AISE MARKERS SERIES

# Benchmark III: Use cases

## Insights from the Global Conference on AI, Security and Ethics 2025

YASMIN AFINA · JAN HENDRIK MANNSPERGER

## 1. Introduction

### 1.1. Context

On 27–28 March 2025, UNIDIR organized its inaugural Global Conference on Artificial Intelligence, Security and Ethics (AISE25), hosted in the Palais des Nations, Geneva. Led by UNIDIR's Security and Technology Programme, the conference provided an agile response to rapid advances in artificial intelligence (AI), which have put this technology at the forefront of today's global policy discussions. AISE25 was held as policymakers and regulators worldwide increasingly recognized the urgency of developing shared understandings, norms and regulations that can transcend national borders and individual interests, including in the context of peace and security.

The conference sought to provide a unique forum for engagement between the multilateral ecosystem and the wider multi-stakeholder community, including academic experts, civil society organizations, industry representatives and research laboratories interested in the governance of AI in peace and security. By jointly analysing and addressing the complex implications of AI for national, regional and global security and resilience, AISE25 enriched dialogue between participants and exposed them to the latest research in the field. This opportunity to exchange and consolidate views on AI in the military domain was timely, with the United Nations General Assembly having requested the views of Member States and other stakeholders as a means of informing discussions during its Eightieth Session, in September 2025 – the deadline for which came just weeks after the conference.[1]

The conference programme was designed to build on the work undertaken as part of UNIDIR's Roundtable for AI, Security and Ethics (RAISE), a multi-year project on multi-stakeholder engagement in this space launched with the support of Microsoft.[2] Specifically, the conference's agenda was primarily organized around the six priority themes identified at the inaugural edition of RAISE, which took place in Bellagio, Italy, in March 2024:[3]

1. Knowledge and capacity-building
2. Trust-building
3. The human element
4. Data practices
5. Life cycle management
6. Addressing destabilization

Combining a series of panel discussions, presentations in the form of thematic deep-dives and lightning talks, as well as a poster exhibition, AISE25 provided a timely platform, open to all, to jointly consider and elaborate on each of these six priority themes while promoting meaningful dialogue and cooperation.

Ahead of the second edition of the AISE, in June 2026, a series of three reports – the first of the AISE Markers series – takes stock of the key takeaways from AISE25 to provide an initial basis and scaffolding for AISE26. By acting as a bridge between editions of AISE, the AISE Markers series will ensure that each conference is built on solid ground and constitutes a natural evolution from the discussions held in the previous conference.

This third report gives a structured account of where use cases – both governance and technological – stood at the time of AISE25. It provides an overview of discussions surrounding the human element in practice, surveys regional perspectives, zooms in on AI for peace operations and humanitarian applications, and then offers cross-cutting observations that may serve as a baseline for AISE26. Accompanying reports cover the conversation on governance and the state of technology.

## 1.2. Why use cases matter for governance

High-level governance principles become testable only when grounded in specific contexts and use cases. In recognition that the question is no longer whether AI will be used in military and security contexts, but rather how, AISE25 consistently revealed a gap between,

1 General Assembly, resolution 79/239, "Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security", 24 December 2024, https://docs.un.org/A/RES/79/239.

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

3 Y. Afina and G. Persi Paoli, *Governance of Artificial Intelligence in the Military Domain: A Multi-Stakeholder Perspective on Priority Areas* (Geneva: UNIDIR, 2024), https://unidir.org/publication/governance-of-artificial-intelligence-in-the-military-domain-a-multi-stakeholder-perspective-on-priority-areas/.

on the one hand, high-level normative frameworks and principles and, on the other, the concrete decisions and uses faced by practitioners, operators and users of these technologies.

This report considers use cases around two dimensions: in addition to technological use cases given through concrete examples of operational application, it also examines governance use cases. Clarity and specificity are key to shaping not only the direction and parameters of on-going and future policy discussions, but they are also requirements for specific compliance measures through both technical and institutional design. This analysis of use cases also reveals variables that high-level principles alone cannot. These include, for instance, the extent to which regional diversity complexifies – or even challenges – universalist governance assumptions, as well as the role of empathy as a building block for governance.

# 2. The human element in practice

The human element sits at the core of many, if not most, policy deliberations on AI in defence and security. Through constructs and framings such as "meaningful human control", "human oversight" or "human judgment", states and the wider multi-stakeholder community have sought to clarify and establish what kind of human element could, or even should, be required in the deployment and use of military AI technologies. Yet, the discussions at AISE25 point to the conclusion that specificity should be the governance priority for on-going and future efforts, rather than terminological consensus as to what that human element means.

## 2.1. What international law currently says and does not say

The human element is often presented as a *sine qua non* condition for compliance with international law. Whether through the exercise of a certain degree of human supervision, oversight or control, and independently of the term and construct selected, this premise lies at the foundation of the views and positions of many states, organizations and experts. From the opposite perspective, there is also the argument that there is "no specific and precise obligation for [the] human element" in current treaty and customary international humanitarian law (IHL).[4] There is also an argument that, under some circumstances, the human element may constitute more of a liability than an asset, with many violations of IHL resulting from human acts and decisions. Independently of whether and how AI systems are used in defence and security contexts, legal responsibility will remain with states and individuals: they hold responsibility over the decision to procure, deploy and use AI systems, including through the conduct of the appropriate risk assessments, due diligence and other measures necessary for compliance. This leads to the view that, ultimately, discussions on the application of IHL are not necessarily about what the law says and provides about the human element, but rather what the human element means for IHL compliance.

These discussions lead to the view that, ultimately, discussions on the application of IHL are not necessarily about what the law says and provides about the human element; rather

4 As noted by Yasmin Afina in day 2's panel on the human element.

they are about what the human element means for IHL compliance. Given states' existing views, national perspectives and interpretations on the application of international law, this analysis of the role of the human element for compliance must consider the technology's entire life cycle: while most on-going discussions focus on deployment and use, there is a need to examine requirements in the pre-deployment stages of these technologies, as well as the end of their life cycle.

