



**DANIELE AMOROSO**

## **La disciplina dei sistemi d'arma autonomi nel diritto internazionale umanitario. Stato dell'arte e prospettive di regolamentazione**

Il presente contributo persegue un duplice obiettivo: da un lato, offrire una ricostruzione del quadro giuridico oggi applicabile ai sistemi d'arma autonomi, con specifico riferimento alle norme del diritto internazionale umanitario (DIU); dall'altro, tracciare le direttrici lungo le quali potrebbe svilupparsi una futura disciplina internazionale. Dopo aver chiarito la nozione di sistema d'arma autonomo, sono analizzate le norme del DIU attualmente applicabili, dedicando particolare attenzione alle revisioni giuridiche *ex art. 36* del Primo Protocollo Addizionale. Si evidenzia come l'autonomia decisionale di tali sistemi imponga di verificarne non solo la liceità intrinseca, ma anche la capacità di rispettare i principi di distinzione e proporzionalità. Sono poi approfonditi gli obblighi derivanti dal principio di precauzione, in relazione alla progettazione dei sistemi e all'interazione uomo-macchina. Nella seconda parte si esaminano le principali proposte normative in discussione a livello internazionale, con un focus sull'approccio "two-tier" del CICR e sugli sviluppi negoziali più recenti, sia in seno al GGE sia presso l'Assemblea Generale dell'ONU.

*Sistemi d'arma autonomi – Diritto internazionale umanitario – Articolo 36 del Primo Protocollo Addizionale alle Convenzioni di Ginevra – Principi di distinzione, proporzionalità e precauzione – Intelligenza artificiale*

### **The Legality of Autonomous Weapons Systems under International Humanitarian Law: What the Law Says and How It Could Be Improved**

This article pursues a twofold objective: on the one hand, it offers an overview of the legal framework currently applicable to autonomous weapons systems (AWS), with a focus on the rules of international humanitarian law (IHL); on the other hand, it outlines possible trajectories for the development of future international regulation on AWS. After clarifying the notion of AWS, the article examines the IHL norms applicable to their use, with particular emphasis on legal reviews under Article 36 of Additional Protocol I. It highlights how the decision-making autonomy of such systems requires not only an assessment of their intrinsic lawfulness but also of their capacity to comply with the principles of distinction and proportionality. The obligations stemming from the principle of precaution are also addressed, especially regarding the design of AWS and the human-machine interaction. The second part explores the main regulatory proposals currently on the table, zooming in on the ICRC's two-tier approach and the most recent developments within the GGE and the UN General Assembly.

*Autonomous Weapons Systems – International Humanitarian Law – Article 36 of the First Additional Protocol to the Geneva Conventions – Principles of Distinction, Proportionality and Precaution – Artificial Intelligence*

L'Autore è professore associato di Diritto internazionale, Università degli studi di Cagliari

Questo contributo fa parte della sezione monografica *Le due facce della rivoluzione digitale, tra emergenze e tecnologie dual-use*, a cura di Paolo Passaglia

**SOMMARIO:** 1. Introduzione. – 2. Definizione operativa e tipologie di sistemi d'arma autonomi. – 3. La *lex lata*: l'art. 36 del Primo Protocollo Addizionale alle Convenzioni di Ginevra. – 3.1. Il principio di distinzione. – 3.2. Il principio di proporzionalità. – 4. Sempre sulla *lex lata*: il principio di precauzione. – 5. Conclusioni *de lege lata*. – 6. Prospettive *de lege ferenda*: verso uno strumento giuridicamente vincolante? – 6.1. L'approccio "two-tier" proposto dal Comitato internazionale della Croce Rossa. – 6.2. Da Ginevra a New York (e ritorno): il progressivo coinvolgimento dell'Assemblea Generale e il *rolling text*. – 7. Conclusioni.

## 1. Introduzione

È ormai ampiamente noto come i recenti sviluppi nel campo della robotica e dell'intelligenza artificiale (IA) abbiano reso possibile la progettazione di sistemi d'arma in grado di selezionare e colpire obiettivi in modo autonomo, vale a dire senza che sia necessario un ulteriore intervento umano dopo l'attivazione. Sebbene il potenziale tecnologico di tali sistemi non sia ancora stato pienamente esplorato, esistono già oggi applicazioni operative concrete che rientrano a pieno titolo nella categoria dei sistemi d'arma autonomi.

Questa circostanza ha stimolato una produzione scientifica vastissima e multidisciplinare, rendendo il tema uno degli ambiti più affollati del dibattito contemporaneo tra studiosi di diritto internazionale, etica applicata e teoria delle relazioni internazionali<sup>1</sup>. La riflessione sull'accettabilità etico-giuridica di questi sistemi, e ancor più sulla

loro disciplina, ha rapidamente travalicato i confini accademici, diventando oggetto di una vivace e persistente discussione politica e diplomatica.

Un ruolo determinante in questa evoluzione è stato giocato dalla società civile organizzata: la mobilitazione, a partire dal 2013, della Campagna internazionale per la messa al bando dei sistemi d'arma autonomi (*Campaign to Stop Killer Robots*, SKR) ha contribuito in modo decisivo a porre la questione all'ordine del giorno dei principali fora multilaterali<sup>2</sup>. In particolare, le implicazioni normative dell'autonomia dei sistemi d'arma sono da quasi un decennio al centro dei lavori di un Gruppo di esperti governativi (GGE) istituito nel quadro della Convenzione su certe armi convenzionali (CCW), e più recentemente sono approdate anche in seno all'Assemblea Generale delle Nazioni Unite.

Ripercorrere per intero questo dibattito, nelle sue articolazioni tecniche, normative e politiche,

1. Si considerino, per limitarci ai lavori monografici che affrontano la questione nella prospettiva del diritto internazionale: COPELAND 2025, VALDARES FERNANDES BARBOSA 2025, KWIK 2024, DE VRIES 2023, BODE-HUELLS 2022, MAURI 2022, SEIXAS-NUNES 2022, SAXON 2021, AMOROSO 2020, MCFARLAND 2020.

2. CARPENTER 2014, p. 88 ss. V. anche, per ulteriori informazioni, il [sito](#) della Campagna.

risulterebbe non solo impossibile nello spazio assegnato, ma anche superfluo, vista la già estesa letteratura esistente. Il presente lavoro si propone dunque un obiettivo più circoscritto: per un verso, offrire una ricostruzione ragionata del quadro giuridico oggi applicabile ai sistemi d'arma autonomi, con particolare riferimento alle norme del diritto internazionale umanitario (o diritto internazionale dei conflitti armati), che per ragioni evidenti ne costituiscono l'ambito regolamentare d'elezione; per altro verso, tracciare alcune direttrici lungo le quali potrebbe svilupparsi, o consolidarsi, una futura disciplina internazionale, esplorandone l'effettiva praticabilità anche in considerazione delle tensioni che, allo stato attuale, dividono la comunità internazionale.

