UNSCR 1540 imposes binding obligations on States to prevent non-State actors from developing, acquiring, manufacturing, possessing, transporting, transferring or using nuclear, chemical and biological weapons and their means of delivery, and to establish effective domestic controls to that end

New technologies and the manipulation of facts in armed conflicts

by Ari Koutale Tila Boulama Mamadou

Framing context

Armed conflict is undergoing a structural transformation driven by rapid technological innovation. Artificial intelligence (AI), autonomous and semi-autonomous weapon systems, algorithmic decision-support tools, cyber operations and digital platforms for information manipulation are increasingly embedded in doctrine and battlefield practice.[1] This transformation does not only change the "means and methods" of warfare; it reshapes how facts are produced, how responsibility is attributed, and how international humanitarian law (IHL) is interpreted and enforced.[2]

IHL is built on a factual architecture: who was targeted, what was known (or reasonably believed) at the time, which precautions were taken, and whether expected civilian harm was excessive in relation to the anticipated military advantage. In conventional settings, those judgements presume human contextual reasoning and a traceable chain of decision-making. In today's conflicts, however, technical systems increasingly mediate both decision and documentation. Algorithmic recommendations can shape targeting decisions; higher autonomy can compress the time available for legal review; and digital manipulation can undermine the reliability of information used by fact-finders and courts.[3]

These dynamics intersect with the preventive logic of United Nations Security Council resolution 1540 (2004). UNSCR 1540 imposes binding obliga-

1 International Committee of the Red Cross (ICRC), International Humanitarian Law and the Challenges of Contemporary Armed Conflicts, 2024.

2 Ibid.

3 Office of the United Nations High Commissioner for Human Rights (OHCHR), Berkeley Protocol on Digital Open Source Investigations, last accessed December 18, 2025.

“The humanitarian cost is borne by people on the ground: civilians who look “abnormal” to an algorithm, humanitarian staff moving near contested areas, and communities whose daily routines are misread as hostile patterns

tions on States to prevent non-State actors from developing, acquiring, manufacturing, possessing, transporting, transferring or using nuclear, chemical and biological weapons and their means of delivery, and to establish effective domestic controls to that end.[4]

This article argues that emerging military technologies generate a dual risk. First, they strain compliance with IHL principles such as distinction, proportionality and precaution by shifting critical judgement toward machine-mediated processes. Second, they destabilise the evidentiary foundations needed to investigate violations, attribute responsibility, and support preventive action under frameworks such as UNSCR 1540.

Autonomous and algorithmic targeting: reconfiguring distinction, proportionality and human responsibility

The targeting cycle is a central pressure point. AI systems are increasingly used to fuse intelligence, prioritise targets, support battle-damage assessment, and accelerate operational tempo. In parallel, some weapon platforms incorporate higher degrees of autonomy. The International Committee of the Red Cross (ICRC) describes autonomous weapon systems as those that can select and apply force to targets without human intervention after activation, based on sensor information and a generalised "target profile", meaning the user may not choose or even know the specific target(s) at the moment force is applied.[5]

Distinction becomes more difficult when classification is statistical rather than contextual. Under IHL, attackers must distinguish between civilians and combatants, and between civilian objects and military objectives. Human decision-makers interpret context and uncertainty, including patterns of civilian life and the presence of protected objects.
Machine-learning systems, by contrast, infer categories from training data. In complex environments, irregular forces, civilians intermingled with fighters, degraded visibility, adversarial deception systems can misclassify in ways that are operationally subtle but legally decisive. The humanitarian cost is borne by people on the ground: civilians who look "abnormal" to an algorithm, humanitarian staff moving near contested areas, and communities whose daily routines are misread as hostile patterns.

> **"Accountability depends on the ability to establish credible facts, yet contemporary conflicts increasingly unfold in an "epistemic battlespace" where information is contested and deliberately manipulated**

Proportionality and precaution are even less amenable to automation. Proportionality prohibits attacks expected to cause incidental civilian harm that would be excessive in relation to the concrete and direct military advantage anticipated. Even if algorithms can estimate blast effects or probabilities, proportionality still demands a qualitative legal judgement about what is "excessive" in context, based on information reasonably available at the time. Likewise, precaution

4 UN Security Council, Resolution 1540 (2004), S/RES/1540 (adopted on 28 April 2004), UN Digital Library record, last accessed December 18, 2025.
5 ICRC, "ICRC position on autonomous weapon systems" definition of AWS and "target profile" language, last accessed 17 December 2025.

requires all feasible steps to minimise harm and to cancel or suspend attacks when circumstances change.[6] High-tempo machine-assisted targeting can narrow the "reconsideration window", the very space in which human judgement is supposed to correct error, uncertainty, or new civilian presence. A core legal risk is quiet norm-drift: allowing the technical limits of a system to redefine what counts as a "feasible" precaution.

