HUMANITARIAN LAW & POLICY

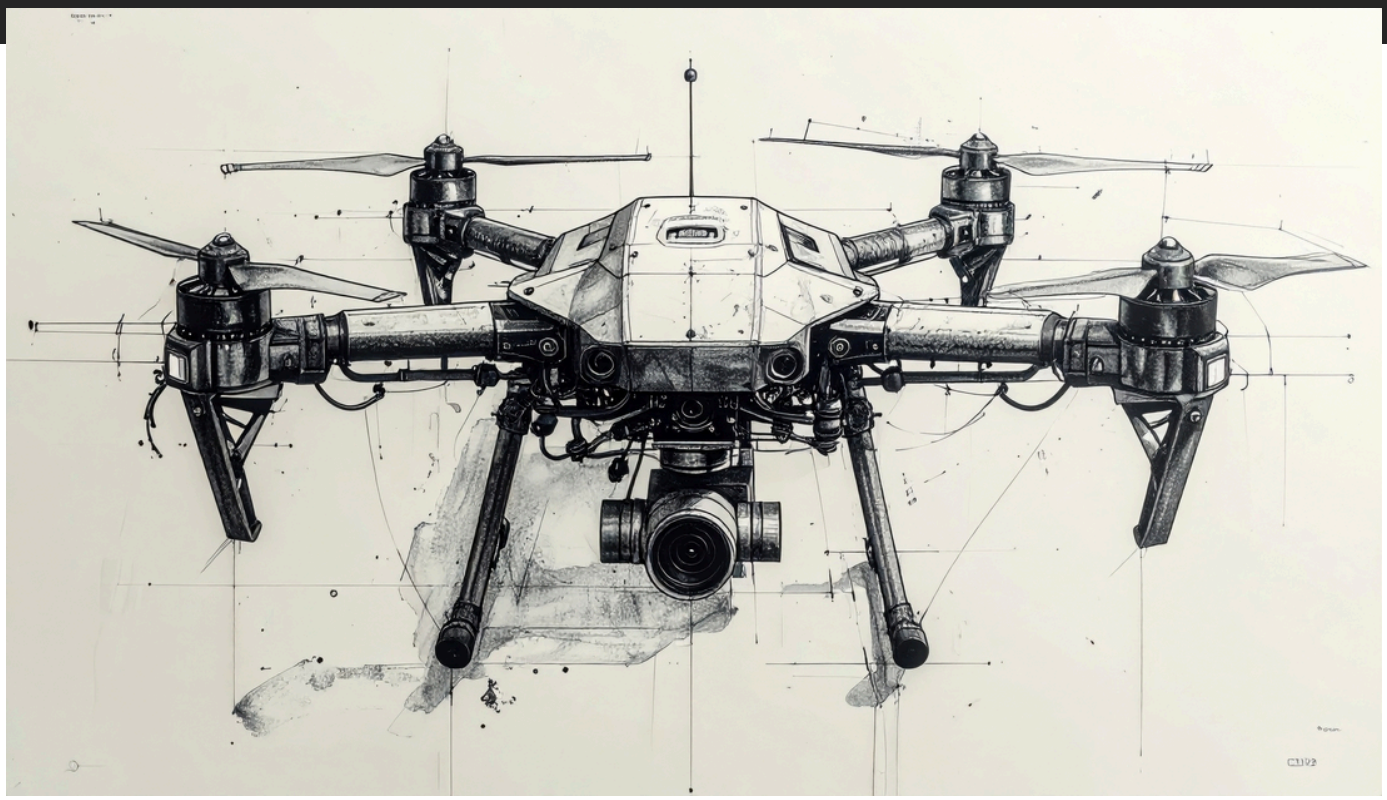
Collaboration without over-reliance: the role of industry in making military AI "lawful by design"

May 26, 2026, Accountability / Analysis / Artificial intelligence in military decision-making / IHL / New Technologies / Weapons

9 mins read

Laura Bruun
Senior Researcher, Stockholm International Peace Research Institute (SIPRI)

Netta Goussac
Special Counsel, Lexbridge; Associate Senior Researcher, SIPRI



In the policy debate on artificial intelligence (AI) in the military domain, there is a growing consensus that international humanitarian law (IHL) must be a central consideration in the design of military AI systems. The imperative to make military AI systems "lawful by design" has, naturally, led to a sharper focus on the role of industry. But what this means in practical terms for AI suppliers — and how states can and should collaborate with industry to strengthen IHL compliance — remains an open question.