Dopo aver chiarito, nel par. 2, la nozione di sistema d'arma autonomo e le principali tipologie oggi esistenti o in via di sviluppo, verranno analizzate, nei par. 3 e 4, le norme del DIU che concorrono a definire la disciplina attualmente applicabile ai sistemi d'arma autonomi. Il par. 3, in particolare, si concentrerà sul ruolo delle revisioni giuridiche dei mezzi di combattimento ex art. 36 del Primo Protocollo Addizionale, evidenziando come l'autonomia decisionale di questi sistemi crei una sfida concettuale inedita, in quanto impone di verificare non solo la liceità intrinseca dell'arma, ma anche la sua capacità concreta di operare nel rispetto dei principi che regolano il *targeting*, *in primis* quelli di distinzione e proporzionalità. Il par. 4 approfondirà, invece, gli obblighi derivanti dal principio di precauzione e le implicazioni di tale principio in relazione alla progettazione dei sistemi e all'interazione tra macchina e operatore umano. Il par. 5 proporrà una sintesi degli obblighi che si possono ricavare dal quadro normativo esistente, soffermandosi in particolare sulle condizioni che consentono, o meno, l'impiego lecito di sistemi d'arma autonomi. Le prospettive di regolamentazione *de lege ferenda* saranno invece affrontate nel par. 6, che illustrerà le principali proposte attualmente in

discussione a livello internazionale, con particolare attenzione all'approccio "two-tier" del Comitato internazionale della Croce Rossa (par. 6.1) e ai più recenti sviluppi negoziali sia in seno al GGE sia in ambito ONU (par. 6.2). Il par. 7 conclude con alcune osservazioni sul futuro della regolazione multilaterale in materia.

## 2. Definizione operativa e tipologie di sistemi d'arma autonomi

Un sistema d'arma autonomo può essere definito come un sistema in grado di selezionare e colpire un obiettivo senza che, dopo la sua attivazione, sia necessario un ulteriore intervento umano. Pur non essendo esente da criticità, in particolare per la sua eccessiva inclusività, questa definizione rappresenta un punto di partenza difficilmente eludibile nella discussione, essendo condivisa da una pluralità di attori istituzionali e non governativi coinvolti nel dibattito<sup>3</sup>. Essa è infatti riprodotta – con variazioni marginali – in documenti e dichiarazioni ufficiali, non solo della Campagna per la messa al bando delle armi autonome<sup>4</sup>, ma anche di soggetti portatori di istanze in tutto o in parte divergenti, come il Dipartimento della Difesa degli Stati Uniti<sup>5</sup> e il Comitato internazionale della Croce Rossa<sup>6</sup>.

È importante sottolineare che il ricorso a sistemi di intelligenza artificiale non costituisce un elemento necessario di questa definizione. L'IA non è, cioè, una componente costitutiva dei sistemi d'arma autonomi, bensì un "abilitatore tecnologico" (*enabler*) della loro autonomia operativa, in particolare nei sistemi più sofisticati e adattivi<sup>7</sup>. In tal senso, la distinzione tra armi autonome "con" o "senza" IA può avere rilevanza funzionale, ma non incide sulla qualificazione giuridica di un sistema come autonomo.

La definizione operativa appena richiamata, come si è detto, è particolarmente ampia. Alcuni sistemi che vi rientrano a pieno titolo sono già oggi in uso, seppure in contesti circoscritti e con finalità

3. Per un tentativo di definizione più raffinato e scientificamente accurato, v. TADDEO-BLANCHARD 2022.

4. HUMAN RIGHTS WATCH 2012.

5. DEPARTMENT OF DEFENSE, Directive 3000.09, *Autonomy in Weapons Systems (2012)*, riformata nel 2023.

6. COMITATO INTERNAZIONALE DELLA CROCE ROSSA 2021.

7. ASARO 2019, p. 60.

prevalentemente difensive<sup>8</sup>. È il caso, ad esempio, dei sistemi di difesa aerea come l'*Iron Dome* israeliano, progettati per reagire in modo automatico a minacce in arrivo,<sup>9</sup> o delle tipologie più avanzate di munizioni “fire and forget”, come il *Brimstone* britannico.<sup>10</sup> Un altro esempio frequentemente citato è quello dei robot sentinella sudcoreani SGR-A1 e Super aEgis II<sup>11</sup>: pur essendo tecnicamente in grado di identificare e colpire autonomamente un obiettivo, per quanto è dato sapere questi sistemi vengono attualmente impiegati in modalità semi-autonoma, nella quale l'individuazione dell'obiettivo viene effettuata dalla macchina, ma la decisione di aprire il fuoco resta affidata a un operatore umano.<sup>12</sup> Particolarmente rilevanti sono poi le cosiddette munizioni circuitanti (*loitering munitions*), come l'*Harpy NG* israeliano, un drone-suicida capace di pattugliare un'area per diverse ore alla ricerca di sorgenti di onde radio su cui abbattersi, senza che l'operatore umano debba intervenire oltre l'attivazione iniziale<sup>13</sup>.

Altri sistemi, attualmente in fase di sperimentazione, sembrano destinati a potenziare ulteriormente l'autonomia decisionale in ambito operativo. Si pensi ai prototipi di caccia autonomi controllati da IA, capaci di ingaggiare vittoriosamente duelli aerei con piloti esperti<sup>14</sup>; o alle ricerche nel campo della *swarm robotics*, che stanno aprendo la strada allo sviluppo dei cd. “sciame” di droni armati, di piccole dimensioni e a basso costo, il cui spiegamento in numero massivo – nell'ordine delle centinaia, se non delle migliaia – dovrebbe consentire di sopraffare le difese aeree nemiche attraverso attacchi di saturazione improvvisi e coordinati<sup>15</sup>.

### 3. La *lex lata*: l'art. 36 del Primo Protocollo Addizionale alle Convenzioni di Ginevra

L'analisi delle norme internazionali applicabili ai sistemi d'arma autonomi, in quanto tecnologie emergenti, non può che prendere le mosse dall'art. 36 del Primo Protocollo Addizionale alle Convenzioni di Ginevra del 1977. Questa disposizione, significativamente rubricata “Nuove armi”, sancisce infatti l'obbligo per le Parti Contraenti di stabilire, “[i]n caso di studio, sviluppo, acquisizione o adozione di una nuova arma”, se il suo impiego sia, “in determinate o in tutte le circostanze, vietato dal [...] Protocollo o da qualsiasi altra norma di diritto internazionale applicabile”.

In quanto disposizione pattizia, l'art. 36 vincola solo gli Stati che hanno ratificato il Primo Protocollo. Va tuttavia rilevato che, tra la manciata di Stati che non sono parte del Protocollo, alcuni tra i più rilevanti dal punto di vista militare – come gli Stati Uniti<sup>16</sup> e Israele<sup>17</sup> – conducono comunque regolarmente revisioni giuridiche dei nuovi mezzi di combattimento<sup>18</sup>. La rilevanza dell'obbligo in questione per i sistemi d'arma autonomi, del resto, è stata espressamente riconosciuta nei Principi Guida adottati per *consensus* dal GGE sui sistemi d'arma autonomi letali, che incorporano – quasi *ad litteram* – il testo dell'art. 36<sup>19</sup>.

Si tratta, all'evidenza, di un obbligo procedurale, volto a garantire che l'introduzione di nuove tecnologie belliche nell'arsenale statale avvenga nel rispetto del diritto internazionale vigente. Tradizionalmente, tale revisione giuridica mira ad accertare la conformità del nuovo sistema con le norme

8. BOULANIN-VERBRUGGEN 2017.

9. *Ivi*, p. 36 ss.

10. *Ivi*, p. 47 ss.

11. *Ivi*, p. 44 ss.

12. PARKIN 2015.

13. BOULANIN-VERBRUGGEN 2017, p. 50 ss.

14. LOSEY 2024.

15. EKELHOF-PERSI PAOLI 2020.

16. *Army regulation 27-53* (1979), riformata da ultimo nel 2019.

