

DYNAMES

Governance di Forze e Sistemi

PAPER & DOSSIER

WHITE PAPER



Armi autonome: la corsa regolatoria americana e il vuoto del diritto internazionale

MIRIAM FOTI

2026



DYNAMES

Paper Series
White Paper

White Paper

ARMI AUTONOME: LA CORSA REGOLATORIA AMERICANA E IL VUOTO DEL DIRITTO INTERNAZIONALE

Un'inchiesta giornalistica, un contenzioso federale ancora pendente, una richiesta di bilancio da 54,6 miliardi di dollari ed un negoziato multilaterale che entra nella sua fase conclusiva proprio in queste settimane: il 2026 ridefinisce, in tempo reale, il confine tra chi decide e chi risponde quando la funzione bellica passa, anche solo in parte, ad un sistema algoritmico.

A CURA DI	Dott.ssa Miriam Foti
DATA	16/08/2026
AREA TEMATICA	Diritto e Legittimità
CATEGORIA	White Paper
VERSIONE	1.0

Documento di analisi strategica

White Paper sulle armi autonome, la regolazione statunitense e il vuoto del diritto internazionale.

Summary

Sommario del White Paper

Executive Summary 4

1. Un sistema già operativo 5

2. Il caso Anthropic: il rifiuto di un fornitore e le sue conseguenze 5

3. La produzione normativa statunitense, e il suo contrappeso contabile 6

4. Ginevra: la fase conclusiva di un negoziato irrisolto 7

5. Il vuoto normativo europeo: dove l'AI Act si ferma alla porta delle caserme 8

6. Regolare in casa non equivale a regolare per tutti 8

7. Il precedente kosovaro: quando la legalità non è sufficiente 9

8. Chi risponde quando la macchina sbaglia bersaglio 10

Note 12

Bibliografia essenziale 13

Profilo autrice 14

NOTA EDITORIALE

Il presente documento è un contributo di analisi strategica pubblicato da DYNAMES. Non costituisce pubblicazione scientifica peer-reviewed, salvo diversa indicazione. Le opinioni espresse sono attribuibili all'autrice e non impegnano necessariamente la linea editoriale della rivista o degli enti collegati.

Executive Summary

Sintesi editoriale DYNAMES

Un'inchiesta giornalistica, un contenzioso federale ancora pendente, una richiesta di bilancio da 54,6 miliardi di dollari ed un negoziato multilaterale che entra nella sua fase conclusiva proprio in queste settimane: il 2026 ridefinisce, in tempo reale, il confine tra chi decide e chi risponde quando la funzione bellica passa, anche solo in parte, ad un sistema algoritmico.

Sintesi editoriale DYNAMES. Il documento esamina l'impiego operativo e la disciplina delle armi autonome nel 2026 attraverso il caso statunitense, il contenzioso Anthropic, il negoziato multilaterale di Ginevra, il quadro europeo e il problema dell'attribuzione della responsabilità. Questa sintesi è redazionale e non sostituisce né modifica il testo dell'autrice, riprodotto integralmente nelle sezioni successive.

Il 27 febbraio il Presidente Trump ha ordinato a tutte le agenzie federali statunitensi la cessazione immediata dell'uso della tecnologia sviluppata da Anthropic; poche ore dopo, secondo la ricostruzione del Wall Street Journal, il Comando Centrale statunitense impiegava ancora il modello Claude per attività di intelligence assessment, identificazione dei bersagli e simulazione di scenari operativi nelle prime fasi degli attacchi contro l'Iran¹. Non si tratta di un aneddoto isolato, bensì della dimostrazione, per così dire, empirica di un problema strutturale: la linea di demarcazione tra funzione di supporto decisionale e sistema d'arma è, nella prassi operativa, assai più labile di quanto le categorie giuridiche vigenti riescano a descrivere.

Ed è entro questa zona grigia — non già teorica, ma già misurabile nei fatti — che si colloca, nel 2026, la disciplina internazionale delle armi autonome, o più precisamente la sua perdurante assenza.

1. Un sistema già operativo

Nella dottrina militare statunitense si definisce Autonomous Weapon System (AWS) un sistema che, una volta attivato, è in grado di selezionare ed ingaggiare un bersaglio senza l'ulteriore intervento di un operatore umano in tempo reale, secondo la definizione contenuta nella Direttiva 3000.09 del Dipartimento della Difesa². Nel luglio 2026 il Corpo dei Marines ha selezionato Bullfrog, sistema deputato al rilevamento ed al tracciamento di minacce aeree — segnatamente di natura droniforme — nonché alla puntatura automatica delle armi da fuoco sul bersaglio identificato. Nel medesimo periodo il Pentagono ha esteso l'impiego del Maven Smart System, sviluppato da Palantir, che connette operatori umani, sensori e piattaforme d'arma in tempo reale — sistema nel quale, sino a poco tempo fa, proprio Claude svolgeva una funzione centrale nell'elaborazione del targeting³.