## 2.2. Evolving views on red lines and limitations

Beyond international law, current ethical arguments bring forth a number of distinct and nuanced considerations for the development, deployment and use of AI in the context of defence and security. Whether it is with respect to the human element or specific red lines on use, the highly distinct views at AISE25 reveal fundamental tensions as to what should or should not be permissible under widely established standards. Some argue that "life-and-death situations must never be left to chance, code, or corporate interests", and "humans must always retain control over decision-making functions – guided by international law, human rights and universal ethical principles".[5] Accordingly, ethical principles do not necessarily provide for red lines for AI that might have lethal capabilities: this view is more nuanced, on the basis that the key would be the preservation of the "appropriate human intention" that may take other forms than "having a human as the final presser of the button".[6] This argument rests on the premise that human–machine interaction can take far more complex forms, particularly considering that the human element can also be multifaceted (e.g., through engagement, control, deliberation, intent or a combination of these).

The traditional ethics of technology provide another dimension to these reflections, treating the deployment of any technology not as a neutral act but as a form of ordering. In other words, deployment is a displacement of power that reshapes the institutional landscape, the distribution of risk and the range of choices subsequently available.[7]

In the context of AI in defence and security, this observation carries a concrete implication for governance: discussions on the ethical dimensions of these technologies, including on what constitutes permissible forms of lethal autonomy, must precede deployment, rather than follow it. The costs of retrofitting ethical and governance considerations onto systems already embedded in operational doctrine and procurement cycles are substantially higher than the costs of integrating them at the design stage.

## 2.3. A framework for actionable human engagement policy

The varied and evolving language surrounding the human element – from "meaningful human control" to "context-appropriate human judgment and control", "human oversight" or "human deliberation" – reflects not only the conceptual complexity of these questions but also, at times, deliberate political choices in how states and organizations aim to frame their positions.[8]

---

5 "UN Secretary General's message to the inaugural Global Conference on AI, Security and Ethics", UNIDIR, 27 March 2025, https://unidir.org/un-secretary-generals-message-inaugural-global-conference-ai-security-ethics/.

6 As noted by Jovana Davidovic (Peace Research Institute Oslo, PRIO) in a dedicated deep-dive on day 2: "On the Purpose of Human Engagement for AI-Enabled Weapons Systems".

7 As noted by Paolo Benanti (Pontifical Gregorian University) in day 2's panel on the human element, drawing on the work of political scientist Langdon Winner on the politics of technology.

8 Davidovic.

Yet, beyond the terminology lies a more fundamental challenge: governance frameworks that do not specify what human engagement is meant to achieve cannot generate meaningful compliance requirements. Three inputs are identified as indispensable for translating the principle of human engagement into policy-relevant requirements: clarity of purpose, of type and of unit of analysis.[9]

## Clarity of purpose

Why is human engagement required? Distinct purposes, guided by safety, attribution of responsibility, respect for human dignity and the preservation of institutional stability, have each been invoked in policy discourse – yet they are rarely distinguished from one another. A governance framework designed primarily to ensure that responsibility can be attributed after the fact may produce very different requirements from one designed around safety, and both may diverge from a framework oriented primarily towards the preservation of human dignity. Conflating these purposes generates governance instruments that are internally inconsistent and, in practice, unverifiable.

## Clarity of type

What kind of human engagement is required? Several distinct forms are present in the literature and in policy discussion: human control (in the sense of a human who can intervene in or halt a system), human judgment (the application of contextual reasoning to a specific decision), human deliberation (the intrinsic or instrumental value of a deliberative process involving human actors) and human intent (ensuring that the outputs of AI-enabled systems remain aligned with the intentions of the commander or operator who set them in motion). The shift away from "meaningful human control" and towards "context-appropriate human judgment and control" that is visible in recent policy documents reflects both substantive and political considerations: substantive in that judgment may more accurately capture what is required in certain operational contexts; political in that it may be more palatable to states resistant to formulations that could be read as constraining operational flexibility.[10]

## Clarity of unit of analysis

At what point in the system, or in the targeting process, should human engagement be required? Using the targeting cycle as a reference framework, AI tools are integrated across multiple stages, from the development of commander intent, through target development, validation, capabilities analysis and mission execution, to post-strike assessment. A requirement for human engagement that specifies neither the stage nor the form of that engagement cannot be operationalized by a military planner, a procurement officer or a legal adviser.[11]

Together, these three inputs provide a diagnostic framework: for any proposed governance measure that invokes the human element as a requirement, it should be possible to answer the questions of why, in what form and at what point. Without these answers, the measure remains aspirational, rather than operational.

---

9 Ibid.

10 Ibid.

11 Ibid. The targeting cycle was referenced as a framework for understanding the various stages at which AI tools may be integrated into military operations and to extrapolate what different types of human engagement could be required or, at least, be desirable.

# 3. Regional use cases

AI adoption varies significantly across regions, shaped by local contexts, technological access and structural conditions. Understanding these underlying factors is therefore essential to any meaningful analysis of current use cases. AISE25 provided local perspectives from regions around the world in order to identifying the current on-the-ground realities that cannot always be seen via outside observation.