In this post, Laura Bruun and Netta Goussac from the Stockholm International Peace Research Institute (SIPRI) argue that while focusing on IHL at the design stage makes sense, it carries the risk that states over-rely on industry to make military "lawful by design". Efforts to elaborate what it means to make military AI "lawful by design" must be grounded in realistic expectations and limits, as well as clear legal responsibilities.

The number of reports about AI being used in targeting in active armed conflicts has mounted in the past few weeks and months, from Gaza to Venezuela to Iran and Lebanon. This is happening while states are still debating how to ensure that military AI is developed and used responsibly. In these discussions, a consensus is emerging that compliance with international humanitarian law (IHL) must be considered at the design stage. Many states agree that decisions made at the design stage, such as those around the selection of training data or object classification methodologies, affect whether a military will be able to employ an AI-enabled capability lawfully in an armed conflict, including in targeting decisions.

While this makes sense, there is a risk that the desire to make military AI "lawful by design" leads to a dynamic in which a state places disproportionate trust in its supplier — usually an industry actor — to deliver IHL-compliant military AI capabilities. As states are set to specify the relationship between AI design and legal

compliance, it is key that such efforts are grounded in a clear understanding of what cannot, and should not, be expected of industry.

What does “lawful by design” mean in practice?

The sharpened focus on IHL at the design stage of military AI systems is linked to the “responsible by design” approach introduced into the policy debate by the Global Commission on Responsible AI in the Military Domain (*GC REAIM*) and the views shared by several states with the *United Nations Secretary-General report in 2025*.

Drawing on the established “X by design” paradigm in technology governance, “responsible” in this context is generally understood to encompass notions of safety, reliability, fairness, transparency, accountability and – importantly – lawfulness. The “responsible by design” approach thus reflects a “lifecycle” approach to lawfulness – including IHL compliance – that can be traced back to the early days of state deliberations on autonomous weapons (*here*).

But what it means to make military AI “lawful by design” in practice remains an open question. Based on existing material, it could include hard-coding translated legal texts into an AI system (see, for example, *Arkins’ “ethical governor*” or *“law-following AI*”), conducting legal reviews and risk assessments during the design phase (as reflected in *REAIM “Pathways to Action*”), or holding industry to certain standards (as reflected in *a recent initiative* by UNIDIR).

One thing is clear: making military AI “lawful by design” requires collaboration with those involved in the design of military capabilities, usually industry actors. Industry actors make choices that are relevant to IHL compliance, for example, around decision thresholds (i.e. the point at which something is classified positively as a target), levels of autonomy, selection of training data, and testing relevant to operational conditions governed by IHL. State collaboration with industry makes sense, not least because implementing IHL obligations can be seen as a collective exercise whereby states could *responsibly rely* on a wide range of actors involved in the design and use phases.

Now, what remains is for states to elaborate on what such collaboration with industry should look like. To this end, it is critical that states are mindful that the desire to *collaborate* with industry does not turn into *over-reliance* on industry.

The risk of over-relying on industry

To avoid the risk of over-reliance on industry actors in pursuit of AI-enabled military capabilities that are lawful by design, states need to keep three things in mind. First, they should avoid unrealistic expectations of industry. Second, they should accept the limits of design decisions. And third, they should acknowledge that the obligation to respect IHL ultimately rests with states, not with industry.

Industry actors are good at many things; navigating the Geneva Conventions may not be one of them

In seeking to ensure IHL compliance, states should have realistic expectations of the (in)ability of industry to undertake tasks involving knowledge of IHL, such as translating legal texts into technical requirements or into testing and evaluation standards. The types of industry actors that states would need to collaborate with to ensure that military AI is “lawful by design” range from legacy primes, defense startups and “neo primes” to big tech and foundation model providers (as explained *here*).

While these companies drive innovation in the field of military AI, they do not necessarily prioritise IHL compliance in the way that states (must) do, so they do not appear to invest in such expertise. Few companies promote themselves as having IHL expertise. For example, references to IHL are largely absent from company websites and recruitment pages, including those of Anduril, Palantir, Helsing, Shield AI, Anthropic, BAE Systems, and Lockheed Martin. And while a handful of companies involved in military AI innovation brand themselves as being “responsible” or “ethical” actors (for example, *Thales*, *Northrop Grumman*, *Palantir*; *Rebellion Defense*, *Helsing*, *Skydio* and *Microsoft*), IHL compliance is not mentioned as a component of such efforts.

Of course, an increased demand by states for IHL expertise may motivate companies to invest in such expertise in the future. But industry actors vary in size and resources (recalling that many key actors are small startups), and not all will necessarily be in a position to invest in such IHL expertise. Lack of IHL expertise will affect when and how states collaborate with industry to ensure that military AI is lawful by design.