Non si verte, dunque, in un'ipotesi meramente teorica. La distinzione tra configurazioni human-in-the-loop, nelle quali l'operatore umano determina l'ingaggio, e configurazioni human-on-the-loop, nelle quali all'operatore residua la sola facoltà di interrompere un'azione già avviata, non costituisce un dettaglio tecnico: incide direttamente sull'individuazione del soggetto chiamato a rispondere qualora il sistema identifichi erroneamente il bersaglio. Ed è esattamente questo il confine che l'episodio iraniano ha reso evanescente: un modello linguistico impiegato per finalità di assessment, identificazione e simulazione non integra, per definizione, un'arma autonoma — ma la distanza tra l'analisi che suggerisce un bersaglio ed il sistema che lo ingaggia si è ridotta al punto che il diritto internazionale non dispone ancora di strumenti idonei a tracciarne il confine con certezza.

2. Il caso Anthropic: il rifiuto di un fornitore e le sue conseguenze

La vicenda che ha condotto Claude nel teatro operativo iraniano origina settimane prima.

Da mesi il Pentagono chiedeva ad Anthropic — fornitore del modello su reti classificate dal 2024 e tra le società che nel luglio 2025 avevano ottenuto un contratto sino a 200 milioni di dollari per l'adozione di capacità di intelligenza artificiale da parte del Dipartimento⁴ — di rimuovere due restrizioni contrattuali: il divieto di impiego dei propri modelli per finalità di sorveglianza di massa su cittadini statunitensi e per sistemi d'arma pienamente autonomi. Anthropic ha opposto rifiuto: il CEO Dario Amodei ha dichiarato che le armi autonome «potrebbero rivelarsi decisive per la nostra difesa nazionale, ma allo stato attuale i sistemi di intelligenza artificiale non offrono affidabilità sufficiente» per governarle in piena autonomia.

Il 27 febbraio la controversia è divenuta di dominio pubblico. Il Presidente Trump ha disposto la cessazione immediata dell'uso della tecnologia Anthropic da parte di tutte le agenzie federali; il Segretario della Difesa Pete Hegseth ha designato la società quale «supply chain risk» per la sicurezza nazionale, precludendo a fornitori e appaltatori della Difesa qualsiasi rapporto commerciale con essa — prima applicazione di tale designazione, di norma riservata a società collegate a governi esteri ostili, ad un'impresa statunitense.

Poche ore dopo, secondo il Wall Street Journal, il Comando Centrale continuava ad impiegare Claude nella pianificazione degli attacchi contro l'Iran: il sistema, secondo fonti del Pentagono, risultava talmente integrato nell'architettura operativa da non poter essere sostituito nell'arco di poche ore.

Anthropic ha adito le competenti autorità giurisdizionali con due azioni parallele — una corte distrettuale in California, una corte d'appello a Washington — deducendo la violazione del Primo Emendamento e delle

procedure legalmente previste per l'esclusione di un fornitore dalle catene di approvvigionamento federali⁵.

Il 26 marzo 2026 la giudice distrettuale Rita Lin ha concesso un provvedimento cautelare, sospendendo tanto la designazione quanto l'ordine di cessazione dei contratti, e qualificando la misura come «classica ritorsione» nei confronti di una società che aveva pubblicamente criticato l'amministrazione; l'8 aprile la corte d'appello ha tuttavia negato la sospensione della pronuncia, ripristinando temporaneamente l'efficacia del divieto⁶.

Il 30 luglio, in sede di udienza sulle mozioni per giudizio sommario, la medesima giudice Lin ha manifestato l'orientamento a rendere permanente il blocco della designazione, dichiarandosi «turbata» dall'ipotesi che la critica pubblica di un'impresa nei confronti del governo possa legittimarne l'esclusione: «se la preoccupazione concerne l'integrità della catena di comando operativa», aveva scritto nella propria precedente ordinanza, «il Dipartimento della Guerra potrebbe semplicemente cessare di avvalersi di Claude».

Una pronuncia definitiva non è ancora intervenuta; nelle more, secondo quanto riferito in udienza dai legali del Pentagono, la dismissione dei prodotti Anthropic dovrebbe completarsi entro il 3 settembre, salva ulteriore evoluzione del giudizio.

3. La produzione normativa statunitense, e il suo contrappeso contabile

Attorno a questa vicenda si muove un Congresso in fase di intensa attività legislativa.