17. JEVGLEVSKAJA 2018, p. 209.

18. Tale circostanza ha suggerito che l'obbligo in questione abbia in realtà natura consuetudinaria (JEVGLEVSKAJA 2018).

19. *Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, 25 settembre 2019, CCW/GGE.1/2019/3, Annex IV, Guiding Principles, Principle (e).

del diritto internazionale umanitario che vietano o disciplinano l'uso di determinate categorie di armi – si pensi, ad esempio, alla messa al bando dei laser accecanti o al divieto delle armi che causano mali superflui o sofferenze inutili. Nel caso dei sistemi d'arma autonomi, tuttavia, la situazione risulta concettualmente più complessa. Come osservato dalla delegazione svizzera nei dibattiti che hanno preceduto l'istituzione del GGE, tali sistemi sono concepiti per portare a termine l'intero ciclo di *targeting* senza intervento umano<sup>20</sup>. Ciò comporta un ampliamento dell'ambito della revisione giuridica: non è sufficiente interrogarsi sulla liceità dell'arma in sé, ma occorre valutare se, e in quali condizioni, essa sia in grado di operare nel rispetto delle norme che regolano il *targeting* secondo il diritto internazionale umanitario<sup>21</sup>.

Ne deriva, pertanto, un arricchimento sostanziale dei criteri da considerare nella revisione giuridica. In particolare, occorre valutare se il sistema sia, in concreto, capace di operare nel rispetto dei principi fondamentali del diritto dei conflitti armati: il principio di distinzione e il principio di proporzionalità<sup>22</sup>. Il terzo cardine del diritto del *targeting* – il principio di precauzione – sarà invece esaminato più avanti, in quanto pone questioni che meritano un approfondimento separato.

Prima di soffermarsi sui problemi specifici posti dall'autonomia dei sistemi d'arma rispetto ai principi appena citati, occorre richiamare un ulteriore aspetto rilevante degli obblighi di revisione giuridica. Tanto l'art. 36 quanto i Principi Guida chiari-scono che un nuovo mezzo di combattimento (e, dunque, un sistema d'arma autonomo) può essere vietato “*in alcune* o in tutte le circostanze”. Tale formulazione consente una valutazione graduata, che non si esaurisce nell'analisi delle caratteristiche intrinseche del sistema, ma considera anche le condizioni concrete del suo impiego. Ne consegue che l'impiego di un sistema d'arma autonomo può essere ritenuto lecito in presenza di determinate condizioni, pur risultando inammissibile in altre.

### 3.1. Il principio di distinzione

Il principio di distinzione rappresenta uno dei cardini del diritto internazionale umanitario e impone alle parti in conflitto l'obbligo di distinguere, in ogni circostanza, tra civili e combattenti, nonché tra beni civili e obiettivi militari. L'art. 48 del Primo Protocollo Addizionale alle Convenzioni di Ginevra del 1977 stabilisce la regola generale secondo cui “le Parti in conflitto devono in ogni momento distinguere tra popolazione civile e combattenti, nonché tra beni civili e obiettivi militari, e dirigere le loro operazioni soltanto contro obiettivi militari”. Questa disposizione è ripresa e specificata negli articoli 51(2) e 52(1) del medesimo Protocollo, che disciplinano rispettivamente la protezione della popolazione civile e dei beni civili.

A queste norme si affiancano altre disposizioni che, pur non essendo volte alla protezione dei civili, pongono analoghi problemi di distinzione, come il divieto di attaccare persone *hors de combat* (art. 41(2) Primo Protocollo), ovvero i combattenti della parte avversa che hanno manifestato l'intenzione di arrendersi o si trovano in stato di incapacità a causa di ferite, naufragio o malattia, a condizione che si astengano da atti ostili e non tentino la fuga<sup>23</sup>.

Il contenuto normativo del principio, almeno nella sua formulazione astratta, appare chiaro. Tuttavia, la sua applicazione concreta è resa complessa dal persistere di zone grigie, in particolare nel contesto degli scenari operativi urbani. In queste situazioni, i combattenti si confondono con la popolazione civile ed i beni civili sono spesso impiegati a fini militari. Si considerino, in particolare, la partecipazione diretta alle ostilità da parte di civili (art. 51(3) del Protocollo), che determina la sospensione temporanea della loro protezione o la qualificazione degli oggetti cd. *dual use* come obiettivi militari legittimi, in virtù della loro destinazione o impiego effettivo, secondo la definizione contenuta nell'art. 52(2) del Protocollo.

20. Svizzera, *Towards a “compliance-based” approach to LAWS*, CCW Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Ginevra, 30 marzo 2016.

21. COPELAND 2025, p. 152.

22. *Ivi*, p. 245.

23. Per una disamina più ampia delle altre norme del DIU che pongono problemi di distinzione v. AMOROSO 2020, p. 56 ss.

Questa complessità si riverbera inevitabilmente sulla capacità dei sistemi d'arma autonomi di rispettare il principio di distinzione. Non vi è dubbio, infatti, che alcuni compiti di identificazione presupposti dal principio di distinzione possano essere eseguiti, con un livello di affidabilità accettabile, sulla base delle tecnologie esistenti. Questo è senz'altro il caso del riconoscimento di oggetti "cooperativi" – ossia quelli che emettono segnali rilevabili, come radar o transponder: l'Harpy NG israeliano, per fare un esempio relativamente noto, è in grado di identificare gli obiettivi su cui farsi esplodere sulla sola base delle onde radio da questi emesse.<sup>24</sup> Lo stesso vale per il riconoscimento degli oggetti che sono obiettivi militari *per natura*, come carri armati, elicotteri o missili. Sia i sistemi di difesa aerea, sia le munizioni "fire and forget" più sofisticate (come il *Brimstone* britannico) sono in grado di identificare autonomamente questa tipologia di obiettivi.<sup>25</sup>

Va detto, peraltro, che diverse norme del DIU mirano proprio a facilitare l'identificazione degli obiettivi legittimi nonché di persone e oggetti protetti. Alcune norme, in particolare, promuovono la "cooperazione" degli oggetti protetti affinché non vengano colpiti erroneamente. L'Allegato I del Primo Protocollo, ad esempio, contiene disposizioni molto dettagliate che regolano l'emissione di specifici segnali luminosi, radio ed elettronici da parte di aeromobili, navi e veicoli medici, al fine di renderli facilmente individuabili dalle Parti in conflitto.<sup>26</sup> Altre norme, invece, richiedono l'esibizione di segni distintivi facilmente distinguibili per segnalare che un certo oggetto è protetto dal DIU. Si pensi, a questo proposito, alle disposizioni che regolano l'uso dell'emblema della Croce Rossa (e simili) per il servizio medico, dello Scudo blu per i beni culturali e dei tre cerchi arancioni che segnalano "opere e installazioni contenenti forze pericolose"<sup>27</sup>. Lo stesso vale per l'obbligo

dei combattenti "di distinguersi dalla popolazione civile [ad es. indossando uniformi] mentre sono impegnati in un attacco o in un'operazione militare in preparazione di un attacco"<sup>28</sup>. Questi obblighi, pur essendo originariamente concepiti per facilitare l'applicazione del principio di distinzione da parte dei combattenti umani, finiscono per semplificare i compiti percettivi assegnati ai sistemi d'arma autonomi, che dovranno essere progettati in modo da essere in grado di riconoscere i segnali (luminosi o di altro tipo) provenienti da oggetti protetti o di identificare i segni distintivi.

Anche alcuni compiti più complessi, ma comunque assistiti da marcatori visivi o comportamentali, possono teoricamente essere affrontati da sistemi autonomi. Si pensi all'identificazione di un combattente che manifesta l'intenzione di arrendersi, a condizione che tale gesto sia chiaramente codificato (ad es., le mani alzate). Secondo alcuni report, la sentinella robotica SGR-A1 sarebbe capace di riconoscere segnali di resa<sup>29</sup>. Va tuttavia sottolineato, però, che tale sistema opera in un contesto, la zona demilitarizzata tra le due Coree, caratterizzato dalla sostanziale assenza di elementi che possano disturbare il riconoscimento autonomo.

