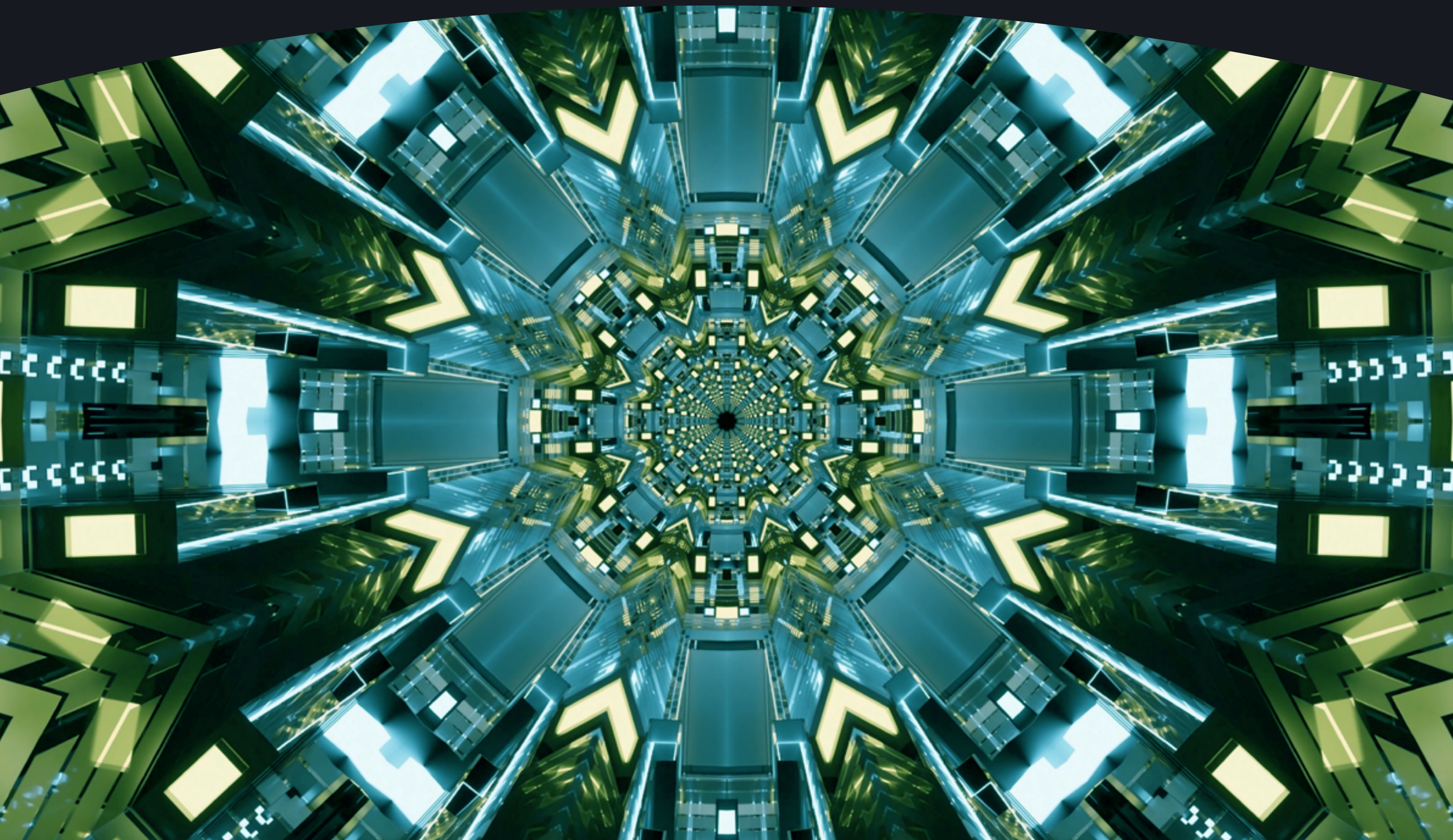
UNIDIR

The Global Kaleidoscope of Military AI Governance

Decoding the 2024 Regional Consultations on Responsible AI in the Military Domain

YASMIN AFINA



Acknowledgments

Support from UNIDIR's core funders provides the foundation for all of the Institute's activities. This policy brief was prepared by UNIDIR's Security and Technology Programme, which is funded by the Governments of Czechia, Germany, Italy, the Netherlands, Norway, the Republic of Korea, Switzerland and the United Kingdom, and by Microsoft.

The author wishes to thank Dr. Giacomo Persi Paoli (UNIDIR) for advice, guidance and support for this programme of work and for his review of this report; as well as Ioana Puscas (UNIDIR), Netta Goussac, Dr. Alexander Blanchard and Major Jamal Mohamed Hassan for their reviews and suggestions. The author also wishes to thank the Republic of Korea and the Netherlands for convening the regional consultations, along with Chile, Costa Rica, Kenya, Singapore and Türkiye, which acted as co-hosts, as well as the Centre for Humanitarian Dialogue for the facilitation of the forecasting exercise that preceded four of the regional consultations.

About UNIDIR

The United Nations Institute for Disarmament Research (UNIDIR) is a voluntarily funded, autonomous institute within the United Nations. One of the few policy institutes worldwide focusing on disarmament, UNIDIR generates knowledge and promotes dialogue and action on disarmament and security. Based in Geneva, UNIDIR assists the international community to develop the practical, innovative ideas needed to find solutions to critical security problems.

Note

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. The views expressed in the publication are the sole responsibility of the individual authors. They do not necessary reflect the views or opinions of the United Nations, UNIDIR, its staff members or sponsors.

About the Security and Technology Programme

Contemporary developments in science and technology present new opportunities as well as challenges to international security and disarmament. UNIDIR's Security and Technology Programme seeks to build knowledge and awareness on the international security implications and risks of specific technological innovations and convenes stakeholders to explore ideas and develop new thinking on ways to address them.

Author

This report was produced by UNIDIR's Security and Technology Programme. It was drafted by Yasmin Afina, who attended all five regional consultations.

Yasmin Afina

Researcher, Security and Technology

Yasmin Afina is a Researcher for the Security and Technology Programme at UNIDIR, where her research covers the intersection between international security, international law and artificial intelligence. Her research experience and interests cover nuclear weapons policy, outer space security, and wider international security and policy issues surrounding emerging technologies, including neurotechnology, quantum technologies, and cyber. She is also a PhD Researcher in law at the University of Essex.

Acronyms & Abbreviations

AI	Artificial intelligence
CCW	(Convention on) Certain Conventional Weapons
ISR	Intelligence, surveillance and reconnaissance
NATO	North Atlantic Treaty Organization
NSAG	Non-State armed group
REAIM	Responsible AI in the Military Domain
UAV	Uncrewed aerial vehicle
UNIDIR	United Nations Institute for Disarmament Research

Contents

Executive Summary

In the run-up to the second iteration of the Responsible AI in the Military Domain (REAIM) Summit, to be held in Seoul, Republic of Korea, on 9–10 September, the Governments of the Republic of Korea and the Netherlands organized, in partnership with Chile, Costa Rica, Kenya, Singapore and Türkiye, a series of five regional consultations on responsible artificial intelligence (AI) in the military domain. Four of the consultations were held in-person and were preceded by a one-day forecasting exercise, facilitated by the Centre for Humanitarian Dialogue, which equipped participants with insights into key policy issues associated with the development, deployment and use of AI in the military domain and, when applicable, the wider security domains.

The present report seeks to capture UNIDIR's main reflections on the key takeaways stemming from the five regional consultations. These consultations did indeed enable the dissection of local contexts, realities and approaches with regards to the responsible development, deployment and use of AI in the military and wider security domains – including the identification of areas of convergence and of divergence at the regional level.

The report first discusses the reflections shared by states on the unique characteristics of AI technologies, and the opportunities that they provide in the military domain. These applications range from supporting, or even enhancing, the conduct of warfare (e.g., enhancing intelligence, surveillance and reconnaissance capabilities, cyber operations, and cognitive electronic warfare), to the organizational level (e.g., training, simulation and planning, and back-end support). Many of these opportunities are even found to extend beyond the military domain and have wider security applications (e.g., to counter organized crime and piracy).

In addition to a host of opportunities, states also discussed and exchanged views on the risks, challenges and implications stemming from the development, deployment and use of AI in the military and wider security domains. Perspectives, viewpoints and experiences were shared on a number of risks, challenges and implications, both technological and non-technological. These include an exacerbation of escalation risks, disruptions to the information environment, compliance with international law and ethical considerations, as well as socio-economic risks. It is important to note that, despite many of these risks being shared across most (if not all) regions, there are variations in states' perceptions of the risks and their subsequent approaches to mitigation of the risks. These differences have an impact on their prioritization too, at both the regional and national levels, and how they are subsequently translated into policy or regulatory responses.

The report then covers points of convergence, six of which emerged from the regional consultations:

a. The opportunities offered by responsible AI in the military and wider security domains
b. Compliance with international law and ethical considerations
c. The human element as a key enabler for accountability and responsibility
d. Multi-stakeholder engagement
e. The need for meaningful regional dialogue to understand local contexts
f. The importance of capacity-building efforts

These six areas were consistently discussed and dissected, to varying degrees, across most if not all regional consultations. Thus, in order to achieve widespread buy-in across regions, each of these six aspects should be considered for future governance discussions and pathways on responsible AI in the military domain. It is, however, important to note that each of these six points of convergence is far from being monolithic and should, rather, be considered as having a nuanced nature: each region may be approaching these six themes with variations and nuances due to a number of factors such as political and cultural divergences, varying security landscape, and different socio-economical contexts.

In the same vein, this report identifies five main points of divergence observed across and even within regions:

a. Unique local contexts and realities
b. Varying regulatory approaches and legal traditions
c. Prioritization and risks perception
d. Resource availability and allocation
e. Desired endgame

These differences, like the convergences, are also due to a host of factors, including those mentioned above. However, it is important to note that these variations in approaches are not inherently harmful to an aligned international approach to responsible AI in the military and wider security domains. At times, these variations and nuances are, in fact, desirable or even necessary to account for regional contexts and realities. Most of the time, as long as there is effective coordination at the international level, and mutual understanding of states' and regions' different approaches and rationales (i.e., drivers and underlying reasons for certain policy directions), the points of divergence will not, in and of themselves, lead to fragmentation. Instead, they will lead to a reinforced and inclusive policy and regulatory landscape that is united in diversity, ultimately enabling and promoting responsible AI in the military and wider security domains.

