Deciding under algorithms: artificial intelligence and the protection of civilian infrastructure in armed conflict

March 12, 2026, Analysis / Artificial Intelligence and Armed Conflict / Artificial intelligence in military decision-making / Conduct of Hostilities / IHL / New Technologies / Technology in Humanitarian Action

🕐 10 mins read

Yéelen Marie Geairon
Legal Expert in IHL and
Emerging Technologies

AI-generated image

Artificial intelligence (AI)-based decision-support systems are increasingly embedded upstream of the use of force, shaping how military actors plan attacks, assessing effects, and anticipating harm. In contemporary urban warfare, where civilian infrastructure forms complex and deeply interconnected systems, these tools are increasingly used to guide decisions with far-reaching humanitarian consequences. This raises critical questions for international humanitarian law (IHL), which requires parties to anticipate and mitigate foreseeable civilian harm when applying the principles of proportionality and precaution, including indirect, cumulative and systemic effects on civilian infrastructure.

In this post, independent legal researcher Yéelen Marie Geairon argues that while AI-enabled decision-support systems do not alter the legal rules governing attacks, they significantly reshape how foreseeability is operationalized in practice. By structuring what decision-makers are able to anticipate, compare and justify ex ante, AI systems recalibrate the factual basis of legal judgment, while also introducing new risks linked to data gaps, opacity and over-reliance on technical outputs. The protection of civilian infrastructure in AI-enabled warfare therefore depends less on technological performance than on the legal discipline, transparency and human judgment with which these tools are embedded in decision-making processes.

ICRC Humanitarian Law & Policy Blog · Deciding under algorithms: artificial intelligence and the protection of civilian infrastructure in armed conflict

Artificial intelligence (AI)-based decision support systems are now fully integrated into contemporary armed conflicts. Beyond the debates on the autonomy of weapons, their deployment is now concentrated upstream of the use of force, in the processes of planning, targeting and evaluating the effects of attacks. Notably, within urban warfare, where civilian infrastructure is a complex and deeply interconnected system, these tools are increasingly being used to guide decisions that could

have major humanitarian consequences. By structuring what decision-makers are able to foresee, compare and justify, AI-enabled decision support systems are quietly but substantively reshaping how international humanitarian law (IHL) is applied to attacks when affecting civilian infrastructure.

From uncertainty to simulation: AI and the new architecture of foreseeability

Recent conflicts in *Ukraine*, *Gaza* and *Sudan* reveal a recurring pattern in contemporary warfare: the most severe humanitarian consequences of attacks on civilian infrastructure arise not from the initial strike, but from the prolonged disruption of interconnected civilian systems. Electricity outages disrupt water supply and hospital operations; degraded water and telecommunications networks undermine healthcare and emergency response; and damaged roads and markets restrict access to food and basic services. International humanitarian law specifically requires that such indirect, delayed and structural effects be taken into account in the ex ante assessment of attacks, particularly in the light of the principles of proportionality and precaution.

It is at this stage of anticipation that AI-based decision support systems are now presented as tools capable of grasping the growing complexity of armed conflict environments.

In practice, this includes the deployment of *Intelligence, Surveillance and Reconnaissance (ISR)* systems which, as an extension of initiatives such as *Project Maven*, rely on the automated analysis of drone and satellite imagery to identify and prioritize objects of interest in densely built environments. In addition, machine learning–based such as *collateral damage estimation* tools, cross-reference data on urban densities, the location of critical infrastructure and different impact scenarios to produce probabilistic projections of the direct and indirect effects of a strike. These data fusion platforms aggregate military intelligence, geographic information systems, open source, and humanitarian data to generate visualizations to inform operational decision-making.

Although the military uses and technical parameters of these tools remain largely undocumented publicly, the underlying logic is not unprecedented. Comparable systems are well understood and widely used in civilian contexts to manage complex and interconnected infrastructure networks, as well as in various safety and crisis management contexts.

In the *energy sector*, AI-based decision support tools are used to model load distribution, anticipate cascading failures, and assess the downstream effects of a power cut on essential services, including water supply, transportation, and healthcare. Similarly, in *disaster risk management* and *urban resilience planning*, machine learning models are used to simulate the humanitarian consequences of infrastructure disruptions and to compare mitigation options within time constraints. These systems are not based on perfect data or the total elimination of uncertainty; rather, they aim to reduce the range of possibilities, highlight systemic vulnerabilities, and structure complex choices.

Transposed to the military sector, these modelling and data fusion technologies appear to be particularly relevant for the evaluation of attacks targeting *dual-use infrastructures*, which are frequent in urban warfare.

For example, an electrical substation can support certain military capabilities while playing a central role in the operation of the civilian grid (e.g. home power supplies, hospitals, water pumping stations, etc.). In a more traditional approach, the ex ante assessment of such an attack would be based on an essentially contextual and prospective assessment. The likelihood, duration and extent of a blackout would remain indeterminate, and the chain effects on essential services would be considered in general, marked by significant uncertainty. Thus, indirect damages, while assessed in principle, would be difficult to specify precisely, leaving the decision-maker with a relatively wide margin of appreciation as to what could reasonably have been anticipated at the time of the decision.

The use of AI-based decision support systems significantly changes this framework.

By modeling the load of the power grid, simulating different outage scenarios and integrating contextual data on the dependence of civil services, the system can estimate the likely duration of the outage, identify the affected areas and project the consequences such as on the water supply or the operation of hospitals.

This dynamic is not limited to energy infrastructure. Telecommunications infrastructures, often also dual-use, play a central role in the protection of civilian populations by providing access to emergency services, coordinating aid efforts and disseminating vital information. When a data fusion platform reveals that the same network supports both military functions and critical civilian communications, simulations can show that the humanitarian impact of a disruption varies greatly depending on the context and time.