Molto più problematici sono, invece, i compiti che implicano valutazioni contestuali e interpretazioni dinamiche. Si pensi al riconoscimento di civili che partecipano direttamente alle ostilità, alla qualificazione di un oggetto come obiettivo militare per uso o scopo (ad esempio, una scuola adibita a deposito di munizioni) o, più in generale, all'applicazione del principio di distinzione in scenari di guerriglia urbana, caratterizzati dalla compresenza di obiettivi legittimi e persone/beni protetti. In questi casi, il rispetto del principio di distinzione presuppone non solo la percezione degli elementi fisici dell'ambiente, ma anche la comprensione della loro funzione, del contesto operativo e dell'evoluzione della situazione – competenze che, allo

24. SCHARRE 2018, p. 84.

25. BOULANIN-VERBRUGGEN 2017, rispettivamente pp. 37-39 e 49-50.

26. Artt. 6-8, 10-11.

27. V., rispettivamente, art. 44 della Prima Convenzione di Ginevra del 1949, artt. 16-17 della Convenzione dell'Aja del 1954 sulla protezione dei beni culturali nei conflitti armati e art. 56, co. 7, del Primo Protocollo.

28. Art. 44, co. 3, del Primo Protocollo.

29. BOULANIN-VERBRUGGEN 2017, p. 45.

stato attuale, sono ancora al di fuori della portata dell'IA<sup>30</sup>.

Tutto ciò non esclude che sviluppi tecnologici futuri possano estendere la capacità valutativa dei sistemi autonomi a scenari oggi considerati fuori portata. L'impiego di architetture basate su *deep learning* e l'addestramento su dataset estesi e realistici potranno consentire una sofisticazione percettiva e inferenziale sempre maggiore. Tuttavia, il salto qualitativo richiesto per colmare il divario tra cognizione algoritmica e giudizio umano in questi contesti operativi resta notevole, e al momento non vi sono evidenze concrete che tale traguardo sia prossimo.

### 3.2. Il principio di proporzionalità

Il principio di proporzionalità, codificato all'art. 51(5)(b) e ribadito all'art. 57(2)(a)(iii) del Primo Protocollo Addizionale, vieta quegli attacchi che ci si attende provochino incidentalmente danni ai civili o a beni civili (c.d. danni collaterali) eccessivi rispetto al vantaggio militare concreto e diretto previsto. Un attacco, pertanto, non diventa illecito per il semplice fatto di causare danni collaterali. Piuttosto, in base al principio di proporzionalità, la sua legittimità è subordinata ad una valutazione *ex ante*, fondata su tre passaggi logici:

- la stima del danno collaterale previsto (in vite umane e beni civili);
- la valutazione del vantaggio militare atteso;
- la determinazione dell'eccessività del primo rispetto al secondo.

Di questi tre passaggi, l'unico che si presta agevolmente ad essere gestito da sistemi autonomi è la stima del danno collaterale. Le forze armate, a partire da quelle statunitensi, fanno già ampio uso di sistemi software per stimare, in sede di pianificazione, l'impatto di un attacco su persone, edifici e infrastrutture civili<sup>31</sup>. Questi strumenti permettono infatti di calcolare, con un certo grado di affidabilità, la portata

dell'esplosione, la propagazione delle schegge, il danno strutturale indotto su edifici contigui e il rischio per persone presenti in aree adiacenti.

Più complessa è la valutazione del vantaggio militare. A differenza di quanto avviene per il principio di distinzione, non è sufficiente stabilire che un determinato bene o una persona costituisce un obiettivo legittimo perché la sua eliminazione o distruzione offre “un vantaggio militare concreto”. Chi decide l'attacco, infatti, dovrà attribuire un valore al vantaggio militare in questione, al fine di confrontarlo con i danni stimati<sup>32</sup>. Qui entrano in gioco fattori non sempre facili da quantificare e soggetti a variazione nel tempo e nello spazio: tra gli altri, l'importanza tattica o strategica del *target*, la fase del conflitto, o il grado di certezza circa il successo dell'azione.

Il passaggio più resistente ad una codificazione algoritmica è però rappresentato dal cuore del principio di proporzionalità: la valutazione dell'eccessività dal danno collaterale. Tale giudizio implica infatti un bilanciamento tra valori disomogenei, come il numero di civili potenzialmente colpiti e l'importanza tattica di un'infrastruttura militare<sup>33</sup>. Di conseguenza, il principio di proporzionalità non si presta ad un'applicazione “binaria”, come avviene per il principio di distinzione, per cui – nonostante le difficoltà che abbiamo visto sopra – è almeno astrattamente possibile stabilire se un bene o una persona siano protetti o meno dal DIU<sup>34</sup>.

Lo standard elaborato nella prassi per orientare l'applicazione del principio di proporzionalità, quello del “comandante militare ragionevole”<sup>35</sup>, si limita invero a definire una zona di ragionevole discrezionalità, che ricomprende una pluralità di decisioni sulla proporzionalità – alcune più attente alla protezione dei civili, altre più orientate alla massimizzazione del vantaggio militare – tutte considerate legittime<sup>36</sup>.

30. HUMAN RIGHTS WATCH 2025, p. 19; COMITATO INTERNAZIONALE DELLA CROCE ROSSA 2024, p. 65.

31. THURNHER 2014, p. 222.

32. HOLLAND 2004, p. 54.

33. FENRICK 1982, p. 94.

34. CANNIZZARO 2014, p. 332.

35. Sul quale v., anche per ulteriori riferimenti, HENDERSON-REECE 2018, p. 845.

36. In questo senso, v. Corte suprema israeliana, *Beit Sourik Village Council v. The Government of Israel*, HCJ 2056/04, 30 maggio 2004, parr. 42 e 46. La scelta tra queste opzioni, anche se libera da vincoli giuridici, è evidentemente gravida di conseguenze e implicazioni morali (KENNEDY 2004, pp. 290-291).

La difficoltà di tradurre in un codice di programmazione uno standard tanto vago nei suoi contenuti applicativi è evidente. Non deve dunque sorprendere che anche autori favorevoli a un impiego robusto dei sistemi d'arma autonomi siano disposti ad ammettere che la valutazione dell'eccessività debba rimanere affidata al decisore umano, che la effettuerebbe in fase di pianificazione dell'attacco<sup>37</sup>.

#### 4. Sempre sulla *lex lata*: il principio di precauzione

Oltre all'art. 36 del Primo Protocollo Addizionale, un'altra norma fondamentale di riferimento, ancora in una prospettiva *de lege lata*, è rappresentata dal principio di precauzione, che impone alle Parti in conflitto di adottare tutte le misure *practicabili* per garantire il rispetto dei principi di distinzione e proporzionalità.

Il principio, codificato all'art. 57 del Primo Protocollo, si traduce essenzialmente in una serie di obblighi positivi, volti a prevenire violazioni del diritto umanitario. Tra questi, rilevano in particolare:

- l'obbligo di fare tutto il possibile per assicurare la corretta identificazione del target come obiettivo militare<sup>38</sup>;
- l'obbligo di fare tutto il possibile per mantenere il controllo sull'esecuzione dell'attacco, così da poterlo annullare o sospendere quando ciò sia necessario per evitare un attacco diretto contro la popolazione civile o danni collaterali eccessivi<sup>39</sup>.