## 3.1. Africa: Counter-insurgency, capacity-building and inclusion

The African continent's engagement with AI in defence and security encompasses a spectrum of applications, from surveillance systems and predictive analytics deployed in counter-insurgency operations to urban security tools aimed at addressing rising rates of violent crime. In the former category, AI tools (including drones, surveillance systems and predictive analytics platforms) have reportedly been deployed in the context of counter-insurgency efforts against non-state armed groups. These enable functions such as tracking communications, predicting potential incidents and analysing encrypted social media activity. In urban settings, facial recognition technology and predictive analytics are also reportedly being employed to identify patterns associated with criminal activity.[12]

While the technological landscape demonstrates a clear uptake on these technologies, the governance landscape presents a markedly different picture. Limited institutional frameworks and national capacity for AI ethics research combined with the relative absence of targeted research funding in this space mean that the governance infrastructure required to support responsible deployment lags significantly behind practice. At the regional level, this observation translates into a structural recommendation: international bodies and research institutions seeking to engage with African states on military AI governance would benefit from directing engagement with regional organizations, including the African Union and subregional bodies such as the Economic Community of West African States (ECOWAS) and the East African Community (EAC). This engagement would act as a means of developing strategic plans that can address the specific security contexts, financing needs and capacity-building requirements of African states.

Two considerations emerge as particularly salient in this context. The first is the inherently dual-use nature of many AI security applications: surveillance technologies initially developed for civilian purposes are increasingly repurposed for security operations, creating governance challenges that straddle the civilian–military boundary. The second, and arguably more fundamental, is the question of empowerment as distinct from inclusion: the objective for African states is not merely to participate in governance conversations led elsewhere, but to be equipped to shape them through the development of the knowledge, infrastructure and human capital required to make independent assessments of AI technologies and their implications for national and regional security.[13]

12 As noted by Baudouin Ngah Akoh (United Nations Development Programme) in day 2's lunch panel on regional perspectives on AI, security and ethics.

13 Akoh.

## 3.2. South East Asia: Strategic positioning and the AI competition crossroads

South East Asia occupies a distinctive position in the global AI landscape: geographically positioned at the intersection of the Pacific and Indian Oceans, it constitutes a region of significant strategic importance in which the implications of the broader technological competition between China and the United States are reported to be acutely felt.[14]

At the same time, the region is characterized by significant internal heterogeneity in terms of AI readiness and governance capacity. Global indices of AI development reveal a gap of over 60 ranks between the most and the least advanced South East Asian states, a differential that has direct implications for the design of governance frameworks expected to apply across the region. Notwithstanding this diversity, regional governance efforts are progressing: the Association of Southeast Asian Nations (ASEAN) Defence Ministers' Meeting has produced a joint statement on cooperation in the field of AI in defence, reflecting a commitment to deepen regional understanding through information exchange and the sharing of best practices and lessons learned.[15]

At the national level, states are developing frameworks for responsible military AI that reflect their own strategic priorities and legal environments. In this context, they have each developed of a set of national principles, encompassing responsibility, reliability, robustness and safety, and then translated them into regional engagement through co-hosted capacity-building consultations. This represents one model for how states at varying levels of AI readiness can contribute substantively to the building of governance norms. At the same time, the translation of such frameworks and commitments into consistent implementation across a region as diverse as South East Asia remains, by all accounts, a long-term undertaking. The structural dimension is also significant: South East Asia's position at the intersection of the United States–China technological competition means that the choices made by the region's states on AI procurement and governance carry geopolitical weight that extends beyond their immediate operational contexts.

## 3.3. Latin America: Emerging defence industry and the research gap

The backdrop to Latin America's engagement with AI in defence and security is distinct from those of other regions: the absence of significant interstate conflicts within the region, combined with the prominence of powerful non-state actors (including transnational organized crime networks, armed groups and those involved in trafficking), shapes both the threat environment and the AI applications most relevant to it. Drone-based monitoring of major ecosystems, AI-supported border management and predictive analytics for law enforcement constitute representative examples of the operational contexts in which AI is being developed or considered.[16]

---

14 As noted by Bagus Jatmiko (Indonesian Navy) in day 2's lunch panel on regional perspectives on AI, security and ethics.

15 As noted by Pak Shun Ng (Ministry of Defence, Singapore) in day 2's panel on knowledge and capacity-building, on Singapore's framework and the ASEAN Defence Ministers' Meeting joint statement. See ASEAN, "Joint Statement by the ASEAN Defence Ministers on Cooperation in the Field of Artificial Intelligence in the Defence Sector", 26 February 2025, https://asean.org/joint-statement-by-the-asean-defence-ministerial-on-cooperation-in-the-field-of-artificial-intelligence-in-the-defence-sector/.

16 As noted by Francisco Rodríguez (FLACSO Ecuador) in day 2's lunch panel on regional perspectives on AI, security and ethics.

A notable finding from regional research is that defence industry and military actors in Latin America do not frame their AI engagement as a process of catching up with technologically advanced states. Rather, they identify an emerging segment of the global military AI industry in which Latin American actors can participate on their own terms, developing capabilities suited to their operational contexts while maintaining that the exercise of human judgment should remain embedded in the command chain. This perspective suggests that governance frameworks designed exclusively around the priorities and concerns of early movers would not apply to the strategic and operational realities of a significant portion of the international community.

A structural challenge persists, however, in the form of restricted access to information. The confidentiality terms associated with many AI development and procurement agreements create significant barriers to independent research on what systems are being developed, how they operate and what risks they may pose. Without improved transparency and information-sharing mechanisms, the assessment of AI applications in this context will remain constrained. As such, frameworks that rely on voluntary disclosure and industry self-reporting alone may be inadequate to assessing AI in the security domain across diverse regional contexts. This subsequently emphasizes the fact that one-size-fits-all recommendations most often produce measures that are neither tailored nor responsive to the specific problems they purport to address.