Il 9 gennaio 2026 il Dipartimento della Difesa — ridenominato "Department of War" quale designazione secondaria ai sensi dell'Executive Order 14347 — ha reso pubblici due memoranda a firma del Segretario Hegseth: il primo stabilisce criteri di valutazione dei modelli di intelligenza artificiale ai fini degli appalti ed impone, entro centottanta giorni, l'inserimento della clausola standard "qualsiasi uso lecito" in ogni contratto avente ad oggetto servizi di intelligenza artificiale, in attuazione della Sezione 1513 del National Defense Authorization Act 2026; il secondo unifica la governance dell'innovazione difensiva sotto un unico Chief Technology Officer⁷, ricollegando la controversia anche al programma di difesa missilistica spaziale "Golden Dome" voluto dal Presidente Trump.

Il 5 giugno un memorandum presidenziale sulla sicurezza nazionale ha incaricato il Pentagono di procedere alla riscrittura della Direttiva 3000.09, testo che dal 2012 disciplina l'autonomia nei sistemi d'arma statunitensi⁸. Sul versante parlamentare, il quadro normativo si è rivelato assai più articolato di quanto risultasse a inizio anno: tra maggio e giugno sono state presentate almeno cinque distinte proposte in materia di intelligenza artificiale militare — quella della senatrice Kirsten Gillibrand, che vieta le armi pienamente autonome salvi i sistemi di difesa locale sottoposti a supervisione; quella del senatore Adam Schiff (l'HALO Act, sulla catena di comando); quella della senatrice Elissa Slotkin (l'AI Guardrails Act); quella del senatore Mark Kelly, presentata congiuntamente al deputato Suhas Subramanyam. È stata tuttavia la proposta dei senatori Chris Coons e Jack Reed, depositata l'8 giugno, ad avere maggiore seguito: il testo — che impone una revisione di livello senior, l'istituzione di un archivio degli incidenti per il tracciamento di malfunzionamenti ed eventi quasi fatali, nonché il divieto assoluto di affidare ad un'intelligenza artificiale le decisioni relative al lancio nucleare — è stato in larga parte recepito dalla maggioranza repubblicana nel testo-base del National Defense Authorization Act per l'anno fiscale 2027, approvato dalla Commissione Forze Armate del Senato il 10 giugno con diciotto voti favorevoli e nove contrari.

Il testo, pur imponendo "livelli appropriati di giudizio umano", conferma nondimeno il sostegno del Pentagono all'impiego di armi letali autonome, richiedendone anzi la "massimizzazione dell'uso"⁹.

Il dato che meglio misura la distanza tra il dibattito regolatorio e l'orientamento effettivo dell'amministrazione è, tuttavia, di natura contabile. La richiesta di bilancio per l'anno fiscale 2027, sottoposta al Congresso il 3 aprile 2026 nell'ambito di uno stanziamento complessivo record da 1.500 miliardi di dollari, alloca 54,6 miliardi di dollari al neocostituito Defense Autonomous Warfare Group (DAWG) — l'organismo che assorbe l'iniziativa Replicator, avviata nel 2023 e formalmente sciolta a fine 2025 per le difficoltà di integrazione industriale incontrate, e che accentra la programmazione dei sistemi autonomi statunitensi.

Si tratta di un incremento di circa il 24.000 per cento rispetto ai 225,9 milioni stanziati per il medesimo organismo nell'esercizio precedente, e di una cifra superiore all'intera richiesta di bilancio del Corpo dei Marines. Mentre il Congresso discute soglie di giudizio umano e archivi degli incidenti, l'esecutivo pianifica,

in termini di risorse effettivamente allocate, quello che l'ex Direttore della CIA David Petraeus ha definito «il più significativo impegno finanziario nella storia della guerra autonoma» — un dato che ridimensiona, quantomeno sul piano della coerenza tra intenzione normativa e prassi di bilancio, la portata delle garanzie procedurali sin qui discusse.

Il percorso legislativo si è, peraltro, successivamente complicato. La Camera ha approvato la propria versione del NDAA 2027 con 216 voti favorevoli contro 212, sostanzialmente lungo linee di partito.

Al Senato, per contro, i Democratici hanno per due volte impedito il passaggio in aula del testo, l'ultima con 50 voti contro 46, insufficienti al raggiungimento della soglia procedurale dei sessanta voti richiesta: non per ragioni attinenti alla disciplina dell'intelligenza artificiale, bensì per l'opposizione al coinvolgimento militare in Iran e per le riserve sulla crescita della spesa complessiva.