The regional consultations demonstrated that responsible AI in the military and wider security domains, and efforts towards its implementation and operationalization, is a priority area for states across all regions. Yet, the international policy landscape and existing approaches are far from being a monolithic block. The multifaceted and complex elements underpinning the present variations are subsequently, more often than not, reflected in the ways in which each state and each region grapples with the development, deployment and use of AI in the military and wider security domains. As stressed throughout this report, however, these differences are not necessarily intrinsically harmful to international alignment and cooperation which, in most circumstances, states acknowledge the importance of in any case. Processes, discussions and frameworks at the international, regional and national levels must not be seen as mutually exclusive and in competition, but rather as complementarity in the kaleidoscope of military AI governance.

As such, the conduct of the regional consultations – which joined information exchange on policies and national viewpoints on the one hand, with discussions framed by expert intervention on the other hand – was not only useful to ensure inclusivity in the REAIM process. These consultations also paved the way for further reflections, at times shared across regions, at times unique to certain regions. As such, the food for thought to consider for the (future) governance of responsible AI in the military domain include:

▶ Prioritize capacity-building and information exchange

States across regions see the establishment of regularly convened, formalized processes and frameworks for information sharing and the exchange of knowledge and best practices, including from the civilian domain, as being critical. Beyond the policy, legal and ethical realms, there is also the desire to gain granularity at the technical level (e.g., on the black box issue, effective integration and interoperability, computing power and data).

▶ Foster the implementation and operationalization of responsible AI through engagement

States have expressed the desire to move beyond norms, principles and political commitments. There is indeed appetite for initiatives that help bridge the gap between the technical and policy communities for operationalization, implementation and execution.

▶ Ensure the complementarity of processes

States have expressed a desire to shed clarity on the respective mandates, roles and responsibilities of each of the many parallel processes and forums, at the international and regional levels both within and outside the United Nations, and ensure that they complement one another instead of brewing competition.

▶ Intensify research efforts and reflections on underexplored areas

A number of themes have emerged as being underexplored in this space and deemed to require further reflections. These include data governance; dual-use technologies; interoperability; computing power; the traceability and transparency of systems; as well as the different dimensions of security in relation to AI development, deployment and use in the military domain (e.g., human security, environmental security and economic security).

▶ Reflect on AI in the "military domain", its boundaries and interplay with wider security applications

While the regional consultations were all conducted around the theme of "AI in the military domain" (with the Latin American and Caribbean segment also including wider security applications), questions arise with regards to the scope of the domain and the extent to which its boundaries are clearly delineated.

1. Background

The inaugural 2023 Responsible AI in the Military Domain (REAIM) Summit, organized by the Netherlands and co-hosted by the Republic of Korea, was a landmark global and multi-stakeholder summit to collectively discuss, deliberate and address the opportunities, challenges and risks associated with military applications of artificial intelligence (AI). The summit has successfully put the topic of responsible AI in the military domain higher on the political agenda of many states. With its Joint Call to Action, endorsed by over 50 states, it underlined the need to further promote initiatives and efforts to foster the responsible development, deployment and use of military AI.

Against this backdrop, and in the run-up to the second iteration of the REAIM Summit in Seoul on 9–10 September 2024, the Netherlands and the Republic of Korea have sought to build on this momentum and to deepen engagement with partners across regions. As such, and in partnership with regional partners, five regional consultations were held in the first half of 2024:

- **Asia:** Singapore, 27 February 2024 (co-hosted by Singapore and in partnership with the S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University, Singapore)
- **South East Europe, the Middle East, the South Caucasus and Central Asia:** Istanbul, Türkiye, 23 May 2024 (co-hosted by Türkiye and in partnership with the National Defence University of Türkiye)
- **Europe and North America:** Virtual, 30 May 2024
- **Africa:** Nairobi, Kenya, 6 June 2024 (co-hosted by Kenya and in partnership with the Kenya Defence Forces and the National Defence University–Kenya)
- **Latin America and the Caribbean:** Santiago, Chile, 14 June 2024 (co-hosted by Chile and Costa Rica and in partnership with the Law Faculty of the University of Chile)[1]

The consultations held in Singapore, Istanbul, Nairobi and Santiago were preceded by a one-day forecasting exercise for the invited consultation participants and additional observers led by the Centre for Humanitarian Dialogue. These exercises equipped participants with insights into key policy issues associated with the development, deployment and use of AI in the military domain and, in Santiago, the wider security domains, as well as the range of practical approaches to responding to their implications in specific situations. These include the implications of procurement from abroad of AI systems for intelligence, surveillance and reconnaissance (ISR) functions; military decision-making in high-tension and time-sensitive situations; as well as the training on public databases and synthetic data of decision-support systems for law enforcement use.

The regional consultations then provided states with the opportunity to share their national views, priorities and engagements with regards to the governance of AI in the military domain and, when applicable, wider security domains. The format of the consultations was critical in providing participants with a

[1] The Latin American and Caribbean segment of the consultations, in recognition of the region's security landscape and realities, was extended beyond the military to also include the broader security domain (including law enforcement, border security, as well as efforts to counter transnational organized crime).

focused framework for exchanges on trends, approaches and the general context in which states seek to implement principles and norms for responsible AI in military applications and, when applicable, wider security applications – all critical viewpoints to inform subsequent discussions and deliberations at the 2024 REAIM Summit.

In addition to state representatives, the forecasting exercises and consultations were also attended by experts and representatives from a range of organizations including UNIDIR; the United Nations Office for Disarmament Affairs; the Geneva Centre for Security Policy (GCSP); the Center for Security and Emerging Technology (CSET); the National Defence University of Türkiye; Bilkent University; the Foundation for Political, Economic and Social Research (SETA); the Centre for Security Cooperation (RACVIAC); the Organization of Turkic States; the National Defence University–Kenya (NDU-K); Stockholm International Peace Research Institute (SIPRI); the African Observatory on Responsible AI; the REAIM Global Commission; as well as the Fundación para la Paz y la Democracia (FUNPADEM).

Specifically, the regional consultations were designed to provide a platform for open discussions, knowledge-sharing and deepening regional understanding on this emerging issue. As such, they were guided and framed by the following questions:

- What are the most important challenges for each region regarding the responsible use of AI in the military domain and, when applicable, the broader security domains?
- What are the factors that facilitate or, conversely, hinder regional and international cooperation with respect to the responsible use of military AI?
- What are the region’s views on – and challenges to – the application of international humanitarian law in military AI?
- What are the different states’ approaches to implementing responsible AI in the military domain? And what good practices can states share with regards to responsible AI in the military domain?
- What are the states’ views on priorities for international discussions on this topic and what are the challenges for the region?
- What should be the next steps for governance of military AI?

The present report seeks to summarize UNIDIR’s reflections on the main findings and takeaways from the regional consultations. Attendance at the five regional consultations has enabled the dissection of local contexts, realities and approaches with regards to the responsible development, deployment and use of AI in the military and wider security domains – including the identification of areas of convergence and divergence at the regional level. In addition, the conduct of these regional consultations has enabled the identification of areas of cross-regional convergence that may, if leveraged and developed appropriately, open up governance pathways at the international level. On the flipside of the coin, there are a number of points about which regions diverge. However, it is important to note that, in many if not most cases, divergences are not necessarily or intrinsically harmful to an aligned international approach to responsible AI in the military and wider security domains (as discussed in Section 5). A comprehensive mapping of areas of divergence can thus support governance initiatives, which will foster cooperation and mutual understanding. Such efforts will ultimately prevent uncoordinated efforts and harmful fragmentation, while enabling unity in the international community and strength in diversity.

2. General Views on AI Fundamentals, Applications and Opportunities

The consultations provided an opportunity for states to share reflections on the unique characteristics of artificial intelligence technologies, and the opportunities they provide in the military domain. These applications range from organizational roles to supporting or even enhancing the conduct of warfare. Specifically, the following applications were among those discussed by states and experts during the consultations, understood as an overview of the main opportunities AI could offer in the military domain:

- Intelligence, surveillance and reconnaissance
- Target recognition and decision-support systems
- Electronic warfare
- Information and cognitive warfare
- Cyber capabilities
- Autonomous and uncrewed vehicles
- Training, simulation and planning
- Predictive maintenance
- Back-end

Many of these opportunities are even found to extend beyond the military domain and to have wider security applications.

States generally agree that AI holds tremendous potential and offers many and deeply transformative opportunities when developed, deployed and used responsibly in the military and wider security domains. Consistent with the general historical trend that technology has a profound impact on warfare and constitutes a force multiplier, there is a shared understanding that AI considerably enhances existing capabilities across a wide range of applications in the military domain and beyond. In fact, states generally agreed that the military use of AI holds the potential to increase efficiency and reach and, if applied to this end, could be used to mitigate some humanitarian harm of armed conflict. Therefore, and in line with the desire and promise to provide those harnessing novel capabilities with competitive advantage and to make war more frictionless overall, many states and experts see AI innovation as an imperative to preserve and maintain international peace and security.

Unlike other technologies, however, both states and experts across regions share the sentiment that AI is distinguished by its general-purpose nature and accessibility (i.e., lower barrier of entry from the availability of low-cost models, if not free open-source AI). These technologies are in fact deployable across domains, both digitally to enable or enhance cyber capabilities and as part of physical, kinetic systems. The majority of states have thus agreed that AI is indeed poised to play a key enabling role in the military and security domains through increased autonomy, speed and scalability, ranging from navigation to data collection, data processing and target identification. The following examples are among the wide range of applications the states have identified where AI could bring tremendous opportunities in the military and security domains.

Intelligence, surveillance and reconnaissance

The ability to collect and process vast amounts of ISR data at speed and at scale will bring tremendous advantages for enhanced situational awareness and, ultimately, to support decision-making. These capabilities are particularly appealing for operations in cluttered environments where, coupled with remote systems such as uncrewed aerial vehicles (UAVs), they enable rapid and reliable target identification and recognition (e.g., using computer vision programmes). In naval warfare, AI enables the collection of intelligence to be continuous and uninterrupted, which enables the processing of vast amounts of data in a timely manner to support military and security operations (e.g., counter-piracy).