Accountability is where the IHL and 1540 lenses converge. Technical systems can diffuse agency across operators, commanders, developers, contractors, data suppliers, and procurement authorities; opacity can frustrate after-action review. Yet IHL and international criminal law remain anchored in human responsibility and traceable decision-making. Meaningful human control is therefore not a purely ethical add-on; it is a practical way of preserving a chain of responsibility, ensuring that a legally accountable human can understand the basis for action, question it, and stop it when legal doubt arises. The ICRC has explicitly called for regulation of autonomous weapon systems and for preserving human control over the use of force.[7]

From a UNSCR 1540 perspective, traceability also supports prevention. Many enabling components of algorithmic targeting high-resolution sensors, geospatial analytics, communications equipment, drone subsystems, and certain software stacks are dual-use and increasingly accessible through commercial markets. Weak controls can enable diversion and capability transfer to non-State actors. Strengthening review, documentation, and human control is therefore simultaneously a humanitarian safeguard and a preventive measure consistent with 1540 obligations.[8]

Manipulation of facts and the crisis of digital evidence in armed conflict

If targeting is one pressure point, evidence is the other. Accountability depends on the

6 International Humanitarian Law and the Challenges of Contemporary Armed Conflicts: Building a Culture of Compliance for IHL to Protect Humanity in Today's and Future Conflicts, Report to the 34th International Conference of the Red Cross and Red Crescent (Power of Humanity), Geneva, 28–31 October 2024, September 2024, last accessed 17 December 2025.

7 Office of the United Nations High Commissioner for Human Rights (OHCHR), Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law, 2022, last accessed 18 December 2025.

8 UN Security Council, Resolution 1540 (2004), S/RES/1540 (adopted 28 April 2004, last accessed 18 December 2025.

ability to establish credible facts, yet contemporary conflicts increasingly unfold in an "epistemic battlespace" where information is contested and deliberately manipulated. Synthetic media, including deepfakes and AI-generated audio, can fabricate incidents, misattribute attacks, or discredit legitimate reporting.[9] The strategic effect is corrosive: even authentic material becomes suspect, witnesses become easier to discredit, and truth becomes harder to prove in legal proceedings.

Computational propaganda adds scale and persistence. Automated accounts, bot networks and targeted disinformation can amplify false narratives, intimidate witnesses, and undermine trust in humanitarian organisations and investigators. The ICRC's Global Advisory Board report on digital threats frames "harmful information" as a key vector of civilian harm and a challenge for protection during armed conflict.[10]

Cyber operations intensify evidentiary vulnerability in a different way: by attacking the infrastructure that stores and transmits proof. Surveillance logs, metadata, communications records and weapon-system telemetry may be deleted, corrupted, encrypted or altered. When audit trails are erased, accountability can become practically "unprovable", not because harm did not occur, but because the evidentiary chain has been compromised.

In this context, "protecting civilians" includes protecting the informational conditions under which protection and accountability can function. When truth can be manufactured or erased at scale, both humanitarian and preventive security frameworks become harder to operationalise.

Digital evidence, verification and legal admissibility in international practice

International justice institutions and fact-finding mechanisms are adapting to the digital turn, but verification and admissibility remain fragile. Digital evidence raises persistent challenges: authentication (is it what it claims to be?), provenance (who handled it and how?), and preservation (was it altered, compressed, or stripped of metadata?). These are not merely technical questions; they affect fairness, reliability, and the probability that perpetrators are held responsible. The ICC's "Unified Technical Protocol (eCourt Protocol)" reflects the Court's move toward standardised electronic handling and submission of evidence and related materials.[11]

The Berkeley Protocol on Digital Open Source Investigations developed through OHCHR responds to these challenges by offering international standards and guidance on identifying, collecting, preserving, verifying and analysing digital open-source information for international criminal, humanitarian and human rights investigations.[12] Its core contribution is methodological consistency: a shared professional baseline that helps judges, commissions, and investigators evaluate how evidence was obtained and tested.

For investigators and humanitarian actors, verification is concrete work: preserving originals, capturing metadata, geolocating images, cross-checking sources, documenting provenance, and securing storage.