4. Ginevra: la fase conclusiva di un negoziato irrisolto

A Ginevra, nel frattempo, il negoziato multilaterale entra nella propria fase più delicata. Il Gruppo di Esperti Governativi (GGE) sulle armi letali autonome, istituito nel 2017 nell'ambito della Convenzione sulle Armi Convenzionali (CCW) del 1980 e presieduto dall'ambasciatore olandese Robert in den Bosch, elabora dal 2021 un "rolling text", ossia un documento recante possibili elementi di un futuro strumento giuridico, costruito attorno al cosiddetto "approccio a due livelli": divieto dei sistemi intrinsecamente incompatibili con il diritto internazionale — in quanto imprevedibili nel funzionamento, non arrestabili una volta attivati, o concepiti per colpire persone come bersaglio primario —, regolamentazione rafforzata di tutti gli altri¹⁰.

Il mandato triennale del Gruppo giunge a scadenza nel 2026, con termine fissato in coincidenza della Settima Conferenza di Revisione della CCW, in programma dal 16 al 20 novembre a Ginevra.

Nel settembre 2025 un terzo dei 128 Stati parte della Convenzione — 42 Paesi, guidati dal Brasile — ha dichiarato il rolling text costituire «una base sufficiente» per l'avvio di negoziati formali su uno strumento vincolante. La sessione del 2-6 marzo 2026 è stata definita dagli osservatori «l'inizio della fine» del mandato del Gruppo, con circa settanta Stati che vi hanno espressamente sostenuto l'apertura di trattative formali.

La seconda sessione del 2026 si svolgerà dal 31 agosto al 4 settembre — ultimo passaggio prima della Conferenza di Revisione di novembre, chiamata a stabilire se conferire al Gruppo un mandato negoziale in senso proprio, se prorogarne il solo mandato di discussione, ovvero se lasciarne decorrere il termine senza esito. La pressione internazionale si è intensificata anche per ragioni contingenti. Nei giorni immediatamente precedenti alla sessione di marzo, la rivelazione del Wall Street Journal sull'impiego di Claude negli attacchi contro l'Iran ha fatto ingresso nella sala negoziale ginevrina: secondo la ricostruzione dell'analista Francesco Vignarca, oltre settanta Paesi — tra cui un gruppo significativo di Stati africani — hanno considerato l'episodio prova che lo sviluppo tecnologico proceda più celermente dei negoziati, sollecitando l'avvio immediato di trattative formali. L'Italia, impegnata nel dossier sin dal 2013 e rappresentata in queste sessioni dall'ambasciatore Leonardo Bencini, ha ribadito la necessità di un quadro normativo «sufficientemente ampio da ricomprendere anche gli sviluppi futuri» di tali tecnologie¹¹.

Permane, tuttavia, il nodo strutturale che ha sinora impedito ogni progresso sostanziale: il GGE delibera per consenso, e a ciascuno Stato è riconosciuto, di fatto, un potere di veto. In assenza di un mutamento sostanziale nel calcolo strategico di Stati Uniti, Russia e Cina — tutti e tre membri permanenti del Consiglio di Sicurezza, tutti e tre impegnati nello sviluppo di sistemi analoghi — un protocollo vincolante entro la scadenza informale del 2026, più volte richiamata dal Segretario Generale delle Nazioni Unite e dalla Presidente del CICR, resta un obiettivo di difficile conseguimento.

5. Il vuoto normativo europeo: dove l'AI Act si ferma alla porta delle caserme

Vi è, in questo quadro, un ulteriore elemento che merita di essere segnalato, e che tocca da vicino l'ordinamento dell'Unione Europea. Il Regolamento (UE) 2024/1689 — il cosiddetto AI Act, pubblicato nella Gazzetta Ufficiale dell'Unione il 12 luglio 2024 ed entrato in vigore il successivo 1° agosto — ha istituito il primo quadro regolatorio organico al mondo in materia di intelligenza artificiale, fondato su un approccio basato sul rischio e su obblighi stringenti per i sistemi ad alto rischio. L'articolo 2, paragrafo 3, esclude tuttavia espressamente dall'ambito di applicazione del Regolamento i sistemi di intelligenza artificiale immessi sul mercato, messi in servizio o utilizzati esclusivamente per finalità militari, di difesa o di sicurezza nazionale — esclusione ulteriormente precisata dal considerando 24, e giustificata dal legislatore europeo mediante il richiamo all'articolo 4, paragrafo 2, del Trattato sull'Unione Europea, che riserva agli Stati membri la competenza esclusiva in materia di sicurezza nazionale¹².