## 3.4.South Asia: Border security and the misidentification risk

The border between India and China has been presented as one of the most operationally demanding environments for AI deployment in the security domain: remote and largely inaccessible terrain, contested territorial claims and a history of periodic escalation define the context in which AI-enabled surveillance and decision-support systems are reportedly being considered for application. Three categories of AI application are particularly relevant:

drone-based AI surveillance for monitoring remote and rugged terrain; facial-recognition technology for identifying potential encroachments or threats; and predictive analytics aimed at anticipating incidents on the basis of historical data and sensor inputs.[17]

Each of these applications raises governance challenges that are specific to this context, but each illustrates broader lessons about the conditions under which AI deployment requires particular caution. Misidentification – whether of own-force assets, civilian populations or non-threatening cross-border movements – carries the potential for negative consequences in an environment where the threshold for escalation may be comparatively low. Cybersecurity vulnerabilities in AI-integrated military systems present an additional dimension of risk: the same connectivity that enables AI-enhanced situational awareness also creates attack surfaces that both states and non-state actors may seek to exploit.

The establishment of a bilateral mechanism dedicated to AI-related issues at the border, covering both technical and legal dimensions, was proposed as a confidence-building measure and a potential stepping stone towards broader multilateral engagement.[18] Confidence-building measures modelled on existing examples from other bilateral security contexts could provide a practical template for managing AI-related incidents before they escalate.

## 3.5. China – United States dynamics and track-2 diplomacy as a use-case governance tool

The governance of military AI in the context of the relationship between China and the United States is widely recognized as a critical determinant of the broader international governance landscape: whether and how the world's two largest economies and most advanced AI powers can develop shared understandings and, eventually, agreed constraints on the military application of these technologies will shape the space available for multilateral governance efforts more broadly.

There is the assumption that geopolitical competition makes meaningful AI governance cooperation structurally impossible.[19] This assumption, while understandable in the light of current tensions, could arguably be counterproductive: it results, on many occasions, in self-fulfilling constraint, standing in the way of opportunities for cooperation that could in fact exist.

Track-2 diplomacy has demonstrated concrete governance value in this context. Academic dialogues involving researchers from both countries have contributed to tangible outcomes, including an agreement to maintain human control over nuclear weapon decisions – a commitment that emerged, at least in part, from the sustained engagement facilitated by informal dialogue processes.[20]

---

17 As noted by Naveen Kumar Samuel Kori (Centre for Public Policy and Good Governance) in day 2's lunch panel on regional perspectives on AI, security and ethics.

18 Kori.

19 As noted by Guangyu Qiao-Franco (Radboud University) in day 2's lunch panel on regional perspectives on AI, security and ethics. For more on the track-2 dialogue between Tsinghua University and the Brookings Institution, see Melanie W. Sisson et al., "Steps Toward AI Governance in the Military Domain", Brookings Institution, 12 November 2025, https://www.brookings.edu/articles/steps-toward-ai-governance-in-the-military-domain/.

20 The 2024 agreement to maintain human control over nuclear weapon decisions was noted by Guangyu Qiao-Franco as a concrete outcome of sustained United States–China track-2 dialogue. For more on the Xi–Biden agreement, see Jarrett Renshaw and Trevor Hunnicutt, "Biden, Xi Agree That Humans, Not AI, Should Control Nuclear Arms", Reuters, 17 November 2024, https://www.reuters.com/world/biden-xi-agreed-that-humans-not-ai-should-control-nuclear-weapons-white-house-2024-11-16/.

These dialogues also bring forth the most basic elements of communication as a governance need that is often overlooked: shared definitions or a common lexicon, in the form of agreed understandings of what terms such as "AI safety", "AI security" and "autonomy" can mean across languages and strategic cultures.

This work towards mutual, fundamental understandings is arguably not a secondary output of governance work; it is, in many respects, its precondition. The experience of sustained bilateral track-2 engagement confirms that dialogue, even under conditions of strategic rivalry, can produce common understandings that formal multilateral processes have not yet achieved.

## 3.6. Small states: The criminal justice AI governance gap

Small island developing states and states with limited AI development capacity have a unique position in the AI governance landscape: they are, in the majority of cases, purchasers and recipients, rather than developers, of AI-enabled security technologies. They are therefore largely dependent on the governance choices made by the developers and suppliers of these technologies, who operate in other jurisdictions. Yet these smaller states' stakes in how these technologies are governed are by no means smaller; in certain respects, their governance challenges are at times more acute.[21]

One dimension of this challenge that received insufficient attention in prior governance discussions concerns the interface between AI applications and the criminal justice system. As AI-enabled surveillance technologies proliferate from military and intelligence contexts into law enforcement – a trajectory anticipated by governance scholars and confirmed by operational practice – the evidentiary and legal framework questions multiply. Whether AI-generated evidence (including outputs from facial-recognition systems or AI-assisted surveillance platforms) would be admissible in domestic courts is a question that sits at the intersection of human rights law, constitutional law and evidence law. It cannot be resolved by governance frameworks designed for military operations alone, and it has direct implications for the rule of law in states where these technologies are deployed.

The prerequisite for meaningful national engagement with these questions is, in the first instance, political: decisions to develop or adopt AI in national security and defence contexts, and the governance frameworks that accompany them, require political authorization at the highest level of government. Capacity-building efforts that focus exclusively on technical or legal dimensions without due consideration for the political prerequisites for governance are liable to produce frameworks that exist on paper but lack the institutional grounding necessary for implementation. Regional cooperation mechanisms for criminal justice offer a potentially valuable model for building governance capacity in states that may still be at an early stage of engaging with these questions at the national level.[22]

## 3.7. What regional diversity reveals for governance

Taken together, the regional perspectives that surfaced at AISE25 point to several conclusions that have direct implications for the international governance of AI in security.