Ciò ha importanti implicazioni per un modello normativo di relazione tra operatori umani e sistemi d'arma. Anzitutto, come si è visto, il principio di precauzione obbliga le Parti in conflitto a mobilitare tutti i mezzi a loro disposizione, ove

ragionevolmente disponibili, per facilitare la corretta qualificazione del *target* come obiettivo legittimo<sup>40</sup>. Se, dunque, il coinvolgimento di operatori umani è idoneo a favorire una maggiore precisione negli attacchi, la loro inclusione nel processo decisionale deve considerarsi obbligatoria in virtù del principio di precauzione. A questo fine, i sistemi devono essere progettati in modo da consentire ad operatori adeguatamente formati di controllarne e limitarne gli effetti.

Il principio di precauzione sollecita quindi una riflessione sul modo in cui le macchine prendono decisioni e, soprattutto, su come possono sbagliare. Anche i sistemi di IA più avanzati, infatti, continuano ad essere soggetti a errori imprevedibili e controintuitivi, talvolta difficilmente comprensibili perfino da coloro che li hanno programmati e istruiti<sup>41</sup>. Si tratta di errori diversi da quelli umani, che non derivano da stanchezza o stress, ma da fallimenti percettivi o cognitivi legati al cosiddetto "gap semantico": l'IA non comprende il significato delle informazioni che elabora, ma effettua correlazioni tra moli enormi di dati a velocità sovrumana<sup>42</sup>.

Questo rende l'IA vulnerabile a distorsioni di contesto, illusioni percettive o manipolazioni intenzionali (*adversarial attacks*)<sup>43</sup> – vulnerabilità cui l'operatore umano è in larga misura immune. Se si considera che, nel contesto di un conflitto armato, questi errori possono tradursi in attacchi diretti contro la popolazione civile (o comunque in gravi danni collaterali), appare evidente quanto sia poco sensato affidare il ciclo di *targeting* interamente alla macchina, così come sarebbe irragionevole escluderla *a priori*. Il principio di precauzione richiede invece un modello di interazione sinergica tra operatore umano e sistema, in cui ciascuno compensi i limiti dell'altro<sup>44</sup>. Naturalmente, ciò

37. V., tra gli altri, ZAJAC 2023; HENDERSON-KEANE-LIDDY 2017, p. 352, SCHMITT-THURNHER 2013, p. 256.

38. Sul punto, si veda la preziosa codificazione del DIU consuetudinario pubblicata dal CICR (HENCKAERTS-DOSWALD-BECK 2005, Regola 16).

39. *Ivi*, Regola 19.

40. MANCINI 2006, p. 278.

41. LIU-WEI-LIU-DAVIS 2025.

42. COMITATO INTERNAZIONALE DELLA CROCE ROSSA 2019, p. 20.

43. Si noti che tali manipolazioni delle immagini sono spesso impercettibili (o comunque irrilevanti) per un essere umano (CHOWDHURY 2025).

44. SHARKEY 2016, p. 23 ss.

non implica un coinvolgimento umano irrealistico: in alcuni contesti, come quello dei sistemi di difesa antimissile (ad esempio, *Iron Dome*), la previsione di un controllo umano accurato sulle singole decisioni di *targeting* comporterebbe rischi inaccettabili per la sicurezza delle persone. Ma proprio perché gli obblighi derivanti dal principio di precauzione sono condizionati alla praticabilità (*feasibility*), il diritto internazionale non esige l'impossibile.

## 5. Conclusioni *de lege lata*

Dall'analisi fin qui svolta emerge con sufficiente chiarezza che il diritto internazionale umanitario, pur non contenendo norme esplicitamente dedicate ai sistemi d'arma autonomi, dispone già di strumenti adeguati a valutarne la liceità. L'art. 36 del Primo Protocollo Addizionale, come si è detto, rappresenta il primo snodo obbligato: in quanto clausola di verifica preventiva, esso impone agli Stati di stabilire se l'introduzione di un nuovo sistema d'arma sia compatibile con le norme esistenti del diritto internazionale, *in alcune o in tutte le circostanze*. Con riferimento ai sistemi d'arma autonomi, ciò presuppone una verifica della capacità di questi sistemi di svolgere il ciclo di *targeting* nel rispetto dei principi di distinzione e proporzionalità. La conclusione cui si è giunti, sulla base delle tecnologie esistenti e note, è che l'impiego di sistemi d'arma autonomi supera il test di legalità di cui all'art. 36 soltanto in presenza di alcune condizioni.

Nei contesti operativi più problematici – in particolare, in scenari urbani caratterizzati dalla prosimità fisica tra obiettivi legittimi e beni protetti – il loro utilizzo richiede infatti estrema cautela. In tali situazioni, la conformità ai principi di distinzione e proporzionalità presuppone valutazioni complesse e contestuali, che allo stato attuale solo un operatore umano può svolgere in modo attendibile. Di conseguenza, la legalità dell'impiego di sistemi autonomi in questi contesti può essere garantita solo a condizione che le decisioni critiche siano assunte in fase di pianificazione, sotto la responsabilità di un comandante umano.

Il ricorso a forme di *targeting* dinamico (vale a dire, non pianificato) da parte del sistema d'arma può essere ipotizzato, in linea di principio, solo in contesti operativi molto specifici, come ambienti desertici o marittimi, in cui:

- l'assenza di civili sia ragionevolmente certa, oppure
- i civili siano chiaramente separati, in termini spaziali o temporali, dagli obiettivi militari da colpire.

In ogni caso, indipendentemente dallo scenario operativo, il diritto internazionale umanitario richiede, in virtù del principio di precauzione, che sia mantenuta una forma di supervisione umana, almeno come *fail-safe*, cioè come meccanismo di sicurezza capace di bloccare o interrompere l'azione del sistema per evitare danni a persone e beni civili<sup>45</sup>.

Affinché il coinvolgimento umano possa svolgere effettivamente questo ruolo, tuttavia, è necessario rispettare due ulteriori requisiti. Da un lato, la formazione degli operatori deve mirare a preservarne il senso critico, ponendo l'accento non solo sulle capacità ma anche sui limiti della macchina, al fine di evitare che soccombano al *cd. bias* dell'automazione. Dall'altro, i sistemi devono essere progettati con un'interfaccia che privilegi la trasparenza e la spiegabilità delle decisioni, così da rendere effettiva l'interazione uomo-macchina anche in condizioni di elevato stress operativo<sup>46</sup>.

## 6. Prospettive *de lege ferenda*: verso uno strumento giuridicamente vincolante?

La disciplina dei sistemi d'arma autonomi resta dunque affidata, per il momento, all'interpretazione di norme generali del diritto internazionale umanitario, in particolare l'art. 36 del Primo Protocollo Addizionale e i principi che governano i processi di *targeting* (distinzione, proporzionalità e precauzione). Nonostante l'assenza di uno strumento giuridico dedicato, da queste norme è possibile ricavare, seppure con qualche sforzo esegetico, una serie di obblighi specifici relativi all'impiego di tali tecnologie che sembrano delineare un quadro regolatorio coerente e, tutto sommato, soddisfacente. Viene

45. SCHARRE 2016, p. 154. La compatibilità tra autonomia operativa e supervisione umana è stata, del resto, espressamente riconosciuta nella Direttiva 3000.09 del Dipartimento della Difesa, che include nella nozione di sistemi d'arma autonomi i sistemi "autonomi con supervisione dell'operatore" (*operator-supervised autonomous weapons systems*). V. Direttiva 3000.09, cit., p. 21.