La conseguenza è che l'Unione, pur avendo elaborato la disciplina più ambiziosa al mondo in materia di intelligenza artificiale, non dispone di alcuno strumento vincolante applicabile ai sistemi d'arma autonomi impiegati dalle forze armate dei propri Stati membri. Il vuoto non trova, peraltro, compensazione organica negli strumenti di finanziamento della difesa comune: il Fondo Europeo per la Difesa esclude sì il finanziamento di armi letali autonome prive di un controllo umano significativo sulla selezione e sull'ingaggio del bersaglio, e una clausola analoga era stata inserita nello strumento per gli appalti congiunti EDIRPA, scaduto alla fine del 2025; il regolamento SAFE (Security Action for Europe), che ne ha preso il posto quale principale veicolo di finanziamento per gli approvvigionamenti comuni, non riproduce tuttavia tale clausola di salvaguardia — omissione che il Belgio ha formalmente stigmatizzato in sede di adozione del testo in Consiglio¹³. Restano naturalmente applicabili il diritto internazionale umanitario, le revisioni giuridiche nazionali sui nuovi sistemi d'arma di cui all'articolo 36 del I Protocollo Addizionale del 1977, e i principi dell'Alleanza Atlantica in materia di uso responsabile dell'intelligenza artificiale; si tratta, tuttavia, ancora una volta, di un mosaico di discipline eterogenee e prive di un vincolo sovranazionale comune.

6. Regolare in casa non equivale a regolare per tutti

È a questo punto che la questione cessa di essere meramente tecnica, per assumere un rilievo propriamente giuridico sotto il profilo della legittimità. Una normativa nazionale, per quanto rigorosa, disciplina la condotta di un unico attore all'interno di un sistema che ne annovera una pluralità: le garanzie procedurali che essa introduce — classificazione del rischio, revisione legale, tracciabilità — restano, per definizione, unilaterali, e producono responsabilità di carattere meramente interno, non già un vincolo opponibile a Stati terzi che sviluppino sistemi analoghi in assenza di un obbligo equivalente.

Il diritto internazionale umanitario presuppone, viceversa, un complesso di regole che vincolano la totalità delle parti in conflitto, indipendentemente dal livello di sofisticazione tecnologica di ciascuna: il principio di distinzione impone di colpire esclusivamente obiettivi militari; il principio di proporzionalità vieta l'attacco qualora i danni collaterali prevedibili risultino eccessivi rispetto al vantaggio militare concreto e diretto atteso; il principio di precauzione impone di verificare la natura del bersaglio e di sospendere l'azione al mutare delle circostanze.

Si tratta di valutazioni che la prassi bellica tradizionale affida ad un operatore umano in tempo reale; quando esse vengono trasferite, anche solo parzialmente, ad un sistema automatizzato, il problema non attiene unicamente alla correttezza formale della decisione algoritmica, bensì all'idoneità del sistema a replicare un giudizio che il diritto internazionale ha sempre presupposto essere umano, contestuale e suscettibile di correzione in corso d'azione.

La medesima frattura strutturale si riscontra sul piano geopolitico. La Repubblica Popolare Cinese persegue la dottrina della "guerra intelligentizzata" sin dal 2019, ribadita dal Presidente Xi Jinping in occasione del XX Congresso del Partito Comunista Cinese nel 2022. Nel gennaio 2026 la National University of Defence Technology ha reso noto, tramite l'emittente statale CCTV, un test nel cui ambito un unico operatore ha garantito la supervisione simultanea di oltre duecento droni, dotati — secondo i ricercatori — della capacità di ripianificare autonomamente le rotte anche in presenza di disturbo elettronico¹⁴. Due mesi dopo, il gruppo statale CETC ha presentato un sistema concorrente, denominato "Atlas": un veicolo di lancio in grado di far decollare 48 droni a intervalli di tre secondi, con un unico centro di comando capace di gestirne sino a 96. Gli analisti del China Aerospace Studies Institute mantengono cautela quanto al passaggio dalla dimostrazione all'effettivo impiego operativo, ma la traiettoria dottrinale — dallo sciame dimostrativo

all'integrazione con i caccia J-20 in funzione di "nave madre" — risulta ormai consolidata nella dottrina militare cinese¹⁵.

La Federazione Russa procede su un binario differente: sistemi meno onerosi, meno spettacolari, ma già concretamente impiegati sul campo. Il 16 marzo 2026 un drone "Lancet", prodotto da ZALA Aero del gruppo Kalashnikov, è precipitato in pieno centro di Kyiv, nei pressi del Monumento dell'Indipendenza, nel corso di un attacco su larga scala. L'intelligence militare ucraina (HUR), esaminando i resti, ha rilevato indizi di una significativa modernizzazione: un modulo di guida basato su un processore Nvidia Jetson, concepito per l'identificazione e il tracciamento autonomo del bersaglio senza comunicazione costante con l'operatore, unitamente a marcature circolari analoghe a quelle utilizzate dai droni russi V2U per il volo in formazione.