Advances in natural language processing also play a critical part in the digital space – for example, to enable multilingual speech or text recognition across digital platforms.

Target-recognition and decision-support systems

In order to complement ISR capabilities, AI-enabled decision-support systems are also increasingly explored as a means to support command and control across all domains of operations. In principle, these capabilities can also be integrated into weapons systems either with or without humans in the loop.

Electronic warfare

The notion of "cognitive electronic warfare" – AI-enabled electronic warfare capabilities – has emerged and has been discussed. Possible applications include, for example, the enhancement of electronic warfare capabilities (including for target identification and engagement), as well as threat detection, interception and disruption within short time frames and at scale.

Information and cognitive warfare

AI offers opportunities in the psychological realm. While AI is generally found to exacerbate disinformation risks and make them more complex, it also holds the potential to counter hostile information operations. For instance, technological solutions are being developed to identify content made by generative AI (e.g., watermarking). Blockchain has also been identified as a promising means to help with authentication to contribute to such counter-disinformation efforts. In addition, these technologies can also help with sentiment analysis and for the prediction of future attacks. In the longer term, AI's convergence with neurotechnology and the rise of brain–computer interfaces is also to be explored (e.g., for human enhancement).

Cyber capabilities

AI is both a force multiplier and a threat multiplier, including in the cyber domain. For instance, AI can create advanced threat-emulation systems that enable better planning and, subsequently, the development of cyber defensive capabilities to safeguard against the increasing sophistication of cyber offensive capabilities (including those AI-enabled). This is particularly important to safeguard states' critical national infrastructure against risks of being exposed to cyber operations. In addition, AI can play a critical role in detecting anomalies at a much faster pace, as well as subsequently enhancing incident responses which, at times, may not even necessarily require humans in the loop.

Autonomous and uncrewed vehicles

These include, in particular, swarming capabilities across all domains of operation.

In addition, because of its general-purpose nature, AI's impact and implications will reach beyond the conduct of warfare. The areas in which the armed forces and other adopting organizations, including law enforcement agencies, will see implications on the organizational level include the following.

Training, simulation and planning

Advances in generative AI hold the promise of providing the armed forces with safer, synthetic environments for the conduct of training exercises and for planning purposes (e.g., to predict the adversary's behaviour, tactics and strategy).

Predictive maintenance

AI can be used to estimate, anticipate and plan when systems would need to undergo repairs or servicing, especially for those capabilities that require regular maintenance (e.g., heavy artillery).

Back-end

AI also holds tremendous potential in enhancing back-end capabilities, ranging from logistics (e.g., ammunition storage and management) to streamlining administrative processes. In addition, armed forces generally consist of a vast pool, and a wide variety, of personnel. AI-enabled systems are expected to greatly affect personnel management as they cannot only automate and streamline a number of human resources-related processes, but can also make the most of the pool of talent available to leverage the personnel's respective skills and knowledge.

States' and experts' deliberations have demonstrated that the above-mentioned AI applications are being increasingly integrated across domains of operation – in the air, on land, at sea, in space and in cyber operations. A number of states across all regions have indeed announced the adoption, to a certain extent, of a number of the above-mentioned capabilities. For instance, the navies of certain African states are using AI-enabled ISR technologies in counter-piracy efforts and, more generally, to track and curb organized crime across borders. UAVs with increasingly autonomous functions are also heavily commercialized (e.g., by Türkiye, Iran and China) and deployed in various armed conflicts around the globe. In Latin America and the Caribbean, AI already plays a critical role in enhancing cyber defensive capabilities – especially to compensate for the lack of resources and personnel that certain states face against the increased sophistication of offensive cyber operations conducted by non-state armed groups (NSAGs) and criminal organizations.

3. General Views on Risks, Challenges and Implications

In addition to a host of opportunities, states also discussed and exchanged views on the risks, challenges and implications stemming from the development, deployment and use of AI in the military and wider security domains. These range from the technological to the legal and the socio-economic.

The following risks, challenges and implications were among those discussed by states and experts during the consultations:

- Cyber risks
- Accessibility
- Risks of misuse
- Conflict and security dynamics
- Characteristics of AI technologies
- Disruptions to the information environment
- Escalation
- Availability of reliable data
- Economic and social implications
- Structural risks
- Overdependency and technical illiteracy
- Compliance with international law and ethical considerations
- Interoperability and integration

In parallel to the opportunities offered by AI technologies, they introduce a host of risks that can be novel or can make existing risks more complex or even exacerbate them. These risks are multidimensional and can stem from a number of factors, including the technology's inherent characteristics, the surrounding conflict dynamics, regional realities on the political and legal fronts, the culture of adoption and use, as well as risks to and from persons in conflict. Disinformation, inadvertent escalation, proliferation to non-state armed groups, as well as the increased vulnerability of systems at risk of adversarial attack constitute some of the many risks shared during the consultations. These risks apply both in armed conflict and below its threshold (i.e., in law enforcement operations, during peacetime), which requires states to be prepared and to have clear risk-mitigation plans in both instances. Some, if not most, of the following risks are generally shared across regions; however, their order of importance and magnitude generally tend to diverge from one another (see Box 1).

Cyber risks

The integration of AI not only across the military domain but also more widely (e.g., in critical national infrastructure) will have a number of security implications. Safeguarding them against cyber operations will require a level of resilience, in addition to dedicated resources and capabilities.

Accessibility

Many factors – including the mass and rapid commercialization of products, the affordability of these technologies and the advent of open-source AI – provide plenty of opportunities to increase the accessibility of AI for security and defence applications. Ensuring that measures are in place to prevent access for misuse, however, will be critical. Policy responses must strike the right balance between the need to curb proliferation risks versus maintaining the opportunities offered by accessible AI (e.g., open-source AI). These measures must not work unfairly at the expense of states with more limited resources and that would (greatly) benefit from lower points of access to these technologies.

Risks of misuse

The dual-use nature of many, if not most, technologies raises concerns related to access and misuse by NSAGs. This is the case, for instance, with the use of generative AI by subcontracted hacktivists to write or enhance malware; another example pertains to the repurposing of uncrewed vehicles. Beyond access concerns, questions also arise with regards to traceability and accountability, both made even more complex in the event of acquisition and use of AI technologies by NSAGs. The acquisition of these technologies and subsequent misuse by NSAGs may cause increased destabilization from hybrid warfare. Furthermore, misuse can also occur in the hands of intended and licit users, which further accentuates the need for traceability and measures for upholding accountability and responsibility of use.

Conflict and security dynamics

The deployment of AI technologies will inevitably affect local and regional dynamics – especially certain sensitive applications, including for border security and electronic warfare. However, the impact that AI will have on conflict and security dynamics remains very much speculative at this stage; although a host of concerns and risks arise, especially in tense environments where distrust prevails (e.g., in the case of erroneous early warning against risk escalations). As such, pre-existing state-to-state relationships and also those between states and other actors (especially for contexts where NSAGs or organized crime prevail) will bear great influence over how these dynamics ultimately unfold.

Characteristics of AI technologies

Given the probabilistic nature of AI technologies and the inherent black box in many, if not most, applications, AI could function unpredictably, unexpectedly and unexplainably. The legal, policy and ethical implications of these technologies across the military and security domains remain to be determined, especially as information and evidence on the concrete impact of these technologies in security and defence remain relatively limited to date.

These concerns, however, ought to be counterbalanced by the inherent uncertainty of warfare – the "fog of war": Thus, in the context of military operations, it would be the commander's responsibility to adopt a holistic approach to assessing the opportunities and risks from the eventual use of certain AI technologies.

Disruptions to the information environment

Artificial intelligence and, in particular, generative AI can erode trust in the information environment, both internally and between states. The veracity and accuracy of military intelligence may be at risk; and the generation of deepfakes can, for example, disrupt intelligence-sharing frameworks between two states.

Escalation

AI will affect perceptions, assumptions and judgments and, thus, can (heavily) affect risk assessments and lead to unintended escalation. Indeed, AI could reduce barriers to the use of force, in addition to opening the door to miscalculations and unintended consequences, which, in turn, can result in escalatory responses. These escalation risks will be increased in contexts of high tension.

Availability of reliable data

Cognizant of the role of data in the development, training, testing and use of AI in the military and security domains, it is critical that high-quality data sets that reflect the local contexts and realities are used for such purposes. Yet, the availability of local data may, at times, be a challenge especially for states where digital transformation is still on-going.

Economic and social implications

The unregulated and unrestrained development, deployment and use of AI in the military domain would have a number of economic and social implications. These can range from public perceptions of the legitimacy of such use to the second- and third-order impact of NSAGs acquiring, deploying and using such technologies (e.g., impact on trade and, thus, the economy). These will ultimately have an impact on the availability and subsequent allocation of resources for upholding responsible practices surrounding AI in the military domain.

Structural risks

A number of risks stem from structural issues that have subsequent implications for responsible practices surrounding AI in the military and wider security domains. These include societal biases in training and testing data; algorithmic bias; the need for contextualized data versus the unavailability of local data; limitations or even scarcity in the financial, human or technical resources available to promote responsible AI; as well as the main players in the local and regional security landscape.

Overdependency and technical illiteracy

AI holds tremendous potential in enhancing situational awareness in conflict and, thus, informing the commander's risks assessments and decision-making. Amidst the increased use and even dependency on AI-enabled decision-support systems, the absence of adequate training and clear rules of engagement, coupled with techno-solutionism, carry risks with regards to the system's reliability, compliance with international law and ethical requirements. Some states were even of the view that such risks could, potentially, go as far as affecting national, regional and international peace and security, for example, in the context of conflict escalation due to overreliance on the output of AI technologies.

Compliance with international law and ethical considerations

States generally prioritize compliance with international law and ethics in their national and regional approaches to responsible AI in the military domain. This is generally reflected in regional and national policies and strategy documents, as well as in states' national positions in regional and international deliberations. However, the development, deployment and use of these technologies, along with their procurement (especially those off-the-shelves) raises questions about the states' ability to comply with these frameworks.