IHL compliance cannot be engineered

The “lawful by design” framing is inviting. It casts IHL compliance as an engineering problem that can be solved. But the framing leaves out part of the picture: design measures alone are not sufficient to ensure compliance with IHL. Compliance with IHL depends on the system, how it is used, and the circumstances it is used in (as explained *here*).

A system that is designed to facilitate IHL compliance – for example through embedding certain red lines – may still be involved in violations of IHL if it is used in circumstances for which it was not designed, in environments that exceed its operational parameters, or by personnel who lack adequate understanding of its capabilities and limitations.

Therefore, to avoid over-relying on industry for IHL compliance, states (and industry) need to keep in mind that IHL cannot be reduced to an engineering problem. Rather, legal compliance is a consideration relevant throughout the lifecycle of an AI-enabled military capability and in the socio-technical institutions (such as militaries) in which such capabilities are developed and used (*here*). Because indeed, while the design phase is an important locus for ensuring that military AI capabilities can be used lawfully, so too are all other phases.

IHL applies to states, not industry

Finally, states should be mindful that IHL applies primarily to states, not to industry actors. Any attempt to define what it means in practical terms to make military AI “lawful by design” must be grounded in the responsibility of states, not industry, to implement IHL obligations. To this end, it is essential that states maintain an independent capacity to assess and verify supplier claims about IHL-relevant properties of military AI capabilities, rather than accepting assurances at face value.

This may mean investing in technical literacy and processes, such as independent testing and evaluation (whether in-house or through trusted third parties), legal reviews, or contractual requirements that mandate disclosure of IHL-relevant information (as argued *here*). It may also require states to translate their legal obligations into meaningful guidance to industry.

Towards responsible collaboration

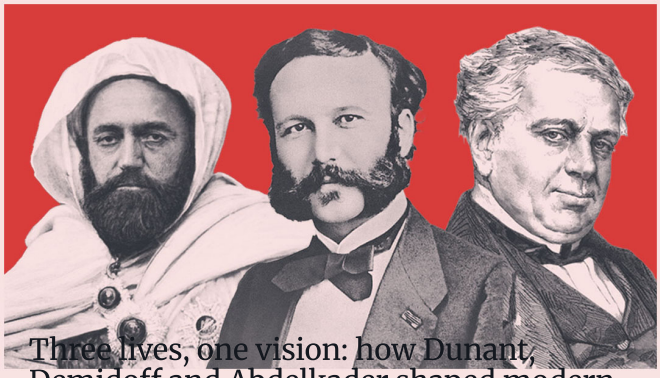
Efforts to ensure that military AI can be used in compliance with IHL will require states to work closely with industry – but not (over)rely on it. *Avoiding* unrealistic industry expectations, *accepting* the limits of design decisions, and *acknowledging* that the obligation to respect IHL ultimately rests with the state are the first steps to ensuring a responsible collaboration with industry actors.

The risk of overreliance arises from an agenda that places design at the centre of legal compliance while, so far, leaving the industry’s responsibilities undefined. As the lawful by design concept gains traction, states should ensure that it strengthens, rather than dilutes, diffuses, or obscures IHL compliance in the context of military AI.

See also

- Yéelen Marie Geairon, *Deciding under algorithms: artificial intelligence and the protection of civilian infrastructure in armed conflict*, March 12, 2026
- Laura Bruun, Marta Bo, *‘Constant care’ must be taken to address bias in military AI*, August 28, 2025
- Joanna L D Wilson, *AI, war and (in)humanity: the role of human emotions in military decision-making*, February 20, 2025
- Elke Schwarz, *The (im)possibility of responsible military AI governance*, December 12, 2024

Tags: AI, armed conflict, Artificial Intelligence, assistance, compliance, conduct of hostilities, IHL, international humanitarian law, law of armed conflict, LOAC, Means and Methods of Warfare, modern warfare, Respect for IHL



Three lives, one vision: how Dunant, Demidoff and Abdelkader shaped modern humanitarianism

15 mins read

Accountability / Analysis / Artificial intelligence in military decision-making / IHL / New Technologies / Weapons

Anastasia Kushleyko, Cédric Cotter & Ahmed Al-Dawoody

The brutal effects of war have long prompted efforts to limit suffering and preserve humanity ...

Enforced disappearances: universal responses to a worldwide phenomenon

14 mins read

Accountability / Analysis / Artificial intelligence in military decision-making / IHL / New Technologies / Weapons Milica Kolaković-Bojović & Grażyna Baranowska

Enforced disappearances remain an issue of profound seriousness, with lasting consequences for the families of ...