---

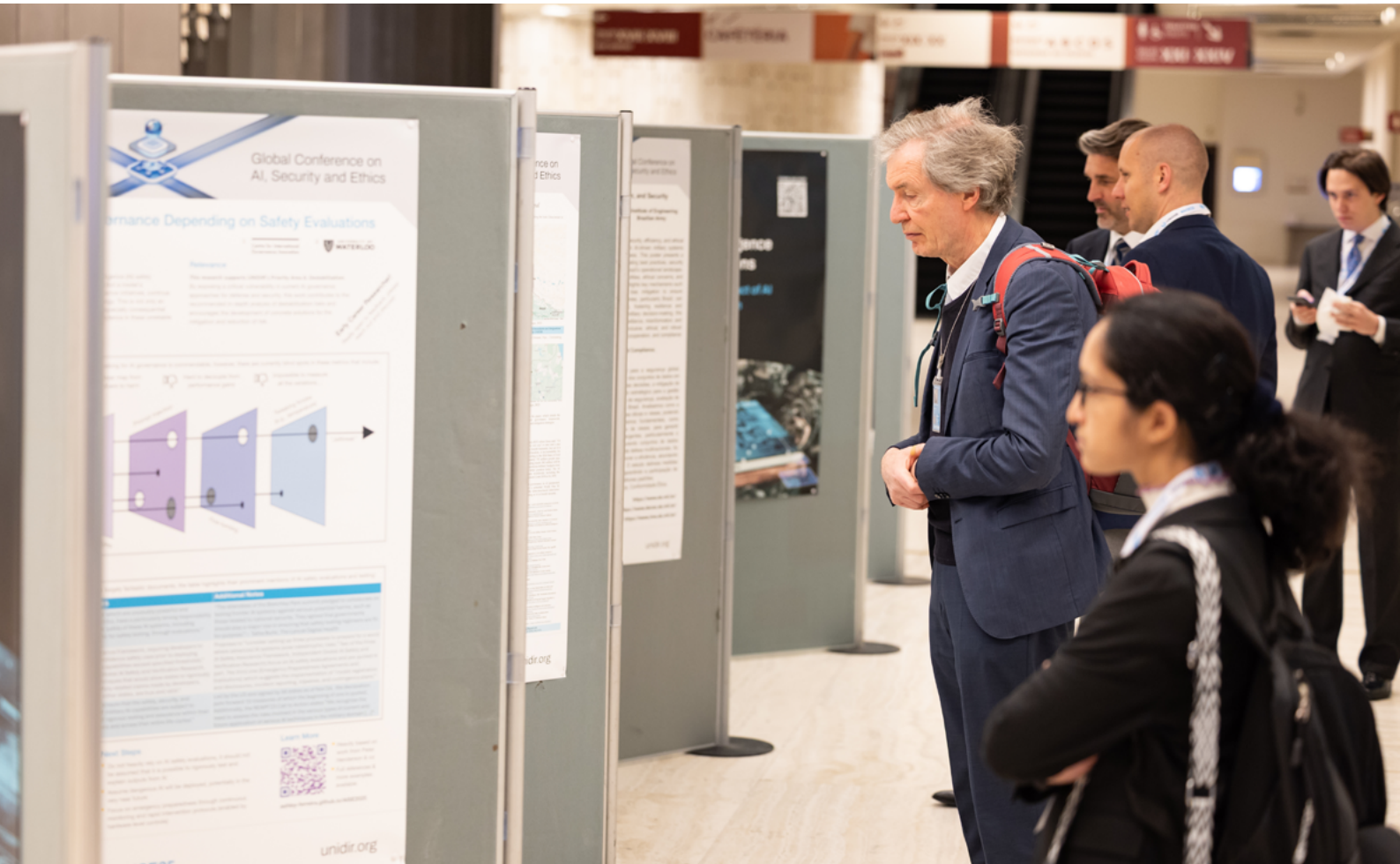
21 As noted by Jason Dass (Office of the Attorney General, Trinidad and Tobago) in day 2's panel on knowledge and capacity-building.

22 Dass.

First, the principal barrier to effective AI governance is not, in most regional contexts, the absence of frameworks or regulatory instruments itself; a wide range of principles, declarations and guidelines already exist. The challenge lies in their operationalization under the specific security conditions, institutional capacities and political constraints of diverse regional contexts. One-size-fits-all governance recommendations, as carefully constructed in multilateral processes, can produce measures that are neither tailored nor responsive to the specific problems they were designed to address. This is especially true if these multilateral process remain dominated by technologically advanced states.

Second, governance frameworks should be anchored across various, appropriate levels. National, bilateral and regional frameworks are not mutually exclusive to international and multilateral governance; rather, regional organizations and subregional bodies offer a necessary intermediary level at which governance principles developed internationally can be adapted to specific threat and security environments and institutional capacities without losing their connection to broader norms. This is not only a matter of efficiency but of legitimacy: governance frameworks that have been developed with meaningful and intentional regional input carry a unique kind of authority with the potential for impact distinct from those developed elsewhere and applied globally.

Third, genuine and meaningful inclusion in governance processes requires more than participation: it requires the development of independent capacity to evaluate and shape the technologies being governed. States that lack the institutional infrastructure to assess AI capabilities on their own terms (including the technical, legal and ethical expertise needed to make independent judgments) perceive themselves as being far from equal participants in governance processes, deliberations and initiatives, regardless of their formal representation in international forums.



# 4. AI for peace operations and humanitarian applications

## 4.1. AI for United Nations peacekeeping knowledge management

The United Nations has nearly eight decades of experience in peacekeeping operations, encompassing 71 missions (as at March 2025) across diverse conflict environments. This represents one of the most substantial bodies of institutional knowledge on the management of complex security operations available anywhere in the world. However, that knowledge is predominantly unstructured: distributed across reports, mission records, debriefs and after-action assessments that are not readily accessible to the uniformed military personnel who most need them, often under time pressure in environments with limited or no connectivity.[23]

An active application of AI in this context involves the use of large language models (LLMs) that have been trained on this accumulated record. This would enable uniformed personnel to access and contextualize relevant past mission experiences. However, a critical design limitation has emerged: LLMs cannot be used for real-time operational decision support, since their training data, by definition, does not include events occurring after the point at which the model was trained. The application is accordingly limited to the retrieval and contextualization of historical experience, in support of human deliberation, rather than as a substitute for it.