46. Su questi aspetti, v. AMOROSO 2020, p. 246 ss.

allora da interrogarsi sulle ragioni – e i meriti – dell'insistenza, da parte della società civile e di un numero sempre crescente di Stati, sulla necessità di adottare uno strumento giuridico *ad hoc*, sia esso un protocollo aggiuntivo alla CCW o un trattato autonomo. Un'insistenza che ha trovato una significativa eco istituzionale nei ripetuti appelli del Segretario Generale delle Nazioni Unite a concludere entro il 2026 i negoziati su uno strumento normativo in materia<sup>47</sup>.

A parere di chi scrive, vi sono almeno tre considerazioni che concorrono a giustificare l'adozione di un nuovo strumento giuridico vincolante.

In primo luogo, uno strumento *ad hoc* potrebbe rendere espliciti e operativamente chiari quegli obblighi che oggi possono essere solo ricavati in via interpretativa. Chi impiega un sistema d'arma autonomo, come qualsiasi militare, ha bisogno di regole che siano immediatamente applicabili senza dover ricorrere a complesse mediazioni ermeneutiche. Un nuovo trattato, inoltre, potrebbe introdurre regole non desumibili dalle norme esistenti, ma fondate su istanze etiche condivise a livello internazionale. È in questo spazio che si colloca la proposta di vietare lo sviluppo e l'impiego di sistemi d'arma autonomi “anti-persona”. Tale proposta poggia, infatti, su preoccupazioni legate al rispetto della dignità umana e al ruolo insostituibile della coscienza morale nell'uso della forza letale<sup>48</sup>. Infine, un trattato dedicato permetterebbe di creare un quadro istituzionale stabile, in grado di garantire non solo l'adozione di regole condivise, ma anche la loro attuazione effettiva attraverso meccanismi di monitoraggio, trasparenza, revisione tra pari e *confidence-building*.

### 6.1. L'approccio “two-tier” proposto dal Comitato internazionale della Croce Rossa

Nel 2021, all'esito di una lunga riflessione, il Comitato internazionale della Croce Rossa (CICR) ha

pubblicato un *position paper* con una proposta di regolamentazione dei sistemi d'arma autonomi che, per la chiarezza della formula e l'autorevolezza della fonte, ha fortemente influenzato il dibattito successivo<sup>49</sup>. L'approccio promosso dal CICR si articola su due livelli (*two-tier approach*): da un lato, il divieto assoluto di sistemi d'arma autonomi “anti-persona” (ovvero: progettati per identificare e colpire esseri umani) e di quelli intrinsecamente imprevedibili; dall'altro, una serie di restrizioni operative volte a garantire un controllo umano contestualmente adeguato sui sistemi d'arma autonomi non coperti dai divieti.

Le restrizioni proposte dal CICR includono: limiti alla tipologia di *target* che possono essere ingaggiati autonomamente (in particolare, soltanto oggetti che costituiscono obiettivi militari *per natura*); limiti alla durata, all'ampiezza geografica e all'entità del loro impiego; limiti relativi al contesto operativo, con la previsione dell'utilizzo di tali sistemi soltanto in teatri caratterizzati dall'assenza di civili o comunque da una netta separazione tra questi e gli obiettivi legittimi. Inoltre, i sistemi – pur essendo autonomi – dovranno essere sempre soggetti alla supervisione di un operatore umano, che avrà la possibilità di interrompere il funzionamento dell'arma esercitando il veto sulle decisioni di *targeting*.

È interessante notare come la proposta del CICR non utilizzi espressamente la nozione di “controllo umano significativo”, che pure aveva orientato (e ancora in parte orienta) la discussione su questo tema<sup>50</sup>. Se questo è vero, è altresì vero che l'approccio *two-tier* è ampiamente compatibile con tale nozione, potendone anzi costituire un'articolazione concreta. Vale la pena osservare, inoltre, che i limiti operativi indicati dal Comitato riflettono e specificano gli obblighi che abbiamo ricavato a partire dall'art. 36 del Primo Protocollo e dal principio di precauzione. L'innovazione più rilevante della proposta del CICR è rappresentata

47. V., da ultimo, *'Politically unacceptable, morally repugnant': UN chief calls for global ban on 'killer robots'*, 14 May 2025, UN News.

48. Su questi aspetti, v. MAURI 2022, SAXON 2021.

49. COMITATO INTERNAZIONALE DELLA CROCE ROSSA 2021.

50. La nozione di controllo umano significativo era stata originariamente introdotta nel dibattito dall'ONG Article36 (ARTICLE36 2013). Più di recente, ad essa si è fatto riferimento anche in ambiti diversi da quello militare (ad esempio, veicoli autonomi, robot chirurgici, etc.). Su questo, v. MECACCI-AMOROSO-CAVALCANTE SIEBERT et al. 2024.

infatti dal divieto generalizzato dei sistemi d'arma autonomi anti-persona, giustificato – come si è detto – da considerazioni etiche, legate alla tutela della dignità umana e all'impossibilità di delegare ad un agente artificiale il potere di decidere della vita e della morte di un essere umano.

## 6.2. Da Ginevra a New York (e ritorno): il progressivo coinvolgimento dell'Assemblea Generale e il *rolling text*

Grazie ad una tenace attività di lobbying da parte della società civile, negli ultimi anni il dibattito diplomatico sui sistemi d'arma autonomi è progressivamente uscito dai confini del GGE di Ginevra, per trovare spazio anche nel più inclusivo foro dell'Assemblea Generale delle Nazioni Unite. Anche se la discussione a New York non ha ancora prodotto sviluppi di rilievo, molto preziosa si è rivelata l'iniziativa, decisa con la Risoluzione 78/241, di affidare al Segretario Generale il compito di raccogliere e sintetizzare le posizioni degli Stati sulla regolamentazione dei sistemi d'arma autonomi<sup>51</sup>.

L'iniziativa ha avuto successo: numerosi Stati hanno trasmesso le proprie osservazioni, tra cui l'Italia, che ha per la prima volta definito in modo relativamente chiaro la propria posizione ufficiale<sup>52</sup>. Quest'ultima presenta numerose convergenze con la proposta del CICR e riproduce, di fatto, un approccio *two-tier*, con alcune differenze rilevanti. In particolare, non si fa alcun riferimento alla necessità di vietare i sistemi d'arma autonomi anti-persona. Questa assenza appare tuttavia in parte compensata dal fatto che, secondo il governo italiano, un sistema d'arma può considerarsi lecito solo se la decisione ultima sull'uso della forza (“the final crucial function”) – e cioè “the decision of whether or not to apply force to a previously identified and selected target” – rimane affidata a un operatore umano<sup>53</sup>. Nella misura in cui un sistema non può ingaggiare un obiettivo già identificato

senza l'autorizzazione di un operatore, ci si colloca chiaramente al di fuori della categoria dei sistemi autonomi propriamente detti, potendosi parlare al più di sistemi semi-autonomi.

L'affiancamento dell'Assemblea Generale al GGE come foro di discussione ha avuto dunque il duplice merito di includere nel dibattito quegli Stati che non sono parte della CCW e di contribuire, grazie al prezioso lavoro di compilazione del Segretariato, alla sistematizzazione delle posizioni emerse nel dibattito, sia individuali sia espresse nel quadro di coalizioni più ampie.

Anche a Ginevra, però, vi è uno sviluppo che vale la pena segnalare. Facciamo riferimento, in particolare, alla circolazione di un “testo in evoluzione” (*rolling text*), redatto su impulso dell'attuale chair del GGE, l'ambasciatore olandese Robert inden Bosch. Il documento è strutturato in cinque box tematici: I (definizioni), II (principi generali), III (divieti e restrizioni operative), IV (misure preventive), V (responsabilità)<sup>54</sup>. Nelle intenzioni del chair, questo documento mira a dare concretezza al dibattito “costringendo” le delegazioni a confrontarsi su un testo concreto, auspicabilmente destinato a confluire in un Protocollo al CCW (o in altro strumento vincolante), invece di continuare ad arrovellarsi in dibattiti inconcludenti su questioni di principio<sup>55</sup>.

