40. Model Artificial intelligence Governance Framework, Second Edition. Available at: https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf

41. European Commission Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF

42. Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions (2020), Use of armed drones for targeted killings, A/HRC/44/38. Available at: https://undocs.org/en/A/HRC/44/38

43. Statement by Palestine at the Sixth Review Conference of the Convention on Certain Conventional Weapons, 16 December 2021. Transcript available at: http://149.202.215.129:8080/s2t/UNOG/RCHCP6-16-12-2021-PM_mp3_en.html

44. Written contribution of Canada to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/Canada-2.pdf

45. Statement by Sierra Leone at the Sixth Review Conference of the Convention on Certain Conventional Weapons, 14 December 2021. Transcript available at: http://149.202.215.129:8080/s2t/UNOG/RCHCP6-14-12-2021-AM_mp3_en.html

46. Joint commentary submitted by Austria, Belgium, Brazil, Chile, Ireland, Germany, Luxembourg, Mexico, and New Zealand to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2020/gge/documents/Various-states_2020.pdf

47. UNESCO Recommendation on the ethics of artificial intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation

48. OECD Recommendation of the Council on Artificial Intelligence. Available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

49. Brazilian Artificial Intelligence Strategy, MCTI Ordinance No. 4.979. Available at: https://www.in.gov.br/web/dou/-/portaria-mcti-n-4.979-de-13-de-julho-de-2021-332164912

50. Artificial Intelligence Strategy of the German Federal Government. Available at: https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf

51. For a Meaningful Artificial Intelligence: Towards a French and European Strategy. Available at: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf

52. Report of the Artificial Intelligence Task Force. Available at: https://dpiit.gov.in/sites/default/files/Report_of_Task_Force_on_ArtificialIntelligence_20March2018_2.pdf

53. The Agenda in Brief - Artificial Intelligence in Mexico: A National Agenda. Available at: https://36dc704c-0d61-4da0-87fa-917581cbce16.filesusr.com/ugd/7be025_85f5cec6ea584d8a842d11ad401c0685.pdf

54. AI - Here for Good: National Artificial Intelligence Strategy for Ireland. Available at: https://www.gov.ie/en/publication/91f74-national-ai-strategy/

55. Dutch Strategic Action Plan for Artificial Intelligence. Available at: https://www.rijksoverheid.nl/ministeries/ministerie-van-economische-zaken-en-klimaat/documenten/beleidsnotas/2019/10/08/strategisch-actieplan-voor-artificiele-intelligentie

56. Artificial Intelligence: a strategic vision for Luxembourg. Available at: https://digital-luxembourg.public.lu/sites/default/files/2020-09/AI_EN_0.pdf

57. EU General Data Protection Regulation, Regulations (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX-:02016R0679-20160504&from=EN

58. The G20 is comprised of Argentina, Australia, Brazil, Canada, China, France, Germany, India, Indonesia, Italy, Japan, Republic of Korea, Mexico, Russia, Saudi Arabia, South Africa, Turkey, the United Kingdom, the United States, and the European Union. G20 AI Principles (2019). Available at: https://www.mofa.go.jp/files/000486596.pdf, p.11-14.

59. European Commission Independent High-Level Expert Group on Artificial Intelligence (2019), Ethics Guidelines for a Trustworthy AI. Available at: https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1

60. Written contribution of the United Kingdom to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/UK.pdf

61. Written contribution of France and Germany to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/France-Germany.pdf

62. Statement by China, Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems, 30 September 2021. Available at: https://media.un.org/en/asset/k1g/k1gjinrfym

63. Statement by Austria at the Sixth Review Conference of the Convention on Certain Conventional Weapons, 13 December 2021. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/RevCon/statements/13Dec_Austria.pdf

64. UNESCO Recommendation on the ethics of artificial intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation

65. European Parliament Committee on Legal Affairs (2021), Report on artificial intelligence: questions of interpretation and application of international law in so far as the EU is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice (2020/2013(INI)). Available at: https://www.europarl.europa.eu/doceo/document/A-9-2021-0001_EN.pdf

66. European Parliament Committee on Civil Liberties, Justice and Home Affairs (2021), Report on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)). Available at: https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_EN.html

67. For a Meaningful Artificial Intelligence: Towards a French and European Strategy. Available at: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf

68. The G20 is comprised of Argentina, Australia, Brazil, Canada, China, France, Germany, India, Indonesia, Italy, Japan, Republic of Korea, Mexico, Russia, Saudi Arabia, South Africa, Turkey, the United Kingdom, the United States, and the European Union. G20 AI Principles (2019). Available at: https://www.mofa.go.jp/files/000486596.pdf, p.11-14.

69. AI Portugal 2030. Available at: https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf

70. Dutch Strategic Action Plan for Artificial Intelligence. Available at: https://www.rijksoverheid.nl/ministeries/ministerie-van-economische-zaken-en-klimaat/documenten/beleidsnotas/2019/10/08/strategisch-actieplan-voor-artificiele-intelligentie

71. Australia's Artificial Intelligence Ethics Framework. Available at: https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles

72. Norwegian National Strategy for Artificial Intelligence. Available at: https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf

73. EU General Data Protection Regulation, Regulations (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX-:02016R0679-20160504&from=EN. For more on the GDPR and autonomy in weapons systems, see, for example, Article 36 (2022), Foundations of the GDPR relevant to discussion on autonomous weapons. Available at: https://article36.org/wp-content/uploads/2022/03/AWS-Foundations-of-the-GDPR.pdf

74. Written contribution of Switzerland to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/Switzerland.pdf

75. Statement by the Group of NAM at the Sixth Review Conference of the Convention on Certain Conventional Weapons, 13 December 2021. Transcript available at: http://149.202.215.129:8080/s2t/UNOG/RCHCP6-13-12-2021-AM_mp3_en.html

76. Written contribution of France and Germany to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/France-Germany.pdf

77. Written contribution of the United Kingdom to the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/UK.pdf

78. See, for example, statement by Argentina on behalf of Argentina, Costa Rica, Ecuador, El Salvador, Palestine, Philippines, Panama, Peru, Sierra Leone, and Uruguay at the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the area of Lethal Autonomous Weapons Systems. Available at: https://media.un.org/en/asset/k13/k13owhtps9

79. Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions (2020), Use of armed drones for targeted killings, A/HRC/44/38. Available at: https://undocs.org/en/A/HRC/44/38

80. Christof Heyns (2016), 'Autonomous weapons systems: Living a dignified life and dying a dignified death', in Bhuta et al. (eds) Autonomous Weapons Systems: Law, Ethics, Policy (Cambridge University Press), 18.

WWW.AUTOMATEDRESEARCH.ORG
INFO@AUTOMATEDRESEARCH.ORG
@AUTOMATEDDR

AUTOMATED
DECISION
RESEARCH