The design of this application also illustrates a governance principle of broader relevance: the introduction of AI into an ecosystem that already reflects historical biases in data collection and documentation requires active mitigation at the design stage, rather than reliance on post hoc correction. The use of synthetic data to counteract biases inherited from legacy data sets represents one possible approach to addressing this challenge.

One constraint is treated as non-negotiable in this application: as the decisions taken with this model will have direct implications for human safety and security – in addition to operational safety, security and success, the maintenance of a human in the decision loop was deemed to be necessary. This commitment is presented, not as a provisional position contingent on the future development of AI capabilities, but as a design principle embedded in the architecture of the application from the outset.

## 4.2. AI in climate risk and vulnerability assessment

Beyond the military domain, AI is being applied to identify and analyse the interlinkages between climate change, environmental degradation, insecurity and socioeconomic vulnerability – dimensions that are increasingly recognized as integral to the broader peace and security landscape. One platform, developed through collaboration between the United Nations Environment Programme (UNEP) and the Food and Agriculture Organization of the United Nations (FAO), aggregates 26 indicators spanning climate and environmental, peace and security, and socioeconomic dimensions. It uses AI-assisted

23 As noted by Avishan Bodjnoud (United Nations Department of Peace Operations) in a dedicated deep-dive on day 2: "Leveraging AI to Support Knowledge Management for Uniformed Military Personnel in UN Peacekeeping Operations".

tools to generate automated analytical reports intended to support evidence-based policymaking in fragile and conflict-affected settings.[24]

This collaboration reveals a particular commitment to open-source methodology: the underlying code and data sets are publicly accessible, enabling any user to verify the analytical basis of the outputs and to interrogate the assumptions embedded in the models. Transparency of methodology here is not merely a feature of the platform's design: it is a governance mechanism in its own right, providing the accountability infrastructure that closed or proprietary systems may not necessarily offer.

Furthermore, the application is explicitly designed as a decision-support tool, rather than a decision-making system: the analytical outputs are intended to inform human judgment, and users are required to take responsibility for any policy decisions derived from the platform's analysis. This framing – of AI as assistant to, rather than substitute for, human deliberation – reflects a governance philosophy that may be directly applicable to certain defence and security contexts, particularly those where the stakes of misuse would be considerably higher.

## 4.3. Learning from adjacent fields: The nuclear analogy and its limits

The history of arms control and non-proliferation offers a repository of institutional and diplomatic experience that the AI governance community has drawn on extensively over the past years. The challenges of managing a transformative dual-use technology with potentially catastrophic security implications are not entirely without precedent, particularly as evidenced by the nuclear field. Yet, the structural differences between the nuclear and AI contexts are sufficiently significant that analogies from nuclear governance must be applied with care and, in several key respects, must resist direct transfer.[25]

The most fundamental structural difference concerns the conditions for proliferation. Nuclear weapon development in the mid-20th century was characterized by specific, restricted and heavy resource requirements and the effective monopolization of relevant knowledge and infrastructure within a small number of state actors. These conditions made supply-side control through material restriction and verification possible and, ultimately, a governance strategy. AI development today operates under categorically different conditions: the most capable models are built on widely available hardware, trained on data that is largely public or commercially accessible, and in many cases made available through open-source repositories. In a domain where the fundamental building blocks are software and data, and where meaningful barriers to entry are low by historical standards, there is no straightforward analogue of the infrastructure of non-proliferation through import and export controls, safeguards agreements and material accounting.

The verification challenge is structural and fundamentally scientific: for nuclear weapons, the physical signatures of testing and the material requirements of production provided the basis for monitoring and verification regimes. AI capabilities leave no comparable signature.

---

24 As noted by Yhasmin Mendes de Moura (Food and Agriculture Organization) in a lightning talk on day 2: "Balancing Algorithms and Interpretation: The Role of AI in Vulnerability and Conflict Assessments".

25 As noted by Joon Baek (Columbia University) in a lightning talk on day 2: "Lessons from the Nuclear Weapons Age for Managing AI Development". For more on the limits of the analogy between the nuclear governance model and that of AI, see Yasmin Afina and Patricia Lewis, "The Nuclear Governance Model Won't Work for AI", Chatham House, 28 June 2023, https://www.chathamhouse.org/2023/06/nuclear-governance-model-wont-work-ai.

Verifying AI capabilities and their compliance with agreed constraints would require a level of access to source code, training data and operational logs that would be intrusive to a degree far beyond existing arms control arrangements. It is presently not feasible without a prior foundation of trust. It is not that verification is impossible, but rather that trust-building must precede it instead of following it – an arguably particularly difficult feature in today's security landscape.

Despite these limitations, what the nuclear experience does offer is evidence that sustained dialogue between adversarial powers can produce concrete outcomes even in conditions of deep strategic mistrust. The mechanisms developed in the nuclear context – risk reduction, hotlines, confidence-building measures and structured dialogue processes – carry lessons for AI governance even where substantive analogies fail.

# 5. Cross-cutting use-case observations

## 5.1. The specificity imperative

Each of the use cases examined at AISE25 brings to the fore the fact that governance frameworks calibrated at a high level of generality are not always necessarily adequate for governing AI deployment in practice. Principles alone are not sufficient to bridge operationalization and compliance requirements: specification of what those principles require in concrete operational contexts will be critical.[26]

This observation has a temporal dimension as well. The traditional ethics of technology attest to the difficulty of reversing design choices once a technology is deployed at scale: the costs of retrofitting governance onto systems already embedded in institutional practice and procurement cycles are substantially higher than the costs of building governance requirements into design from the outset (see Section 2.2).[27]

This argument applies both to the development of specific AI systems and to the construction of governance frameworks: principles established before deployment carry a different kind of force, and a different practical reach, from those formulated in response to problems that have already manifested. As such, governance efforts should engage with use cases early and specifically, and they should resist the temptation to defer to principles as a substitute for the harder work of specifying what those principles require in context.