Nella sua versione più recente, il testo adotta un impianto solo parzialmente riconducibile all'approccio *two-tier*. Sotto il profilo dei divieti, infatti, ci si limita a ribadire norme la cui applicabilità ai sistemi d'arma autonomi è del tutto pacifica (ad esempio, il divieto di causare mali superflui e sofferenze inutili), mentre non vi sono riferimenti al divieto di sistemi d'arma autonomi anti-persona. Molto più interessante, invece, è la previsione del requisito del “giudizio e controllo umano appropriati al contesto”, il cui contenuto è chiarito attraverso l'indicazione di una serie di misure volte a garantire tale controllo (e giudizio), tra cui la

51. *Lethal autonomous weapons systems*, 22 December 2023, UN Doc. A/RES/78/241, par. 2.

52. *Lethal autonomous weapons systems. Report of the Secretary-General*, 1 July 2024, UN Doc. A/79/88, p. 63 ss.

53. *Ivi*, p. 64.

54. GGE on LAWS, *Rolling text, status date: 12 May 2025*. I titoli dei box tematici sono stati aggiunti dall'autore sulla base del loro contenuto.

55. Si tratta di una scelta che, almeno nel breve periodo, ha già dato i suoi frutti, stimolando una discussione vivace e proficua, dopo anni di stallo (VARELLA 2025).

prevedibilità e tracciabilità degli effetti; la limitazione dei parametri operativi (*target*, durata, area geografica); la previsione di dispositivi di disattivazione automatica; il controllo sulle capacità di auto-apprendimento<sup>56</sup>.

Permangono tuttavia ambiguità. Il testo non chiarisce cosa renda “appropriato” il controllo umano in un determinato contesto, né quali prerogative decisionali debbano essere assicurate agli operatori umani (mera pianificazione? Supervisione? Diritto di veto? Approvazione delle singole decisioni di *targeting*?). A questo fine, potrebbe essere utile riprendere la proposta, formulata in altra sede<sup>57</sup>, di modulare il livello di controllo umano normativamente richiesto alla luce di tre fattori:

- *cosa* deve fare il sistema, ovvero la natura dei compiti (offensivi/difensivi) e dei target assegnati (ad esempio, obiettivi militari per natura oppure membri di un gruppo armato di opposizione);
- *dove* verrà impiegato, ovvero le caratteristiche dell'ambiente operativo (ad esempio, presenza/assenza di civili);
- *come* funziona il sistema, ovvero le capacità senso-motorie e quelle di elaborazione delle informazioni che il sistema mette in atto per eseguire la missione assegnata (ad esempio, capacità di *swarming*, *loitering* e apprendimento automatico).

Maggiori sono le preoccupazioni etico-giuridiche sollevate dalla missione e dallo scenario operativo, più intensa sarà la forma di controllo richiesta affinché possa essere considerata “appropriata al contesto”. Ad esempio, un sistema d'arma progettato per identificare civili che partecipano direttamente alle ostilità in un contesto urbano dovrà essere soggetto ad un controllo più stretto rispetto a quello previsto in relazione ad un sistema che deve intercettare missili nemici in ambiente desertico. Lo stesso dicasi per la presenza o l'assenza di caratteristiche tecniche, come quella di operare in “sciame”, che possano determinare il sorgere di comportamenti “emergenti” e dunque non prevedibili. La presenza di dette caratteristiche, infatti, giustifica un controllo più stringente da parte dell'operatore umano.

## 7. Conclusioni

Una componente ampia e sufficientemente rappresentativa della comunità internazionale – 127 Stati, cui si aggiungono la Santa Sede e la Palestina – si è ormai pronunciata a favore dell'adozione di uno strumento giuridicamente vincolante sui sistemi d'arma autonomi<sup>58</sup>. Si tratta, per intenderci, di una cifra analoga a quella degli Stati parte della Convenzione sulle armi convenzionali (CCW), anche se non vi è coincidenza tra i due insiemi di Stati. A questa evoluzione si affiancano gli altri segnali positivi richiamati sopra: il graduale consolidarsi di un linguaggio comune all'interno del GGE grazie alla circolazione del *rolling text* e, sul piano più politico, l'impegno diretto del Segretario Generale delle Nazioni Unite a promuovere e facilitare la conclusione di un nuovo strumento giuridico sui sistemi d'arma autonomi entro il 2026. Tutto ciò lascia intendere che l'adozione di un trattato in materia, per quanto ancora difficile, non può più essere considerata una chimera.

Tuttavia, permangono ostacoli che non è possibile ignorare. Tra i 54 Stati che non hanno assunto una posizione chiara e, soprattutto, tra i 12 che si sono dichiarati apertamente contrari figurano numerosi Stati “militarmente significativi”, secondo la terminologia del Preambolo della CCW: tra gli altri, Stati Uniti, Russia, Regno Unito, Israele, India e Australia. In tale contesto, non è escluso che l'adozione di un trattato vincolante – se e quando dovesse avvenire – possa replicare la vicenda del Trattato sulla proibizione delle armi nucleari (TPNW). Quest'ultimo, pur avendo ricevuto il sostegno iniziale di una coalizione di 138 Stati, è stato finora ratificato da poco più della metà di questi e da nessuna delle potenze nucleari, limitandone fortemente – fino quasi all'irrilevanza – l'impatto normativo.

Di fronte a questa prospettiva, e tenuto conto della profonda crisi che ormai da anni affligge il multilateralismo, vale la pena esplorare – quanto meno nel breve e medio termine – soluzioni alternative. Sotto questo aspetto, una strategia di *second best* potrebbe consistere nel promuovere l'emersione e il chiarimento delle norme indicate

56. *Rolling text*, cit., III.6.

57. AMOROSO-TAMBURRINI 2021.

58. Maggiori informazioni sono reperibili nel sito di [Automated Decision Research](#).

nelle pagine che precedono favorendo la convergenza progressiva su buone prassi unilaterali, formalizzate nei manuali militari nazionali di DIU. Al momento, l'unico manuale militare nazionale che contiene disposizioni specifiche sui sistemi d'arma autonomi è quello statunitense<sup>59</sup>. Tuttavia, altri Stati stanno attualmente aggiornando i propri manuali di DIU – tra cui l'Italia e il Regno Unito – e tale processo potrebbe offrire un'occasione preziosa per includere disposizioni analoghe, auspicabilmente ispirate all'approccio proposto dal Comitato internazionale della Croce Rossa.

Se sufficientemente omogenee, queste prassi potrebbero contribuire in misura significativa

alla formazione di nuove norme consuetudinarie, colmando parzialmente il vuoto normativo e preparando il terreno per un futuro trattato. La giurisprudenza internazionale ha riconosciuto nei manuali militari una fonte rilevante per l'accertamento degli elementi costitutivi della consuetudine<sup>60</sup>. In questa chiave, essi non devono essere visti solo come strumenti interni di addestramento ma, in attesa di un trattato ampiamente condiviso, anche come veicoli di produzione normativa “dal basso”, capaci di anticipare e orientare sviluppi giuridici futuri.