Interoperability and integration

The dependence of many states on foreign technologies, especially from certain regions, raises the question of coordination, interoperability, reliability (i.e., against the need for contextualized training), safety and security (against different testing standards). The interaction between novel and legacy systems adds another layer of complexity and uncertainty with regards to interoperability.

BOX 1.

Varied Approaches to Risks, Challenges and Implications

While many, if not most, of the concerns about the risks of AI in the military domain are generally shared across regions, they tend to be approached and prioritized differently at the regional or even subregional level. These divergences are reflected, for instance, in states' assessment with regards to their likelihood and (potential) impact; the key actors; types of harm involved against local contexts; orders of consequences; as well as their time frame (i.e., whether they are considered as short-, medium- or long-term risks). Some regions may prioritize risk-reduction efforts and measures in the realms of legal compliance and ethics; others' policies may revolve further around increasing national security.

Perceived risks differ too at the cross-regional and subregional levels. In Asia, a number of states would prioritize risks surrounding the misuse of these technologies and their appropriation by and proliferation to "malicious" actors. Others are more focused on the humanitarian impact and risks stemming from these technologies. Risks associated with strategic stability, nuclear affairs and regional stability are also among those raised as a priority by some. However, it is important to note that the differences in risk perceptions and prioritization were not necessarily found to be mutually exclusive by states within the region.

Similarly, in Africa, some states would prioritize addressing the (existential) risks stemming from the possible integration of AI into nuclear command, control and communications systems; while others are of the view that these risks are less relevant for the region and that other risks prevail. In Latin America and the Caribbean, a distinct emphasis is put on the risk of proliferation of these technologies into the hands of non-state armed groups and their subsequent use to exacerbate organized criminal activities. Additionally, the region is also particularly concerned by the far-reaching impact of AI on the advancement of the United Nations' Sustainable Development Goals (SDGs); as well as the preservation of human rights in the light of the development, deployment and use of AI in the military and wider security domains and its subsequent impact on political, socio-economic and cultural rights.

4. Key Policy Priorities: Points of Nuanced Convergence

Through the five consultations, six points of convergence were identified as surfacing and applying across regions:

> **Facets of responsible AI practices in the military and wider security domains: Taking stock of opportunities**

> **Compliance with international law and ethical considerations**

> **The human element as a key enabler for accountability and responsibility**

> **Multi-stakeholder engagement**

> **The need for meaningful regional dialogue to understand local contexts**

> **The importance of capacity-building efforts**

There are, however, nuances within and across regions in these commonalities. These variations are due to a host of factors, including political and cultural differences, varied socio-economical contexts, as well as differences in legal traditions.

At the cross-regional level, a number of commonalities, patterns and areas of convergence emerged from the consultations. There are, generally, shared sentiments between states on the following six thematic areas:

a. Facets of responsible AI practices in the military and wider security domains: Taking stock of opportunities

b. Compliance with international law and ethical considerations

c. The human element as a key enabler for accountability and responsibility

d. Multi-stakeholder engagement

e. The need for meaningful regional dialogue to understand, and factor in, local contexts – which is key to promoting, enabling and operationalizing the responsible development, deployment and use of AI in the military and wider security domains

f. The importance of capacity-building efforts

Together, these similarities constitute building blocks for the responsible development, deployment and use of AI in the military domain. However, they are far from being monolithic and there are nuances and variations across regions, despite shared patterns and the existence of common threads. In fact, akin to the general observation on risks, regions may be approaching these areas of convergence differently, hence they are "nuanced".

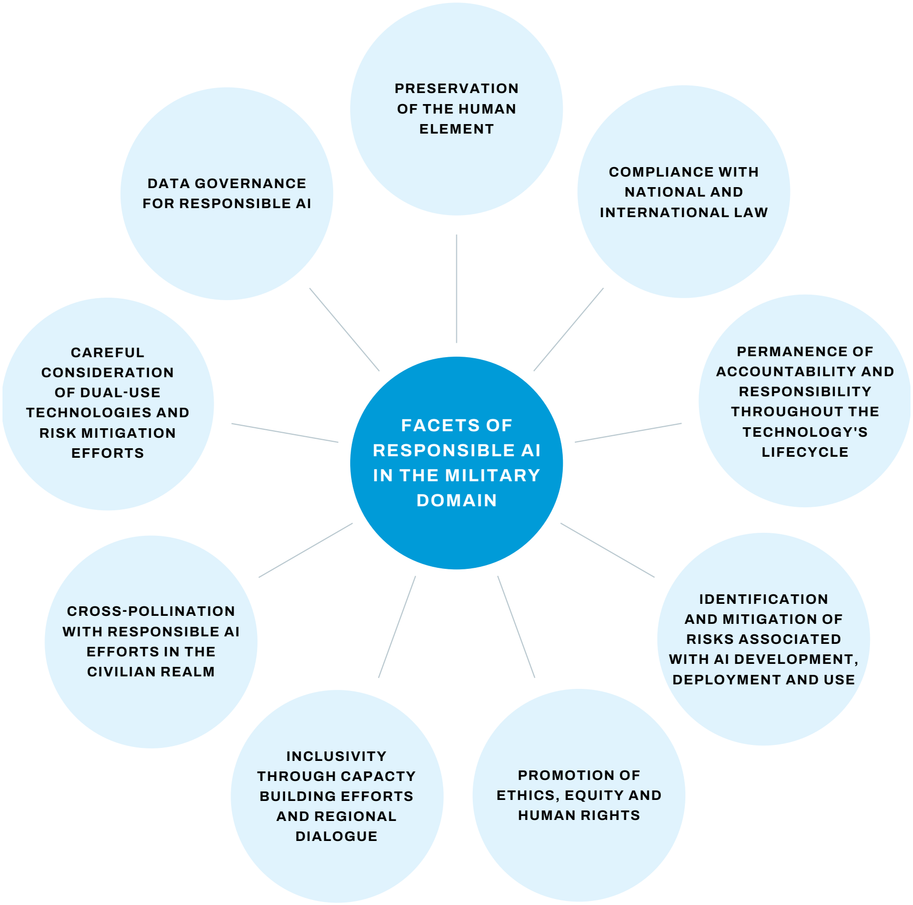
4.1 Facets of Responsible AI Practices in the Military and Wider Security Domains: Taking Stock of Opportunities

There is agreement, across regions, that AI in the military domain offers a host of opportunities, from enhancing ISR capabilities to supporting command and control, logistics and planning, as well as personnel management. Integration of AI across all functions and domains can, indeed, act as a force-multiplier, with many even considering the adoption of these technologies as a necessity to respond to the evolving nature of warfare and conflict (e.g., with the rise in hybrid warfare and asymmetrical threats). Beyond combat, AI is also seen as a critical component of the conduct of (collective) self-defence, and even conflict prevention. These opportunities extend far beyond the military domain and concern wider security applications, including to support law enforcement, bolster border security, enhance counter-piracy efforts, as well as aid with humanitarian and relief efforts in response to natural disasters.

Yet, there is also the shared sentiment that responsible behaviour and practices must be upheld while these opportunities are being explored and harnessed and must remain central to the development, deployment and use of AI in the military domain. While there is no consensus as to what “responsible” means, Figure 1 shows elements and considerations that constitute some of the facets of responsible AI in the military domain shared across regions.

FIGURE 1.

Facets of Responsible AI in the Military Domain


FACETS OF RESPONSIBLE AI IN THE MILITARY DOMAIN
PRESERVATION OF THE HUMAN ELEMENT
COMPLIANCE WITH NATIONAL AND INTERNATIONAL LAW
PERMANENCE OF ACCOUNTABILITY AND RESPONSIBILITY THROUGHOUT THE TECHNOLOGY'S LIFECYCLE
IDENTIFICATION AND MITIGATION OF RISKS ASSOCIATED WITH AI DEVELOPMENT, DEPLOYMENT AND USE
PROMOTION OF ETHICS, EQUITY AND HUMAN RIGHTS
INCLUSIVITY THROUGH CAPACTY BUILDING EFFORTS AND REGIONAL DIALOGUE
CROSS-POLLINATION WITH RESPONSIBLE AI EFFORTS IN THE CIVILIAN REALM
CAREFUL CONSIDERATION OF DUAL-USE TECHNOLOGIES AND RISK MITIGATION EFFORTS
DATA GOVERNANCE FOR RESPONSIBLE AI

4.2 Compliance with International Law and Ethical Considerations

The overwhelming majority of states across regions place compliance with international law and ethical considerations, a key complementary element to the latter (see Box 2) as a central component of their governance approaches to AI in the military, and even wider security domains.

There is indeed a shared sentiment that international law is an important framework that must be upheld throughout the life cycle of AI technologies. Among other things, international law considerations must be considered from the earliest stages (i.e., design, development and testing), which would require efforts to "translate" international law into technical requirements in order to frame and shape the pre-deployment stages of these technologies. In addition, international law – and in particular international humanitarian law and international human rights law – ought to inform, or even frame and shape, procurement processes as many states are increasingly considering purchasing AI-enabled capabilities. To this end, guidance documents and frameworks with best practices for legal compliance would be useful, in particular for states without vast resources and capacity available for such efforts.

From a policy standpoint, however, there are nuances across regions in states' approaches to international law. States in Latin America and the Caribbean, for example, generally dedicate more attention and efforts to foster compliance with, and uphold, international human rights law. While states in all regions acknowledge the importance of the latter, international humanitarian law tends to overwhelmingly dominate the policies and discourse of states in regions other than Latin America and the Caribbean and, as discussed below, Africa. This approach is reflective of the regional security landscape, where transnational efforts at combatting organized crime prevail and in the light of international human rights law's applicability both in and outside conflict. Nevertheless, this is not to discount the importance of international humanitarian law in the Latin American and Caribbean region. These states' approach to and interpretation of international humanitarian law, however, tends to differ from that of other regions by revolving around its underlying spirit – that is, minimizing the effects of war and protecting civilians – in hand with the predominating perception that AI further dehumanizes war and conflict. In contrast, other regions would tend to approach international humanitarian law from a permissibility and compliance standpoint.