## 5.2. The accountability chain in practice

A consistent finding across the use cases is that the accountability challenges posed by AI in defence and security are not confined to the point of use: they extend across the full technology life cycle, from design and development through procurement, testing, deployment and, eventually, decommissioning. Governance frameworks that focus exclusively on the moment of operational decision-making capture, at most, one link in an accountability chain that begins considerably earlier.[28]

---

26 Davidovic.

27 Benanti.

28 This theme was addressed across multiple sessions; in particular by Drexx Laggui (Cybercrime Investigation and Coordinating Center, Philippines) in a dedicated deep-dive on day 1 on "AI on the Battlefield: Building Trust and Accountability Through Digital Forensics", and Alexi Drew (International Committee of the Red Cross) in the facilitation of the life cycle management session on day 1.

This observation has practical implications for how accountability requirements are formulated. The question of who bears responsibility when an AI-enabled system produces an unlawful outcome is not answered simply by examining the decision-making process at the time of the incident; it requires an examination of the choices made in design, the adequacy of testing, the parameters set in the operational specification, the accuracy of the representations made by the developer to the procuring authority, and the decision-making process surrounding the deployment of these technologies.

Distributing accountability in this way (i.e., across multiple actors and multiple stages of the life cycle) is far from being a means of evading it. It is a more accurate reflection of how responsibility is actually constituted in complex sociotechnical systems and across the many layers of the AI stack. This argument was echoed by conference participants who engaged in a scenario involving AI-enabled targeting and civilian harm: they established that responsibility would most likely rest with more than one actor, a view that converges with the technical and forensic dimensions of the accountability question discussed at AISE25.[29]

## 5.3. Capacity asymmetry as the defining use-case governance challenge

Across the regional use cases considered at AISE25, a structural asymmetry is visible to which governance frameworks focused primarily on the concerns of technologically advanced states tend to give insufficient weight: the states most consequentially affected by the operational deployment of AI-enabled security systems are, arguably and in a significant number of cases, not the states developing them.[30]

As such, states that lack the institutional, technical and human capital infrastructure to independently assess AI systems (i.e., to evaluate their capabilities, verify their operating parameters and assess their compliance with applicable legal and ethical requirements) are not, in any meaningful sense, in a position to govern those systems, regardless of their formal participation in governance processes. They may be present in multilateral and international forums, and they may endorse declarations and guidelines; but without the capacity to translate those instruments into domestic governance requirements and to enforce those requirements in procurement and deployment, the gap between principle and practice will persist.

Addressing this asymmetry requires both technical and political components. The technical dimension involves sustained investment in knowledge infrastructure: universities, research institutions, government bodies and legal communities with the capacity to engage with AI governance as a substantive domain. The political dimension requires that existing governance processes actively create space for perspectives that emerge from different operational contexts and threat environments, and that the development of governance norms reflects, rather than presupposes, that diversity.

---

29 The distributed nature of responsibility was reflected in live audience polling conducted during Davidovic's deep-dive on day 2.

30 This observation was particularly evident in the regional perspectives shared in day 2's lunch panel, and in the contributions of Ng and Dass.

## 5.4. AI applications beyond weapons: The underrepresented use-case space

A dimension of the AI-in-security landscape that had received insufficient attention in governance discussions prior to AISE25, and that the conference helped to surface, is the range of AI applications in the military and security domain that do not involve weapon systems or lethal force, but that nonetheless carry significant governance implications. These applications include command and control, intelligence analysis, logistics, training systems, knowledge management, humanitarian assessment and conflict forecasting. They represent a category of AI deployment that is, in several respects, more mature, more widespread and more immediately consequential for the day-to-day operation of security institutions than the autonomous weapon applications that dominate governance discourse.[31]

This observation is not merely empirical: it has normative implications for governance design. The current architecture for governance frameworks tends to be shaped in significant part by the history of discussions on lethal autonomous weapon systems (LAWS) in the Convention on Certain Conventional Weapons. If this continues, these frameworks will systematically fail to address a broader and arguably more immediately relevant category of AI deployment. Worse, they may create and establish the misconception that AI applications outside the lethal category are governance-free. There will thus be no framework for the accountability questions that arise when, for instance, an AI-assisted intelligence system produces an erroneous assessment that contributes to a policy decision with significant humanitarian consequence.

The use cases presented at AISE25 in this category – on AI for peacekeeping knowledge management, climate risk and vulnerability assessment, and conflict forecasting (see Section 4) – each illustrate, in different ways, what responsible AI application looks like when governance considerations are embedded in design from the outset. They also suggest the practical value of building governance capacity in less contested application spaces as a foundation for addressing more sensitive ones.

# 6. Conclusion: Carrying the conversation forward from AISE25 to AISE26

AISE25 served as a useful waypoint to map the use-case landscape for AI in defence and security as at March 2025: not a comprehensive survey, but a grounded cross section of operational contexts, governance challenges and analytical frameworks that together illuminate both the diversity of AI's applications in this domain and the systemic features of the governance challenges they collectively pose.