Su 62 componenti elettronici identificati, la maggior parte risultava di origine occidentale — Stati Uniti, Svizzera, Cina — nonostante il regime sanzionatorio vigente¹⁶. Taluni analisti ucraini hanno posto in dubbio che si trattasse effettivamente di un Lancet, stante la sua limitata autonomia di volo; resta il dato per cui Mosca lavora, dichiaratamente, a rendere i propri sistemi d'attacco sempre meno dipendenti dal pilotaggio in tempo reale.

7. Il precedente kosovaro: quando la legalità non è sufficiente

La distinzione tra legalità e legittimità non costituisce acquisizione recente del diritto internazionale.

Nel 1999, su iniziativa del Primo Ministro svedese Göran Persson, fu istituita la Commissione Indipendente Internazionale sul Kosovo, presieduta dal giudice sudafricano Richard Goldstone — già Procuratore Capo del Tribunale Penale Internazionale per l'ex Jugoslavia — unitamente al politico svedese Carl Tham, con l'incarico di valutare l'intervento della NATO del 1999, condotto senza l'autorizzazione del Consiglio di Sicurezza.

Nel rapporto pubblicato nel 2000, la Commissione è pervenuta ad una conclusione volutamente non lineare: l'intervento era da ritenersi "illegale ma legittimo". Illegale, in quanto privo della base giuridica richiesta dall'articolo 2, paragrafo 4, della Carta delle Nazioni Unite, e non riconducibile né all'eccezione della legittima difesa di cui all'articolo 51, né ad un'autorizzazione del Consiglio di Sicurezza ai sensi del Capitolo VII; legittimo, in quanto rispondente ad un'esigenza umanitaria ampiamente riconosciuta dalla comunità internazionale, in assenza di alternative praticabili.

Il rapporto della Commissione ebbe un'eco duratura: le riflessioni ivi contenute in ordine agli obblighi successivi ad un intervento contribuirono, negli anni seguenti, ad orientare l'elaborazione della dottrina della Responsibility to Protect, successivamente recepita, nel 2005, dal Documento Finale del Vertice Mondiale delle Nazioni Unite. Giova, tuttavia, precisare che il rapporto non fu inteso, né dalla dottrina prevalente né dalla prassi successiva, quale fondativo di un diritto consuetudinario all'intervento umanitario unilaterale: cautela confermata, nel 2010, dal parere consultivo della Corte Internazionale di Giustizia sulla dichiarazione di indipendenza del Kosovo, che si astenne dall'attribuire all'intervento valore di precedente.

La transizione da un atto controverso ad un principio internazionalmente riconosciuto richiese, in altri termini, un processo deliberato di ratifica collettiva — il Vertice Mondiale del 2005 — e non si produsse per mera accumulazione di prassi unilaterale.

Applicata alle armi autonome, la medesima distinzione può essere utilmente ribaltata.

Una normativa pienamente legale — adottata secondo le procedure costituzionali statunitensi, vincolante per il Pentagono, azionabile dinanzi alle competenti autorità giurisdizionali, come dimostra la stessa vicenda Anthropic — non esaurisce, di per sé, la legittimità nel significato proprio del diritto internazionale: quella che deriva dal riconoscimento condiviso, dalla reciprocità degli obblighi, dalla possibilità, per gli altri Stati, di fare affidamento su uno standard comune. Una disposizione pienamente legale nell'ordinamento interno può, sul piano internazionale, risolversi in un atto unilaterale tra i molti: rispettabile, financo apprezzabile, ma non costitutivo di un obbligo opponibile a chi non l'abbia adottata.

Trattasi di un rovesciamento significativo, proprio in quanto controintuitivo. Nel caso del Kosovo, la comunità internazionale si trovò dinanzi ad un'azione priva di base legale, ma ampiamente percepita come legittima, e la relativa eco normativa si consolidò soltanto per il tramite di un atto collettivo di ratifica.

8. Chi risponde quando la macchina sbaglia bersaglio

Il diritto internazionale umanitario dispone già di una dottrina volta ad attribuire la responsabilità nell'ipotesi in cui un subordinato commetta una violazione: la responsabilità di comando, che il I Protocollo Addizionale alle Convenzioni di Ginevra del 1977 costruisce attraverso una combinazione normativa spesso appiattita, nella trattazione divulgativa del tema, sulla sola norma di chiusura. L'articolo 86 pone il fondamento sostanziale della responsabilità, imputando al superiore le violazioni commesse dal subordinato ove il primo sapesse, o disponesse di informazioni tali da consentirgli di concludere, nelle circostanze del momento, che quella violazione si stava per commettere o si stava commettendo, e non abbia adottato tutte le misure fattibili in suo potere per impedirla o reprimerla; l'articolo 87 vi si affianca imponendo ai comandanti un dovere positivo — prevenire, reprimere, denunciare — la cui inosservanza integra, essa stessa, l'omissione rilevante ai sensi dell'articolo 86.