Here again, AI does not decide on the attack, but it modifies the informational content on which the attack decision is based, by making visible effects that would otherwise have been difficult to anticipate within tight operational deadlines.

The illusion of precision: data gaps, blind spots and model bias

Despite this opportunity, the deployment of reasoning towards probabilistic and systemic analyses is not without generating significant legal and humanitarian tensions.

By translating complex urban realities into models, scores, and probabilities, decision support systems can create illusions of *control* and *objectivity*. Similarly, the limitations of the data available in conflict contexts, marked by their instability, partial nature and inadequacy in the face of rapidly changing situations, can lead to an underestimation of delayed but serious humanitarian harms. Indeed, in the context of armed conflict, practices such as improvised infrastructure repairs, the use of alternative networks, or the rapid adaptation of populations to severe disruptions are often not captured in available datasets. *Humanitarian experience* illustrates this gap, for example, through the reliance on informal routes, frequently absent from databases, for the delivery of aid and access to essential services.

As a result, a decision heavily based on such simulations may overlook serious humanitarian consequences, not because they were excluded in the balancing of the expected concrete and direct military advantage against civilian harm, but because they were insufficiently represented, or even made invisible by the model, thus distorting the parameters on which the decision-making process is based.

Against this backdrop, the issue is not only the technical reliability of AI-enabled decision-support tools, but the degree of legal diligence with which human judgment is exercised when applying the principles of IHL.

Under international humanitarian law, the application of proportionality and precaution hinges on what can reasonably be anticipated at the time of decision-making. This assessment becomes particularly complex where attacks affect civilian infrastructure whose humanitarian significance lies less in immediate physical damage than in cumulative and downstream effects.

Schools illustrate this difficulty. While not vital installations in a strict sense, they structure daily civilian life and support a range of social functions whose disruption may generate long-term humanitarian harm. A strike on a military objective in the immediate vicinity of a school can have consequences that go far beyond the physical destruction of the building, including the prolonged disruption of education, the loss of community centres or temporary shelters, and the lasting disruption of essential social services. These effects, which are often diffuse and indirect, are difficult to quantify and are therefore likely to be underestimated in analyses that focus on immediately measurable indicators.

When AI-based decision-support systems structure how foreseeability is operationalized, by privileging certain forms of harm over others, they shape the factual horizon within which proportionality and precaution are applied. Legal standards remain unchanged, but the scope of what is recognized as foreseeable civilian harm and its assessment must be recalibrated.

Human judgment under algorithmic influence

It is in this context that the broader question of the use of AI-based decision support systems arises. The risk is not so much that these tools will replace the human decision-maker, but that *legal reasoning will tend to retreat behind a technical output perceived as neutral or scientifically indisputable*, reducing the capacity for critical and contextual appreciation required by IHL at the time of decision-making. In this sense, AI does not create new legal obligations, but it contributes to redefining and raising the level of due diligence expected in the application of existing obligations, while influencing the conditions under which decisions can be reviewed and assessed a posteriori, in particular with regard to the liability of decision-makers.

International humanitarian law remains clear: *responsibility for the use of force rests with human actors and cannot be transferred to a machine*. However, when decisions are profoundly structured by algorithmic systems, the ex post assessment of their compliance with IHL depends on the ability to trace how the available information was interpreted, weighted, and integrated into the final decision. In the absence of sufficient safeguards in terms of transparency, documentation, and traceability, the issue is not so much the automation of decision-making as the weakening of the practical conditions for establishing whether the requirements of proportionality and precaution have been met, including for the purposes of establishing liability in cases of serious violations of IHL.

In this context, the protection of critical civilian infrastructure depends less on the technical performance of AI than on how these tools are embedded within rigorous, critical and well-documented legal decision-making processes. As AI increasingly shapes how harm is anticipated before the use of force, it can either narrow legal reasoning to numerical abstractions that obscure the lived reality of civilian populations or support a more demanding exercise of judgment.

When integrated without safeguards, AI risks masking indirect and systemic harm. When embedded in transparent, contestable and legally disciplined processes, it can instead strengthen the diligence of decision-makers in anticipating such harm and promote operational choices that more faithfully reflect the fundamental objectives of international humanitarian law: the protection of life, human dignity and essential civilian services.

See also

- Laura Bruun and Marta Bo, *'Constant care' must be taken to address bias in military AI*, August 28, 2025
- Joana 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
- Wen Zhou and Anna Rosalie Greipl, *AI in military decision-making: supporting humans, not replacing them*, August 29, 2024

Tags: AI, armed conflict, Artificial Intelligence, Civilians, compliance, conduct of hostilities, IHL, international humanitarian law, Means and Methods of Warfare, modern warfare, Urban Warfare and Violence

Engaging non-state armed groups on the protection of missing people and their

The adoption of the 1949 Geneva Conventions: a huma...

families

🕐 11 mins read

 Analysis / Artificial Intelligence and Armed Conflict / Artificial intelligence in military decision-making / Conduct of Hostilities / IHL / New Technologies / Technology in Humanitarian Action Tilman Rodenhäuser & Ximena Londoño

The ICRC continues to witness unacceptable levels of suffering when the law designed to protect …

colonial continuity

🕐 12 mins read

 Analysis / Artificial Intelligence and Armed Conflict / Artificial intelligence in military decision-making / Conduct of Hostilities / IHL / New Technologies / Technology in Humanitarian Action Srinivas Burra

More than seven decades after their adoption, the four Geneva Conventions of 1949 remain foundational …