AISE26 will convene having inherited several of the unresolved questions that AISE25 identified, while also benefiting from a denser institutional landscape. Some of the key governance milestones that have been achieved between AISE25 and AISE26 include:

- The publication of the United Nations Secretary-General's report on AI in the military

31 Drew.

domain and its implications for international peace and security[32]

- The adoption by the United Nations General Assembly of resolution 80/58, a follow-up to resolution 79/239 on AI in the military domain and its implications for international peace and security,[33] which led to the organization of a three-day informal exchange on this issue to be held in Geneva on 15–17 June 2026[34]
- The organization of the third Responsible AI in the Military Domain (REAIM) Summit, held in A Coruña, Spain, in February 2026
- The publication of the Strategic Guidance Report of the Global Commission on Responsible AI in the Military Domain (GC REAIM)[35]
- The launch of the UNIDIR-led project for the development of a Framework of Responsible Industry Behaviour for AI in the Military Domain, to be developed in partnership with the Office of the United Nations High Commissioner for Human Rights (OHCHR) and in consultation with industry representatives and governments[36]

Against this backdrop of advancing governance activity, several questions that AISE25 paved the way for are particularly well-suited to structured engagement at AISE26 including, among others:

- What human element would be required across the spectrum of AI applications in the military and security domains? What level of specificity would be required and on what basis? What are the implications of recent technological advancements, particularly with the advent of agentic AI?
- How can the international community address the implications of AI beyond weapon systems? At what level should they be addressed, and by whom?
- What should each layer of AI governance in the context of international peace and security address? What would be appropriate to address at the multilateral, international, regional and/or national levels? How can the international community ensure that each layer complements and reinforces one another, and what would coherence across processes, initiatives and frameworks require?
- How can the international and wider multistakeholder community move beyond principles and work towards their implementation and operationalization? What specific guidance would be required for the operationalization of existing commitments by all stakeholders involved (e.g., procurement officers, commanders, legal advisers, developers and technology providers)?

---

32 General Assembly, "Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security", Report of the Secretary-General, A/80/78, 5 June 2025, https://docs.un.org/A/80/78.

33 General Assembly, resolution 80/58, "Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security", 1 December 2025, https://docs.un.org/A/RES/80/58.

34 Office for Disarmament Affairs, "UNODA Science, Technology and International Security Unit – Meeting", Meeting Place, 2026, https://meetings.unoda.org/unoda-stu-meeting/unoda-science-technology-and-international-security-unit-meeting-2026.

35 Global Commission on Responsible Artificial Intelligence in the Military Domain (GC REAIM), *Responsible by Design: Strategic Guidance Report on the Risks, Opportunities, and Governance of Artificial Intelligence in the Military Domain* (The Hague: GC REAIM, September 2025), https://hcss.nl/wp-content/uploads/2025/09/GC-REAIM-Strategic-Guidance-Report-Final-WEB.pdf.

36 UNIDIR, "Framework of Responsible Industry Behaviour for AI in the Military Domain", 2026, https://unidir.org/framework-of-responsible-industry-behaviour-for-ai-in-the-military-domain/.

## Acknowledgments

Support from UNIDIR's core funders provides the foundation for all of the Institute's activities. This report was produced by UNIDIR's Security and Technology Programme, which is supported by the Governments of France, Germany, Italy, the Netherlands, the Republic of Korea and Switzerland, and Microsoft for its work on AI and autonomy. In addition, the Global Conference on AI, Security and Ethics 2025 was also supported by Advai.

UNIDIR extends its most sincere gratitude to all speakers, moderators, poster authors and audience members for their insightful presentations, comments and other contributions, which form the foundation of this report. UNIDIR also extends its profound gratitude to the members of the jury who helped review the abstracts received from UNIDIR's open call for proposals, which contributed to the conference's high calibre and the quality of the discussions: Dr. Alexi Drew, Major Jamal Mohamed Hassan, Calum Inverarity, Michael Karimian and Dr. Magdalena Pacholska. The evaluations were presented in the jury members' independent capacity and do not necessarily reflect the views or opinions of the jury members or of the organizations with which they work.

The authors wish to thank Dr. Giacomo Persi Paoli (UNIDIR) for advice, guidance, vision and support for the Global Conference and for his review of this report. Thanks for their support throughout the organization of the conference are also due to Jessica Lee Abowitz, Sapar Annayev, Asa Cusack, Jessica Espinosa Azcárraga, Anna Grangier, Edward Madziwa, Federico Mantellassi, Claudia Marquina and Mireia Mas Vivancos (UNIDIR), as well as to Elucidate Studios for the design support.

## About UNIDIR

UNIDIR is a voluntarily funded, autonomous institute within the United Nations. As one of the few policy institutes worldwide that focus on disarmament, UNIDIR generates knowledge and promotes dialogue and action on disarmament and security. Based in Geneva, it assists the international community in developing the practical, innovative ideas needed to address critical security problems.

## About the UNIDIR Security and Technology Programme

Contemporary developments in science and technology present both new opportunities and challenges to international security and disarmament. UNIDIR's Security and Technology Programme aims to build knowledge and awareness about the international security implications and risks associated with specific technological innovations. It also convenes stakeholders to explore ideas and develop new approaches to address these issues.

## 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 necessarily reflect the views or opinions of the United Nations, UNIDIR, its staff members or sponsors.

## Citation

Yasmin Afina and Jan Hendrik Mannsperger, "AISE Markers Series – Benchmark III: Use Cases, Insights from the 2025 Global Conference on AI, Security and Ethics", Geneva: UNIDIR, 2026.

## About the authors

This report was produced by UNIDIR's Security and Technology Programme. It was drafted by Dr. Yasmin Afina and Jan Hendrik Mannsperger.

## Photo credit

Cover photo generated with AI. Credit: Adobe Stock / Dmitry. All other photos featured were taken by Pierre Albouy at the Global Conference on AI, Security and Ethics 2025, 27–28 March 2025, Geneva. Credit: UNIDIR / Pierre Albouy.

## Acronyms and abbreviations

| | |
|---|---|
| **AI** | Artificial intelligence |
| **AISE** | Global Conference on AI, Security and Ethics |
| **ASEAN** | Association of Southeast Asian Nations |
| **IHL** | International humanitarian law |
| **LAWS** | Lethal autonomous weapon systems |
| **LLM** | Large language model |
| **RAISE** | Roundtable for AI, Security and Ethics |

UNIDIR

@unidir
/unidir
/un_disarmresearch
/unidirgeneva
/unidir

Palais des Nations
1211 Geneva, Switzerland



UNIDIR.ORG