Le due disposizioni, lette congiuntamente e riconosciute quale diritto consuetudinario dalla Norma 153 dello studio sul diritto internazionale umanitario consuetudinario elaborato dal Comitato Internazionale della Croce Rossa, compongono l'impianto della responsabilità di comando: un comandante che impartisce direttive, un subordinato che le esegue e che può, in linea di principio, essere istruito, corretto, sanzionato.

La questione posta dalle armi autonom è se lo standard di conoscenza richiesto dall'articolo 86 risulti predicabile allorché il "subordinato" non sia una persona, bensì un modello che elabora bersagli in base a parametri appresi. Si consideri, a titolo esemplificativo, l'ipotesi di un sistema quale Bullfrog, schierato per l'intercettazione di droni ostili, che identifichi erroneamente quale minaccia un velivolo civile penetrato per errore nello spazio aereo controllato, ed apra il fuoco: risponderebbe, ai sensi dell'articolo 86, l'operatore che ha attivato il sistema, ignaro del difetto e privo di elementi per dubitarne? Ovvero il personale tecnico che, mesi prima e in un contesto del tutto estraneo al teatro operativo, ha definito i criteri di addestramento del sistema medesimo — soggetto che l'articolo 87, calibrato su una catena di comando militare, non è strutturato per raggiungere?

Il quesito non è meramente teorico neppure sul piano della responsabilità penale individuale.

Lo Statuto di Roma della Corte Penale Internazionale, all'articolo 28, distingue infatti due fattispecie con soglie soggettive diverse: la lettera a) attribuisce al comandante militare una responsabilità fondata su uno standard colposo — sapeva, o avrebbe dovuto sapere, in ragione delle circostanze del momento —, sostanzialmente allineato al parametro dell'articolo 86 del I Protocollo Addizionale; la lettera b), riservata ai superiori non militari, richiede invece che questi abbiano consapevolmente ignorato informazioni che indicavano chiaramente la violazione, soglia prossima al dolo eventuale.

Se il personale tecnico che ha addestrato il sistema fosse qualificato come superiore non militare ai sensi della lettera b) — ipotesi tutt'altro che scontata, stante la sua estraneità a qualsivoglia catena di comando in senso proprio —, la sua responsabilità penale individuale richiederebbe la prova di una conoscenza consapevolmente elusa, non della semplice prevedibilità dell'errore. Dunque, la difficoltà di soddisfare un simile standard anche nei confronti di un comandante umano trova puntuale riscontro nella prassi della Corte Penale Internazionale. Nel caso Bemba — l'unico, ad oggi, in cui la CPI abbia applicato l'articolo 28 sino a una condanna — la Camera d'Appello ha ribaltato, nel giugno 2018, la condanna inflitta in primo grado a Jean-Pierre Bemba Gombo quale comandante delle truppe del Mouvement de libération du Congo per crimini commessi nella Repubblica Centrafricana, non già per l'assenza di un controllo effettivo sui subordinati, bensì per l'insufficienza della prova che l'imputato avesse omesso di adottare tutte le misure necessarie e ragionevoli per prevenire o reprimere quei crimini.

Parte della dottrina propende per un'estensione, in via analogica, di logiche prossime alla responsabilità da prodotto. Il filosofo Robert Sparrow, cui si deve la formulazione originaria del concetto di "responsibility gap" applicato ai sistemi d'arma autonomi, ha per primo posto in evidenza come l'assenza di un agente umano individuabile renda strutturalmente problematica l'attribuzione della responsabilità, tanto morale quanto giuridica¹⁷. Sulla medesima linea, il giurista Thompson Chengeta ha proposto modelli di "split-responsibility", nei quali l'onere risulta ripartito tra l'operatore, il comandante ed il soggetto industriale. Taluni autori, infine, richiamano il regime convenzionale sulla responsabilità civile in materia nucleare — segnatamente, la Convenzione di Parigi del 1960 — quale modello praticabile: responsabilità oggettiva del produttore per i difetti di progettazione, temperata da tetti risarcitori e da meccanismi assicurativi obbligatori.