Human rights also play a critical role in Africa, particularly within the framework of the African Charter on Human and Peoples' Rights; the underlying reasons, however, vary. In the light of the strong emphasis in Africa on contextualization, this approach enables the consideration of the wider socio-economic implications stemming from the development, deployment and use of AI in the military domain.

States in Asia, while affirming the importance of ensuring compliance "by design" and throughout the technology's life cycle, put great emphasis on the importance of realistic monitoring and enforcement mechanisms based on scientific evidence, through the input and expertise of the technical community. Furthermore, there are specific concerns with regards to the implications that AI technologies will have for the application of international maritime law, which many in the region believe will require further reflection.

Meanwhile, states in South East Europe, the Middle East, the South Caucasus and Central Asia are particularly concerned about the applicability and subsequent application of international law to non-state actors. These concerns are twofold. First, as non-state armed groups are increasingly seeking to adopt and use AI technologies to support their military operations, these groups' liability, and the ability to hold them accountable, will depend on a number of variables such as the political will of parties involved, the resources available, as well as accessibility to conflict areas for investigation and evidence collection. Second, in addition to state responsibility and that of individuals (e.g., the commander, through international criminal law), the development, deployment and use of AI in the military domain raises questions about corporate accountability. This question is of particular relevance for incidents resulting from a technology's malfunction: Many states have indeed shared concerns with regards to the black box nature of AI systems that will make it difficult, if not impossible, to ensure traceability and thereby establish accountability and responsibility. This conundrum is made more complex in situations where states are procuring capabilities from foreign defence contractors or technology companies – an issue many in other regions are encountering too.

BOX 2.

Ethics: A Key Element Complementary to International Law

Beyond international law, states across regions generally shared the view that ethics constitute an important approach to complement international law for the governance of AI in the military domain. While there is no consensus as to the key ethical principles that underpin the "responsible" development, deployment and use of AI in the military domain, the following elements constitute some of the main considerations shared by states throughout the consultations:

- Traceability
- Explainability
- Equity
- Accountability
- Humanity
- Fairness
- Responsibility
- Transparency

4.3 The Human Element as a Key Enabler for Accountability and Responsibility

Building on the previous point, states generally agree that accountability and responsibility must be upheld in the context of the development, deployment and use of AI in the military and wider security domains. There are a number of ways through which accountability and responsibility can be established. On an individual basis, commanders will always be responsible for the military operations that they helm. States may also be held liable for attributable internationally wrongful acts in the development, deployment and use of AI in the military domain. Beyond these, a number of states across regions are also keen on exploring the extent to which corporate responsibility can be upheld, including the individual responsibility of developers involved in the design, development and testing of these technologies, especially with regards to incidents.

The human element is often raised as a key enabler to establishing and maintaining accountability and responsibility. To this end, two particular aspects of the human element are central:

a. Human judgment, intervention, oversight and control
b. The effective sensitization and training of individuals involved throughout the life cycle of the technology – from its design and development to its testing, deployment and subsequent use

Generally, human–machine teaming, with varying degrees of human judgment, intervention, oversight and control, is seen as necessary to maintain accountability and responsibility in decision-making. Certain states would even argue that there may be circumstances in which humans in the loop are not necessarily feasible, nor desirable for those deploying these systems; the deployment of AI-enabled cyber defence capabilities being one of such examples. Nevertheless, commanders will always, ultimately, maintain individual responsibility for military operations, including those that involve AI-enabled technologies and including operations in the cyber realm. Clarity on the extent to which humans are involved and can exercise judgment, intervention, oversight and control over the technology can only help with establishing and maintaining accountability and responsibility.

Furthermore, and in order to enable effective and reliable human–machine teaming, there is the shared sentiment that the sensitization and training of individuals involved throughout the life cycle of the technology must be prioritized. This implies raising awareness among technologists and developers on legal requirements and ethical considerations in order to factor these “by design”; end users will also need to have undergone training in order to fully understand the technology, including its capabilities and limitations. The latter is particularly important in order to maintain the accountability and responsibility of commanders with regards to risks assessments and the eventual decision to deploy and use AI-enabled technologies for military operations.

In addition, consultation participants from all regions drew attention to the rigorous deliberations undertaken in this space, specifically in the context of autonomous weapons systems both within the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts on Lethal Autonomous Weapons Systems and within the United Nations General Assembly. A number of states have discussed (or, in certain cases, expressed the desire to reflect on) how to connect and bridge the discussions occurring in different forums, as well as what lessons could be drawn for future, broader discussions on AI in the military and wider security domains.

4.4 Multi-Stakeholder Engagement

States generally recognize the value of multi-stakeholder and cross-sectoral engagement to promote responsible AI in the military domain. A number of (varying) incentives to conduct such engagement have been put forward, including the following examples.

Compliance by design

Multi-stakeholder engagement, especially with industries, provides the platform needed to take stock of legal frameworks and ethical considerations, and "translate" them into technical requirements and solutions in order to foster compliance by design.

Effective implementation and operationalization of responsible norms and behaviour

Engaging the multi-stakeholder community will ensure that policy discussions, deliberations and solutions remain grounded, realistic and evidence-based. This is a critical feature in the light of the technical nature of AI integration in the military domain and in addition to its potentially profound implications for peace and security. As such, multi-stakeholder platforms provide room for evidence-based forecasting, risk assessments and the subsequent development of risk-mitigation measures.

Risk mitigation

Platforms for cross-sectoral and multi-stakeholder exchange of good practices and information will be key for risk mitigation, especially to feed into and consolidate compliance efforts from the "upstream" stages in the technology's life cycle.

Cross-pollination from other security fields

A number of states and experts alike see multi-stakeholder engagement as an opportunity to draw lessons from the governance of other security fields, such as the nuclear and cyber domains. Cross-sectoral dialogue, coupled with comparative studies, could subsequently inform and provide inspiration, as appropriate, for governance efforts surrounding responsible AI in the military domain.

Inclusivity

Perspectives drawn from industry, civil society and academia will ensure that governance solutions for responsible AI in the military domain are inclusive, equitable and holistic; and that there is clarity as to what role each stakeholder can play in their implementation and operationalization.

Capacity-building and awareness-raising

Multi-stakeholder dialogue bridges perspectives and efforts between and across communities, from the scientific and technical circles to those in the policymaking, legal and ethical spaces. Multi-stakeholder consultations will thus raise awareness on alternative perspectives and approaches to responsible AI in the military domain. These will need to be coupled with capacity-building efforts to foster mutual understanding and governance pathways that are actionable and holistic. It is hoped by many states and experts alike that these will subsequently pave the way for risk-mitigation and governance measures that are coordinated across regions and across sectors, in addition to solidifying incident responses and fostering preparedness globally.

For instance, certain states in Latin America and the Caribbean and in Asia are of the view that there is a need to increase the awareness of industry actors of applicable laws and the subsequent implications for the design and development of AI technologies in the military domain. For instance, companies should consider whether the development of systems to support war efforts could be considered as direct participation in hostilities and, thus, would make developers lawful targets under international humanitarian law in an armed conflict. There is indeed no clarity as to whether companies, especially small and medium enterprises, have in-house capacity and processes to test against states' legal obligations and ultimately foster compliance.

There are, however, diverging views as to how multi-stakeholder engagement should be conducted, and why. In fact, while states generally agree that it is at least important to consider multi-stakeholder perspectives, regional and national sensitives cause disagreements as to the best approach to this issue and the motivations. States' views particularly diverge with regards to the level of involvement by industry: despite the common understanding that parallel discussions with industry and the overall technical and expert community are much needed, the level to which they should be involved in governance discussions and deliberations remain contested. Some are of the view that norm-setting and policymaking exclusively remain the sovereign prerogative of states and, thus, their development, negotiation and adoption must continue to be done by states, while companies, civil society and academia will play a critical role in helping implement norms and principles. Others were of the view that input from industry and civil society organizations is essential and must be embedded within formal deliberations and processes; a number of states have raised and discussed the model followed in multilateral discussions on cyber that involve, to a certain extent, industry representatives who are provided space to share knowledge, technical expertise and perspectives, while ensuring that states maintain their sovereign prerogative for policymaking and regulation-making. These variations were not only present across regions but also among states within the same region.

As such, different forms of multi-stakeholder engagement have emerged. For instance, a number of states are keen to explore formal public–private partnerships to promote the responsible development, deployment and use of technologies. Informal forms of multi-stakeholder engagement were also discussed, such as joint capacity-building activities (e.g., multi-stakeholder tabletop exercises), as well as informal exchanges of best practices. A number of states across regions have expressed support for initiatives akin to UNIDIR's Roundtable for AI, Security and Ethics (RAISE), which provides for an independent and neutral platform that generate cross-regional and cross-sectoral perspectives, recommendations and action points on the governance of AI in the military and wider security domains.[2] The Global Commission on Responsible AI in the Military Domain (GC-REAIM), established as an outcome to the first REAIM Summit, constitutes another example of such efforts that has been deemed as promising by certain states.

[2] UNIDIR, 'RAISE: The Roundtable for AI, Security and Ethics', https://unidir.org/raise/

4.5 Meaningful Regional Dialogue and Frameworks

There is a general appetite for maintaining dialogue at the regional level on AI governance in the military and security domains, as well as the exchange of information and best practices. Far from supplanting multilateral processes, such regional dialogues would in fact complement and consolidate the discussions held at the international level on the issue of AI in the military and wider security domains (e.g., within REAIM, within the framework of the United States-led Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, as well as within the First Committee of the United Nations General Assembly). While higher-level political discussions can take place in multilateral and international forums, parallel regional deliberations provide states with an avenue for further granularity.

Certain regions have specifically stressed more than others the importance of contextualization – that is, putting an emphasis on regional contexts and realities to inform and shape policy and regulatory responses. This sentiment was particularly prominent among African states, as they consider contextualization as a *sine qua non* condition to all governance approaches and solutions to promote responsible AI in the military domain. Contextualization is indeed key to considering local technological, geopolitical and security realities – all of which require dialogue at the regional level in order to capture their breadth and depth.