## Riferimenti bibliografici

- D. AMOROSO (2020), *Autonomous Weapons Systems and International Law. A Study on Human-Machine Interactions in Ethically and Legally Sensitive Domains*, ESI/Nomos, 2020
- D. AMOROSO, G. TAMBURRINI (2021), *Toward a Normative Model of Meaningful Human Control over Weapons Systems*, in “Ethics & International Affairs”, vol. 35, 2021, n. 2
- ARTICLE36 (2013), *Killer Robots: UK Government Policy on Fully Autonomous Weapons*, April 2013
- P. ASARO (2019), *What is an ‘Artificial Intelligence Arms Race’ Anyway?*, in “I/S: A Journal of Law and Policy for the Information Society”, 2019
- I. BODE, H. HUELLS (2022), *Autonomous Weapons Systems and International Norms*, McGill-Queen's University Press, 2022
- V. BOULANIN, M. VERBRUGGEN (2017), *Mapping the Development of Autonomy in Weapon Systems*, SIPRI, 2017
- E. CANNIZZARO (2014), *Proportionality in the Law of Armed Conflict*, in A. Clapham, P. Gaeta (eds.), “The Oxford Handbook of International Law in Armed Conflict”, Oxford University Press, 2014
- C. CARPENTER (2014), *“Lost” Causes: Agenda Vetting in Global Issue Networks and the Shaping of Human Security*, Cornell University Press, 2014
- K. CHOWDHURY (2025), *Adversarial Machine Learning: Attacking and Safeguarding Image Datasets*, in “arXiv”, 31 January 2025
- COMITATO INTERNAZIONALE DELLA CROCE ROSSA (2024), *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*, 2024
- COMITATO INTERNAZIONALE DELLA CROCE ROSSA (2021), *Position on Autonomous Weapon Systems*, 2021
- COMITATO INTERNAZIONALE DELLA CROCE ROSSA (2019), *Autonomy, artificial intelligence and robotics: Technical aspects of human control*, 2019

59. Dipartimento dello Difesa USA, *Law of War Manual*, June 2015 (aggiornato nel luglio 2023), par. 6.5.9.

60. V., ad esempio, Tribunale internazionale per la ex-Jugoslavia, *The Prosecutor v. Dusko Tadić*, IT-94-1-AR72, Appeals Chamber, Decision, 2 ottobre 1995, parr. 92, 99, 108, 118, 131, 580.

- D. COPELAND (2025), *A Functional Approach to the Legal Review of Autonomous Weapon Systems*, Brill, 2025
- B. DE VRIES (2023), *Individual Criminal Responsibility for Autonomous Weapons Systems in International Criminal Law*, Brill, 2023
- M. EKELHOF, G. PERSI PAOLI (2020), *Swarm Robotics. Technical and Operational Overview of the Next Generation of Autonomous Systems*, UNIDIR, 2020
- W.J. FENRICK (1982), *The Rule of Proportionality and Protocol I on Conventional Warfare*, in "Military Law Review", vol. 98, 1982
- J.-M. HENCKAERTS, L. DOSWALD-BECK (2005), *Customary International Humanitarian Law*, Cambridge University Press, 2005
- I. HENDERSON, K. REECE (2018), *Proportionality under International Humanitarian Law: The "Reasonable Military Commander" Standard and Reverberating Effects*, in "Vanderbilt Journal of Transnational Law", vol. 51, 2018, n. 3
- I. HENDERSON, P. KEANE, J. LIDDY (2017), *Remote and Autonomous Warfare Systems: Precautions in Attack and Individual Accountability*, in J.D. Ohlin (ed.), "Research Handbook on Remote Warfare, Edward Elgar", 2017
- J. HOLLAND (2004), *Military Objective and Collateral Damage: Their Relationship and Dynamics*, in "Yearbook of International Humanitarian Law", vol. 7, 2004
- HUMAN RIGHTS WATCH (2025), *A Hazard to Human Rights. Autonomous Weapons Systems and Digital Decision-Making*, 28 April 2025
- HUMAN RIGHTS WATCH (2012), *Losing Humanity. The Case against Killer Robots*, 2012
- N. JEVGLEVSKAJA (2018), *Weapons Review Obligation under Customary International Law*, in "International Law Studies", vol. 94, 2018
- D. KENNEDY (2004), *Dark Side of Virtue: Reassessing International Humanitarianism*, Princeton University Press, 2004
- J. KWIK (2024), *Lawfully Using Autonomous Weapon Technologies*, Springer, 2024
- M. LIU, J. WEI, Y. LIU, J. DAVIS (2025), *Human and AI Perceptual Differences in Image Classification Errors*, in T. Walsh, J. Shah, Z. Kolter (eds.), "Proceedings of the AAAI Conference on Artificial Intelligence", vol. 39, 2025, n. 13
- S. LOSEY (2024), *US Air Force stages dogfights with AI-flown fighter jet*, in "Defense News", 19 April 2024
- M. MANCINI (2006), *Air Operations against the Federal Republic of Yugoslavia (1999)*, in N. Ronzitti, G. Venturini (eds.), "The Law of Air Warfare: Contemporary Issues", Eleven International Publishing, 2006
- D. MAURI (2022), *Autonomous Weapons Systems and the Protection of the Human Person. An International Law Analysis*, Edward Elgar, 2022
- T. MCFARLAND (2020), *Autonomous Weapon Systems and the Law of Armed Conflict. Compatibility with International Humanitarian Law*, Cambridge University Press, 2020
- G. MECACCI, D. AMOROSO, L. CAVALCANTE SIEBERT et al. (eds.) (2024), *Research Handbook on Meaningful Human Control of Artificial Intelligence Systems*, Edward Elgar, 2024
- S. PARKIN (2015), *Killer robots: The soldiers that never sleep*, in "BBC news", 16 July 2015
- D. SAXON (2021), *Fighting Machines: Autonomous Weapons and Human Dignity*, University of Pennsylvania Press, 2021

- P. SCHARRE (2018), *Army of None. Autonomous Weapons and the Future of War*, W.W. Norton & Company, 2018
- P. SCHARRE (2016), *Centaur Warfighting: The False Choice of Humans vs. Automation*, in “Temple International & Comparative Law Journal”, vol. 30, 2016, n. 2
- M.N. SCHMITT, J.S. THURNHER (2013), *“Out of the Loop”: Autonomous Weapon Systems and the Law of Armed Conflict*, in “Harvard National Security Journal”, 2013
- A. SEIXAS-NUNES (2022), *The Legality and Accountability of Autonomous Weapon Systems. A Humanitarian Law Perspective*, Cambridge University Press, 2022
- N.E. SHARKEY (2016), *Staying the Loop: Human Supervisory Control of Weapons*, in N. Bhuta, S. Beck, R. Geiß, et al. (eds.), “Autonomous Weapons Systems: Law, Ethics, Policy”, Cambridge University Press, 2016
- M.R. TADDEO, A. BLANCHARD (2022), *A Comparative Analysis of the Definitions of Autonomous Weapons Systems*, in “Science and Engineering Ethics”, vol. 28, 2022
- J.S. THURNHER (2014), *Examining Autonomous Weapon Systems from a Law of Armed Conflict Perspective*, in H. Nasu, R. McLaughlin (eds.), “New Technologies and the Law of Armed Conflict”, T.M.C. Asser Press, 2014
- L. VALDARES FERNANDES BARBOSA (2025), *Autonomous Weapons Systems and the Responsibility of States. Challenges and Possibilities*, Routledge, 2025
- L. VARELLA (2025), *Editorial: Disarmament is Always the Smartest Choice*, in “CCW Report”, 2025
- M. ZAJAC (2023), *AWS compliance with the ethical principle of proportionality: three possible solutions*, in “Ethics and Information Technology”, vol. 25, 2023, n. 1