9 OHCHR, Berkeley Protocol on Digital Open Source Investigations, 2022.

10 Protecting Civilians Against Digital Threats During Armed Conflict : Final report of the ICRC's Global Advisory Board on Digital Threats During Armed Conflicts, 19 October 2023.

11 International Criminal Court (ICC), Unified Technical Protocol ("E court Protocol") for the Provision of Evidence, Witness and Victims Information in Electronic Form.

12 OHCHR, Berkeley Protocol on Digital Open Source Investigations, 2022.

Where resources are limited, uneven capacity produces uneven evidentiary quality and, ultimately, uneven access to justice. The ICRC has also highlighted that digital threats target humanitarian organisations themselves, making cybersecurity and information integrity part of protection and operational security.[13]

> **When audit trails are erased, accountability can become practically "unprovable", not because harm did not occur, but because the evidentiary chain has been compromised**

States can support evidentiary resilience by investing in digital forensics, standardising domestic procedures for receiving and protecting digital evidence, and enabling secure cooperation with civil society and technology companies that may hold relevant data. For UNSCR 1540, these investments align with national implementation measures that strengthen investigative capacity and improve prevention and enforcement against non-State actors.[14]

Toward an integrated normative framework: IHL, AI governance and UNSCR 1540

Technological change does not diminish legal responsibility; it increases the need for adaptive governance. The ICRC's 2024 "Challenges Report" underlines how contemporary conflicts raise evolving issues for IHL, including those linked to new technologies.[15] Building on this, a coherent response requires connecting four priorities.

Modernise weapons and methods review. Where Article 36 review obligations apply, they should address AI-enabled systems through realistic testing, evaluation of failure modes, and scrutiny of human–machine interaction, including how systems behave under uncertainty and in civilian-dense environments.[16] Reviews should explicitly consider evidentiary consequences: whether the system's operation will be reconstructible after the fact.

Operationalise meaningful human control. Human control must be substantive: informed operators, real intervention capacity, and command responsibility that does not defer to machine outputs. This is essential to preserve IHL's structure of accountability and civilian protection.[17]

Design for accountability and evidentiary resilience. Where possible, systems and processes should generate auditable records (logs, secure storage, documentation of targeting decisions and data inputs), while protecting sensitive information. These measures support IHL investigations and strengthen States' capacity to detect misuse, diversion, and illicit technical modification.

Align preventive controls with UNSCR 1540 realities. UNSCR 1540 requires States to adopt and enforce effective measures to prevent proliferation to non-State actors,

13 ICRC, Protecting Civilians Against Digital Threats During Armed Conflict, 2023.

14 UN Security Council, Resolution 1540 (2004), S/RES/1540.

15 ICRC, International Humanitarian Law and the Challenges of Contemporary Armed Conflicts (2024) sections addressing evolving challenges/new technologies and IHL.

16 International Humanitarian Law and the Challenges of Contemporary Armed Conflicts Building a Culture of Compliance for IHL to Protect Humanity in Today's and Future Conflicts, September 2024

17 ICRC Position on Autonomous Weapon Systems, Geneva, 12 May 2021.

including appropriate domestic controls.[18] In an era of AI- and cyber-enabled conflict, this implies expanding national risk assessments to include software and data-driven tools; strengthening export, brokering and end-use controls for relevant dual-use technologies; and developing structured cooperation with the private sector, which often holds key technical knowledge and data relevant to prevention.

If law is to remain credible, it must remain operational. The objective is not to "ban technology", but to ensure that innovation does not erode the legal duties that protect civilians and enable responsibility and prevention.

ABOUT THE AUTHOR

Ari Koutale Tila Boulama Mamadou is a diplomat, First Secretary at the Permanent Mission of Niger to the United Nations in New York, and focal point for Niger to the Security Council›s 1540 Committee. He is in charge of disarmament and international security, including counter-terrorism. He also handles legal issues and ICT-related matters. In the academic field, he holds a Master's degree in International Relations, specializing in Security Studies, Conflict Resolution, and Peace Policies, and another Master's degree in Development Communication. He has also obtained several certifications in international human rights law and transnational organized crime. He is currently a doctoral candidate, studying the impact of new digital technologies on international security, including cyber operations and hybrid warfare. He is affiliated with the Centre for African Studies at Leiden University. Previously, he worked for five years on the implementation of a reintegration programme for former Boko Haram child soldiers in Niger.

18 UN Security Council, Resolution 1540 (2004), S/RES/1540.