There are a number of ways and forms in which such regional dialogue can be conducted. Regional organizations and processes constitute one key platform to enable the convening of such a dialogue, eventual alignment (e.g., through a regional-level policy or strategy document) and the subsequent establishment of enforcement mechanisms. Members of the North Atlantic Treaty Organization (NATO) value, for instance, the alliance's efforts to enable coordinated and concerted approaches to AI development and innovation: beyond dialogue within the alliance, initiatives and frameworks such as the Defence Innovation Accelerator for the North Atlantic (DIANA) and NATO's AI Strategy are considered as key influencing factors. The latter, for example, places key ethical principles (e.g., traceability and explainability) high on member states' political agenda. To go even further, during the consultations certain NATO members expressed an interest in establishing a regulator at the organizational level, with a supervisory board or authority supported by select experts, including from academia and the industry. This body would enforce legal and ethical requirements in addition to conducting independent investigations, cognizant of the sensitivities of military operations.

In parallel, while African states generally acknowledge the diversity of viewpoints within and across the region, the local context is such that there is a shared desire for a united regional front. It is therefore a priority to arrive at a collective vision on responsible AI in the military domain that fosters regional cooperation, mutual trust, shared understandings, and the meaningful exchange of information, best practices and resources. States generally align their national approaches and policies to the efforts done at the regional level; at the time, many Africa states were awaiting the formal adoption of the African Union's Continental AI Strategy, which would then be integrate into their national policies and legislation.[3] As such, states in Africa share the desire to raise awareness, at the international level, of this unique approach that promotes unity in diversity, all the while respecting states' national sovereignty.

[3] At the time of the consultations, on 6 June 2024, the African Union's Continental AI Strategy had not yet been adopted. The Strategy was adopted on 17 June 2024. African Union, "African Ministers Adopt Landmark Continental Artificial Intelligence Strategy, African Digital Compact to drive Africa's Development and Inclusive Growth", Press release, 17 June 2024, https://au.int/en/pressreleases/20240617/african-ministers-adopt-landmark-continental-artificial-intelligence-strategy.

In addition to dialogue within institutions (e.g., within NATO or the African Union, as discussed above), there is an appetite for thematic discussions conducted at the regional level. Reaching beyond the policy, legal and ethical realms, there is, for instance, a desire to gain granularity at the technical level. For example, a number of African representatives expressed support for thematic discussions and working groups on such issues as the black box, effective integration and interoperability, computing power and data. These could convene technical representatives from states within the region to inform and consolidate policy recommendations and operationalization pathways. With regards to data, for instance, the establishment of a dedicated regional and technical working group to exchange and dissect specific aspects (e.g., data availability and quality in each state, data-protection frameworks, as well as data-sharing agreements) will be important to foster responsible practices surrounding AI in the military domain. Focused discussions on hardware and computing power will also be key to promoting equitable and safe technology transfers while curbing risks of proliferation and misuse.

Finally, joint investment frameworks, whether at the regional or even bilateral level, can also constitute a key avenue for meaningful dialogue and cooperation. States in the Gulf have a cooperation and joint investment programme into AI technologies, including the military domain, which constitutes one such example. Additionally, interstate dialogue, confidence-building measures and agreements at the intermediary levels (e.g., between ministries of foreign affairs) must also be explored, especially to circumvent political blockages at the higher level).

4.6 The Importance of Capacity-Building Efforts

Beyond dialogue, states generally emphasize the need to invest in capacity-building, technical literacy and education, at both the regional and national levels. States are of the view that high levels of investments in AI must, necessarily, be coupled with high levels of investments in education to train, preserve and maintain the human capital needed to foster responsible practices surrounding the development, deployment and use of AI in the military and wider security domains.

Beyond the need to foster and retain AI talent, training is generally perceived as key in the context of capacity-building initiatives and, ultimately, to foster the responsible development, deployment and use of AI in the military and security domains. End-users, supervisors, operators, commanders, and even those procuring and evaluating those systems must go through extensive training that enables them to leverage the opportunities that these technologies have to offer and, conversely, exercise control when such need arises. Furthermore, training is considered as key to effective human–machine teaming and collaboration and, ultimately, responsibility over the deployment and use of AI systems. As such, a number of states across regions have stressed the importance of mandatory and robust training of armed forces personnel to ensure responsible practices surrounding AI in the military domain.

To this end, platforms, efforts and regular convenings to facilitate information-sharing and the exchange of best practices will be key. In addition, investment in talent and public education must also be prioritized to enable mutual learning and cross-pollination between the civilian and public sectors, ultimately paving the way for robust, resilient and responsible AI systems. Capacity-building needs, however, differ across regions and actors, and local contexts vary. As such, capacity-building efforts and initiatives must be adapted accordingly in order to ensure effective knowledge transfer and information-sharing.

States are not the only target audiences. In fact, states have expressed a desire to intensify capacity-building efforts for industry representatives with regards to international law obligations, especially as dual-use technologies are growing in prominence. Not only will these efforts help foster compliance by design, they will also raise awareness of the potential implications that international law may have for their activities (e.g., whether the development of systems to support war efforts could be considered as direct participation in hostilities and, thus, would make these entities lawful targets under international humanitarian law in armed conflict).

A number of capacity-building models were explored and discussed during the consultations (see Figure 2).

FIGURE 2.

Capacity-Building Models Explored and Discussed over Consultations

TRAINING BY INTERNATIONAL AND REGIONAL ORGANIZATIONS	INTERNAL CAPACITY-BUILDING	PUBLIC-PRIVATE PARTNERSHIPS
• In Latin America and the Caribbean, the Organization of American States (OAS) has conducted a number of capacity-building efforts in the cyber field over recent year, which many states have found to be extremely useful. There is appetite for such efforts to expand beyond the cyber realm and to encompass AI issues as well. • Targeted training for the diplomatic corps (as provided by UNIDIR), which certain states have expressly commended, was also found to be of value, especially when these thematic discussions are contextualized against the state of affairs at the multilateral level.	• There is appetite among most states for capacity-building exercises within and across government agencies. In addition to the "usual" ministries that address the issue of AI in the military domain (i.e., ministries of foreign affairs and of defence), internal capacity-building will not only ensure that this issue is addressed holistically and across the board. • Such efforts are also seen as necessary in the light of internal differences in capacity that often lead to diverging, or even competing, approaches across ministries and state agencies.	• Public–private partnerships have also been explored as a means to establish capacity-building and information-exchange mechanisms between governments and the private sector, thus enabling cross-pollination. • On a less institutionalized level, a number of states from South East Europe, the Middle East, the South Caucasus and Central Asia also expressed interest in the conduct of joint multi-stakeholder tabletop exercises, ultimately paving the way for cross-sectoral collaboration.

5. Key Policy Priorities: Points of Divergence

Five points of divergence surfaced during the regional consultations:

- Unique local contexts and realities
- Varying regulatory approaches and legal traditions
- Prioritization and risk perception
- Resource availability and allocation
- Desired endgame

These divergences, however, are not harmful in and of themselves to the possibility of an aligned international approach to responsible AI in the military and wider security domains. At times, these divergences may even be desirable, or even necessary.

A number of divergences have emerged across and within regions. However, it is important to note that these variations in approaches are not inherently harmful to an aligned international approach to responsible AI in the military and wider security domains. At times, these variations and nuances are, in fact, desirable or even necessary to account for regional contexts and realities. Most of the time, as long as there is effective coordination at the international level and mutual understanding of states' and regions' different approaches and rationales (i.e., drivers and underlying reasons for certain policy directions), the points below will not necessarily lead to fragmentation. Instead, they will create a reinforced and inclusive policy and regulatory landscape that is united in diversity, ultimately enabling and promoting responsible AI in the military and wider security domains.

5.1 Unique Local Contexts and Realities

The first primary area of divergence relates to the unique local contexts, realities and strategic priorities across and within regions. A discussion on these variations across regions can indeed help states understand the rationale, motivations and overall background that may influence the approach of other states, or even a whole region, to responsible AI in the military domain: this mutual understanding can then help provide for an environment conducive to trust and productive dialogue.

In Asia, maritime security is a particularly central (and contentious) subject. Due in particular to the geopolitical landscape of the region, states are particularly keen to leverage the opportunities that AI has to offer to bolster their maritime borders and their security and sovereignty at sea. In addition, the unique geographical traits of the maritime environment in the region are such that the integration of AI and autonomous functions in military assets is of high interest (e.g., because communication links are difficult, if not impossible, to establish at sea, especially in the high seas). It is also partly due to these unique geographical traits that states are particularly insisting on the conduct of clear and robust acceptance tests, especially for off-the-shelves capabilities local procured from abroad: ensuring the system's ability to operate and adapt, cognizant of local environments, constitutes a paramount step to mitigate risks of unintended consequences upon its deployment and use.

Similarly, states in South East Europe, the Middle East, the South Caucasus and Central Asia often have to grapple with cluttered and inaccessible environments – this time on land, and in particular in the context of military operations against non-state armed groups. As such, a number of states and regional experts have discussed the prominence of UAVs in the region, as attested through the widespread provision of such capabilities, including with autonomous functions, by either Türkiye, Iran or China, and their use. In addition, the conduct of future joint military operations within the region (e.g., as part of NATO for members of the alliance) also raises a number of questions and concerns with regards to the interoperability of AI technologies across states, but also with legacy systems.

In parallel, the African military theatre is uniquely characterized by an array of issues and considerations, including language barriers, ethnics tensions, as well as competition for resources. The Geneva Academy's Rule of Law in Armed Conflicts (RULAC) project estimates that there are currently over 30 non-international armed conflicts in the continent, which attests to the growing number of NSAGs operating across the region.[4] Coupled with the rapid technological growth that the continent is facing, as well as regional regulatory frameworks (e.g., the African Charter on Human and People's Rights) and legal traditions, these unique traits attest to the need for a continued meaningful regional dialogue. This would indeed allow for the promotion and implementation of norms and behaviour for responsible AI in the military and wider security domains at the regional (and, subsequently, national) level, infused with local contexts and realities, allowing for their effective operationalization and enforcement.

In Latin America and the Caribbean, there is the widespread and general view that, because the security landscape in the region focuses more on countering organized crime and less on interstate conflicts, AI applications for national security (e.g., through law enforcement and border security) are critical to the region – and more so than in the military domain. The deployment and use of AI technologies can even, at times, be perceived as compensating for capacity issues faced by smaller states in the absence of the talent and human resources at the disposal of armed forces and national security agencies (e.g., to increase surveillance and cyber capabilities). Additionally, one notable case is that of Costa Rica, which uniquely does not have an army and, thus, the priority lies in fostering peaceful AI applications. As such, the region's approach to its concerns about the development, deployment and use of AI technologies in security and defence will be different from that of other regions. These concerns include accessibility and proliferation issues; governance of dual-use technologies; as well as the social and economic impacts of AI proliferation and use by NSAGs.

[4] Geneva Academy, RULAC, https://www.rulac.org.

5.2 Varying Regulatory Approaches and Legal Traditions

States generally agree on the need to align AI development, deployment and use with international norms and principles, societal values, legal requirements and ethical considerations. Compliance with international law features prominently in states' national approaches to promoting responsible AI in the military domain; and there is a general acknowledgment of the importance of an environment conducive to compliance. This is reflected in states' positions, national strategy documents and respective policy frameworks. Their development, adoption, implementation and review are indeed considered, by most states, as key not only to establish their approach and interpretation, but also to upholding international and regional peace and security, international law, as well as ethical values such as transparency, accountability and humanity.

However, while states agree on their applicability (and their need to be prioritized), the interpretation and application of international law and ethics are done in such a way that regional, national and local (i.e., subnational) contexts will inevitably lead to variations in approaches, interpretation and prioritization to international law – even if these regions are facing similar concerns. One stark example of such variations pertains to the African; South East European, Middle Eastern, South Caucasus and Central Asian; and Latin American and Caribbean regions. These three regions have consistently expressed concern about NSAGs, a topic that has inevitably been discussed in all consultations. Nevertheless, their respective approaches to international law and main concerns present variations (see Box 3). While they may not, in themselves, necessarily be harmful, knowledge of those nuances can be useful in better understanding these regions' individual approaches to the same question.

BOX 3.

One Issue, Three Regions, Three Approaches: The Question of Non-State Armed Groups

AFRICA

In Africa, there is a strong emphasis on upholding human rights, particularly those enshrined in regional frameworks such as the African Charter on Human and Peoples' Rights. In fact, while it is acknowledged that international humanitarian law applies in times of armed conflict, there is also the recognition that international human rights law also applies in peace time (e.g., for law enforcement and counter-terrorism operations). As such, it generally confers more protection on individuals than does international humanitarian law through a host of rights, including the right to life and protection against discrimination. Additionally, the African Charter on Human and People's Rights, in its Article 23(1), provides for the right to national and international peace and security – a protection specifically conferred from post-colonial experiences.

These legal traditions within the region must indeed be contextualized against historical considerations and wider societal implications while scrutinizing the development, deployment and use of AI in the military domain – such as the impact on economic and social rights and trade. This wider context must also take into account states' adjacent technological concerns and policy priorities, which may have direct and indirect implications for AI governance pathways. These include digital connectivity and inclusion, the promotion of digital public goods, the need for an effective architecture for digital cooperation, as well as the development and maintenance of resilient data infrastructures.

SOUTH EAST EUROPE, THE MIDDLE EAST, THE SOUTH CAUCASUS AND CENTRAL ASIA

In parallel, states in South East Europe, the Middle East, the South Caucasus and Central Asia have shared concerns with regards to the complex implementation and enforcement of international law, as non-state armed groups also seek to adopt and use AI-enabled technologies. While it is clear that international humanitarian law would apply to situations meeting the required threshold for non-international armed conflicts, states are particularly concerned with these groups' liability and the ability to hold them accountable. This will depend to a great extent on a number of factors, including the political will of parties involved, the resources available, as well as accessibility to conflict areas for investigation and evidence collection.

These issues, albeit predating AI and applying more generally in contemporary warfare, are further intensified as non-state armed groups acquire, deploy and use AI technologies in their operations. Furthermore, there is also the acknowledgment that, below the threshold of non-international armed conflict, international human rights law applies along with national laws. This is particularly important in the light of the growing prominence of "grey zone" conflicts – those that occur neither in peacetime nor in active conflict (e.g., cyber operations). This issue also featured prominently during the consultations held in Singapore. As such, further research on the applicability, application and subsequent enforcement of international law in various conflict settings and shedding light on accountability for "grey zone" operations was seen as a priority by states in South East Europe, the Middle East, the South Caucasus and Central Asia, in addition to considering compensation mechanisms for civilian victims.

LATIN AMERICA AND THE CARIBBEAN

In Latin America and the Caribbean, there is a general perception that, in the absence of interstate conflicts, the military paradigm is less prominent than that of law enforcement or, at least, must be complemented by the latter. In fact, countering organized and transnational crimes constitutes the main priority of most, if not all states, within the region. As such, for states in the region, compliance with international humanitarian law may not be as much of a concern compared to enforcing international human rights law, the applicability of which extends beyond armed conflicts onto peacetime. In fact, outside armed conflict – and thus including law enforcement operations – a series of human rights are relevant with regards to the development, deployment and use of AI for security applications.

The most cited and studied right in this context is the right to life. This is enshrined in many human rights treaties, including the Universal Declaration of Human Rights; the International Covenant on Civil and Political Rights (ICCPR); as well as the American Convention on Human Rights. In many of these treaties, what is prohibited is a deprivation of life that is “arbitrary”, meaning that it complies with neither international rules and standards pertaining to the right to life nor domestic law. Compliance with the right to life applies as well in the context of AI use, both in conflict (subject to its interplay with international humanitarian law) and outside conflict (e.g., for law enforcement); this assertion increasingly features both at the national level from states’ positions and internationally.

For example, the United Nations Human Rights Committee issued, in 2019, a General Comment on the Right to Life, which asserts that “States parties engaged in the deployment, use, sale or purchase of existing weapons and in the study, development, acquisition or adoption of weapons, and means or methods of warfare, must always consider their impact on the right to life.”[5] The report then raises the example of autonomous weapons systems, which the Committee has determined should not be developed and deployed, either in times of war or in peacetime, unless it has been established that their use is consistent and conforms with the right to life (i.e., the development and use of these systems would not lead to the arbitrary deprivation of the right to life), along with other relevant norms of international law.

[5] Human Rights Committee, “Article 36: Right to life”, General Comment no. 36, CCPR/C/GC/36, 3 September 2019, https://undocs.org/CCPR/C/GC/36, paragraph 65.

These differences must not be seen as synonymous to a lack of shared understanding, at the higher level, of the importance and applicability of international law. Compliance with international law has consistently emerged as a key priority for all regions, with the desire by states to ensure its respect across the life cycle of the technologies in question and to disseminate its key principles among all. Variations in interpretation, however, merit further attention and study to ensure coordination and alignment – ultimately promoting, not undermining, international law.

Beyond legal traditions, variations can also be observed in the policy landscape: states within and across regions tend to be at different stages in the development, adoption and implementation of AI policy or strategy documents. Some already have in place established frameworks and may even be in the midst of review of their strategy, while others are still in the process of developing one. At times, there is evidence that instruments and positions adopted at the regional or even international level trickle down to influence national approaches. In Africa, for example, many states in the process of developing a national strategy on AI in security and defence were waiting for the African Union's Continental Strategy on AI, which they will then integrate into their national policies and legislation. Similarly, a number of NATO member states have acknowledged the influence exercised by the alliance's Strategy on AI, which plays an important role in shaping their approach to the governance of AI in the military domain.

5.3 Prioritization and Risk Perceptions

While states are generally aligned on the perceived opportunities and risks arising from the development, deployment and use of AI in the military domain, their prioritization tends to differ among states. Some may confer more value on legal compliance and ethics, with an emphasis on international law, human dignity and fairness. Others will prioritize national security and ensuring safe and secure environments; meanwhile, reducing the risks of civilian casualties may be central to some states. Similarly, a host of risks have been shared across regions, from concerns of proliferation to NSAGs to that of inadvertent (nuclear) escalation. Yet, their perception (e.g., perceived weight and scope), interpretation and prioritization differ from one another. Within Africa, for instance, some would prioritize the (existential) risks stemming from integrating AI technologies into nuclear command, control and communications systems, while others see it as less relevant for the region. In Asia, some states prioritize risks surrounding the misuse of these technologies and their appropriation by and proliferation into "malicious" actors. Others focus more on the humanitarian impact and risks stemming from these technologies. Risks associated with strategic stability, nuclear affairs and regional stability are also among those raised as a priority by some. Echoing the points made in Subsection 5.1, variations in local contexts and realities constitute a decisive factor in states' policy approaches; their opportunities and risks perception and prioritization are not spared from this.

Interestingly, while these differences can, at times, be a source of friction and can stand in the way of consensus (or, at least, alignment on certain issues), it is also important to note that these differences are not necessarily harmful in and of themselves. For African states, these differences do not necessarily affect the general desire for a united continental front on the governance of AI in the military domain.

5.4 Resource Availability and Allocation

While states share the desire to promote responsible AI in the military domain, the availability and allocation of resources dedicated to such efforts differ from one state to another. These differences in resource allocation are reflected in the approaches adopted to promoting and implementing responsible practices surrounding AI in the military domain: some states may have dedicated teams exclusively for AI-related affairs, while other armed forces (if not most) may have personnel tasked with emerging technologies more generally – ranging from AI to the safeguarding of critical national infrastructure against cyber operations

Another example relates to capacity-building: while all states generally emphasize its importance, the extent to which dedicated resources and funding have been allocated to such efforts diverge from one state to another. Some may have formalized institutions for education and capacity-building (e.g., with established centres of excellence); such efforts, however, require significant resources that not all states may necessarily be able to allocate and prioritize yet. Nevertheless, there is a general acknowledgment, in Africa at least, that institutional capacity and regional cooperation will be key to mutual assistance and, ultimately, to foster localized AI solutions to shift away from overdependency on Western technologies.

5.5 Desired Endgame

While all states present at the regional consultations agreed on the importance of responsible development, deployment and use of military AI, there is disagreement as to what processes and outcome are needed to uphold and ensure such responsible behaviour. For instance:

- Some are of the view that a new international treaty dedicated to AI in the military domain is needed; others were more keen to focus on the issue of lethal autonomous weapons systems first, in the light of the work done on this area for over a decade within the framework of the CCW. In parallel, a number of states were of the view that these discussions must not be mutually exclusive, and governance deliberations on AI in the military domain must be conducted in concert with efforts within the CCW framework; other states were of the view that applicable international law is sufficient.
- Specifically on the issue of lethal autonomous weapons systems, some were of the view that an additional protocol to the CCW would be most valuable; others criticized the limited membership of the CCW regime and highlighted the need for a wider scope beyond lethal autonomous weapons systems.
- Some are of the view that taking the conversation outside the United Nations is of value and offers flexibility and inclusivity, while others insist on maintaining the discussions within the United Nations in order to not jeopardize the work of the organization.

Beyond the *what*, differences have also emerged with regards to the *why* – that is, the underlying reasons and rationale behind their desired end-result – and the *how*. For instance, in Asia, some were of the view that certain intended uses must be prohibited due to the risks of their humanitarian impact (e.g., those with highly destructive capabilities), while those simply with the intention to "neutralize" the adversary may be allowed and even, at times, prioritized. On the other hand, others were of the view that this very much remains on a case-by-case basis, and as such, potential opportunities that AI technologies have to offer must not be jeopardized by blanket prohibitions.

In South East Europe, the Middle East, the South Caucasus and Central Asia, there is a shared sentiment that formalized processes and frameworks are needed to regulate or, at least, govern and promote responsible AI in the military domain. In fact, states in the region generally acknowledge the need for a certain form of regulation surrounding the development, deployment and use of AI in the military domain. As such, while the development of an international treaty is an option, other means such as a code of conduct were perceived as being useful too. Such a framework would enable states to establish shared principles, akin to those of NATO's AI Strategy – although in the absence of a shared alliance underpinning such a document, its development, adoption and subsequent implementation will require tremendous effort, investment and political will. Key principles may include accountability and responsibility, including for unintended consequences. However, it is important to think beyond the substance during the processes needed to lead up to the adoption of such shared principles. As such, beyond international processes and frameworks, regional means and military alliances must also be explored as a way of formalizing shared principles and fostering cooperation in this space – including for capacity-building. An effective combination of both international and localized approaches will thus ensure complementarity and an ecosystem of processes and frameworks conducive to responsible AI in the military domain.

States in Africa generally prioritize the development, adoption and implementation of a regulatory instrument at the regional level – perhaps even more so than at the international level. In the light of Africa's unique policy landscape and the desire for unity within the continent, regional processes would be conducive to an agreement on unique elements that may not be attainable at the international level. In fact, there is the view that, while international frameworks are important, they must remain general, with regional tools providing more room for granularity and specificity. The national level will then subsequently be the most specific with regards to the implementation of these international and regional policies. States must thus also prioritize the development, adoption, implementation and review of a national strategy document on AI in security and defence.

Latin America and the Caribbean is generally divided into two blocs, each approaching the question of AI in the military and wider security domains slightly differently – albeit complementarily. Some states, including Brazil, Chile and Mexico, prioritize an approach centred around legal requirements and technical solutions; other states, including Costa Rica, prioritize an approach centred around the political process and norm-setting as an outcome. There is also a third group of states that, despite sharing the view that regulations are much needed, have voiced concerns of prohibitions eventually standing in the way of their combating adversaries – especially NSAGs that, themselves, may have adopted and may be using AI to support their operations. These approaches, however, are not mutually exclusive and certain states' policies may be a hybrid of two or more of these.

In Europe and North America, states are divided in their preference for hard versus soft law as the way ahead for the governance of AI in the military domain. At one end of the spectrum, states that prioritize hard law draw inspiration from the European Union AI Act and its risk-based approach. States at the other end of the spectrum are generally wary of overregulation and express concerns about hampered innovation. The idea of an international treaty, or even any instrument or document that may pave the way for such treaty, is this not desirable for these states.

6. Conclusion and the Way Ahead

Responsible AI in the military and wider security domains, and efforts to implement and operationalize it, is an area of priority for states across all regions. Yet, the international policy landscape and approaches surrounding AI in the military domain are far from uniform. Across and within regions, there are a number of convergences and divergences. The former are generally not even uniform and may be nuanced across and within regions. The divergences can also vary, whether examined at the cross- or intra-regional level.

These nuances are (partly) explained by the variations across regions in local contexts, realities, legal traditions, culture, strategic priorities, security landscapes and socio-economic situations. These multifaceted and complex elements are subsequently, more often than not, reflected in the ways in which each region and each state grapple with the development, deployment and use of AI in the military and wider security domains. As stressed throughout this report, however, these differences are not, in and of themselves, necessarily harmful to international alignment and cooperation. Moreover, states acknowledge the importance, in most circumstances, of such cooperation. Processes, discussions and frameworks at the international, regional and national levels must not be seen as mutually exclusive and in competition but, rather, as complementarity in the kaleidoscope of military AI governance.

As such, the conduct of regional consultations, joining information exchange on policies and national viewpoints on the one hand with discussions framed by expert intervention on the other hand, was not only useful to ensure inclusivity in the REAIM process. These consultations also paved the way for further reflections, at times shared across regions, at times unique to certain regions. As such, the food for thought to consider for the (future) governance of responsible AI in the military domain includes the following elements.

6.1 Prioritize Capacity-Building and Information Exchange

There is a shared sentiment, across regions, that capacity-building is a key priority. As such, the establishment of regularly convened, formalized processes and frameworks for information sharing and the exchange of knowledge and best practices, including from the civilian domain, will be important. Beyond the policy, legal and ethical realms, there is also a desire to gain granularity at the technical level: Close examination of specific technical issues (e.g., the black box, effective integration and interoperability, computing power and data) will be key for the formulation of concrete policy recommendations and operationalization pathways. As such, the establishment of standardized and benchmarked training (e.g., through certification) will also be important to ensure coordination and alignment within and across regions.

6.2 Foster the Implementation and Operationalization of Responsible AI through Engagement

States have expressed the desire to move beyond norms, principles and political commitments: there is indeed appetite for initiatives that help bridge the gap between the technical and policy communities for operationalization, implementation and execution. To this end, establishing a mechanism to enable collaboration between these communities will be key. For example, drawing on lessons from cyber, a model similar to the cyber policy community's engagement with computer emergency response teams (CERTs) could be adopted. The challenge will be to keep the discussions grounded, evidence-based and accessible to all communities and stakeholders involved.

In addition to norms and principles, there is an appetite from states to unpack technical approaches and solutions to uphold international law and ethical requirements. This includes the need to uphold transparency and accountability through evaluations and the auditability of systems. Smaller states in the Caribbean have consented to the intervention of major states in supporting efforts at combatting crime. Such interventions may not always be accompanied by use of force; however, they are presented as often using AI technologies. Yet, this is frequently done in a very opaque manner, resulting in power imbalances and the inability to uphold accountability and responsibility over the use of these technologies by the intervening states. Transparency standards for transnational operations would, in this sense, be useful – especially for smaller states consenting to foreign intervention due to limitations in resources and capabilities. As such, the exchange of best practices and, more generally, the establishment of international and regional cooperation mechanisms that are multidisciplinary and inclusive will be key.

6.3 Ensure the Complementarity of Processes

The parallel processes and efforts in this space range from REAIM to the United Nations High-Level Advisory Board (HLAB) on AI, the United Nations Security Council discussions, the US Political Declaration, the CCW Group of Governmental Experts on Lethal Autonomous Weapons Systems, discussions within the United Nations General Assembly's First Committee, initiatives within NATO, as well as the Freetown and Belén Communiqués. Given this constellation, states have expressed the desire to shed clarity on the respective mandates, roles and responsibilities of each process and forum at the international and regional levels, and to ensure that they complement one another instead of encouraging competition. In addition, there is an appetite to explore possibilities for cross-pollination and the sharing of lessons learned from other arms control processes and security fields (e.g., the Montreux Document as well as the work done on responsible behaviour in cyber and in space).[6] To this end, coordination and clear communication will be key.

[6] The Montreux Document on Private Military and Security Companies provides for the pertinent international legal obligations and good practices for states related to the operations of private military and security companies during armed conflict. While not legally binding, the Montreux Document intends to provide clarifications as to how certain well-established rules of international law apply to states in their relations with private military and security companies, and their operation during armed conflict – particularly under international humanitarian law and international human rights law. See: ICRC, "The Montreux Document on Private Military and Security Companies" (2009), https://www.icrc.org/en/publication/0996-montreux-document-private-military-and-security-companies.

6.4 Intensify Research Efforts and Reflections on Underexplored Areas

In addition to the recommendations above, a number of themes have emerged as underexplored in international discussions surrounding AI in the military and wider security domains. Among these are data governance; dual-use technologies; interoperability; computing power; the traceability and transparency of systems; as well as the different dimensions of security in relation to AI development, deployment and use in the military domain (e.g., human security, environmental security and economic security). While policy deliberations at the higher level are necessary, these discussions must be supported through further research efforts and examination of specific points that, unpacked, will ensure that policy outcomes are well-informed, robust and comprehensive.

6.5 Reflect on AI in the "Military Domain", its Boundaries and Interplay with Wider Security Applications

While the regional consultations were all conducted around the theme of "AI in the military domain" (with the Latin American and Caribbean segment also including wider security applications), questions arise with regards to the scope of this domain and the extent to which its boundaries are clearly delineated. Efforts to combat organized crime are considered to be an integral part of the military domain in certain regions, while others would categorize such operations as law enforcement. Depending on the region, there would thus be differences in the categorization of AI applications for combatting crime (i.e., as to whether or not they would fall within or outside the scope of "AI in the military domain"). Further reflections on what "AI in the military" domain consists of, its boundaries, its interplay with wider security applications, and which applications would fall within the scope of future governance frameworks would indeed be useful to keep the discussions focused and targeted.

UNIDIR

@unidir
/unidir
/un_disarmresearch
/unidirgeneva
/unidir

Palais des Nations
1211 Geneva, Switzerland

WWW.UNIDIR.ORG