Trattasi, in ciascun caso, di estensioni ben note al diritto interno in altri ambiti, ma che il diritto internazionale umanitario — costruito su soggetti statali e su catene di comando militari — non ha ancora integrato in modo sistematico. Sino a quando tale lacuna non sarà colmata, permane quello che la letteratura specialistica definisce *accountability gap*: il rischio, cioè, che la responsabilità per un errore della specie descritta non ricada, in concreto, su alcun soggetto — né sull'operatore, né sul comandante, né sull'impresa che ha sviluppato il sistema.

La legittimità, nel diritto dei conflitti armati, non si esaurisce nella mera esistenza di una procedura di controllo interno. Essa coincide, piuttosto, con la possibilità di attribuire la responsabilità in modo prevedibile, verificabile ed opponibile, a prescindere dal soggetto che abbia sviluppato od impiegato il sistema — Stato, produttore o impresa tecnologica. Le iniziative legislative statunitensi sin qui esaminate, pur nella loro eterogeneità, condividono un limite comune: esse rispondono all'interrogativo "come regoliamo noi stessi", non già a quello, ben più complesso, "come garantiamo che chiunque impieghi tali sistemi risponda secondo il medesimo standard". Ciò non implica una valutazione critica del valore, in sé considerato, della regolamentazione interna: una direttiva che imponga revisione legale, tracciabilità e supervisione umana rappresenta un progresso reale rispetto all'assenza di regole — a condizione, tuttavia, che non resti smentita, sul piano sostanziale, da scelte di bilancio di segno opposto, come dimostra il caso del DAWG.

Ricondurre, comunque, l'autoregolamentazione nazionale alla legittimità internazionale costituirebbe un errore di impostazione. Il vuoto determinato da un negoziato multilaterale bloccato non si colma sommando discipline nazionali, per quanto numerose od accuratamente redatte — né, tantomeno, esclusioni regolatorie come quella dell'articolo 2, paragrafo 3, dell'AI Act europeo: esso si colma soltanto laddove sussista un vincolo che tutti gli attori rilevanti — inclusi, ormai, i fornitori privati di tecnologia — siano tenuti a rispettare, a prescindere dalla propria collocazione geopolitica o dal proprio modello imprenditoriale.

La Conferenza di Revisione di Ginevra, in novembre, dirà se il 2026 sarà ricordato come l'anno in cui tale vincolo ha iniziato a prendere forma, ovvero come l'ennesima occasione mancata. Nelle more, la questione giuridica di maggior rilievo permane aperta: non già se un sistema abbia rispettato la legge del proprio Stato di appartenenza, bensì se, nel colpire un bersaglio, esso abbia rispettato una legge effettivamente vincolante per la totalità dei consociati.

Note

1. Wall Street Journal, ripreso da CBS News, «Anthropic's Claude AI being used in Iran war by U.S. military, sources say», 2-3 marzo 2026; Bloomberg Opinion, «Iran Strikes: Anthropic Claude AI Helped US Attack. But How Exactly?», 4 marzo 2026; The Times of Israel, 2 marzo 2026.
2. U.S. Department of Defense, Directive 3000.09, «Autonomy in Weapon Systems»; JURIST, «Autonomous Weapons Are Outrunning the Law That Governs Them», luglio 2026 (su Bullfrog).
3. Racism and Technology Center, «Anthropic is powering the horrific strikes on Iran», 9 marzo 2026 (sul ruolo di Claude nel Maven Smart System di Palantir).
4. Congressional Research Service, IN12669, «Pentagon-Anthropic Dispute over Autonomous Weapon Systems: Potential Issues for Congress», aprile 2026 (contratto da 200 milioni di dollari, luglio 2025; dichiarazioni di Dario Amodei).
5. Pearl Cohen, «Anthropic Sues Department of Defense Over Supply Chain Risk Designation», marzo 2026; PBS NewsHour, «Anthropic sues in federal court to reverse Trump administration's 'supply chain risk' designation», 9 marzo 2026.
6. CNBC, «Anthropic loses appeals court bid to temporarily block Pentagon blacklisting», 8 aprile 2026.
7. National Defense Authorization Act per l'anno fiscale 2026, Sez. 1513.
8. National Security Presidential Memorandum, 5 giugno 2026 (riscrittura della Direttiva 3000.09).
9. Congress.gov, S.Rept. 119-127, National Defense Authorization Act for Fiscal Year 2027; CRS IN12704, «FY2027 NDAA: Status of Legislative Activity».
10. Digital Watch Observatory, «GGE on lethal autonomous weapons systems», aggiornato ad agosto 2026; United Nations Office for Disarmament Affairs, calendario delle sessioni GGE 2026.
11. Rappresentanza Permanente d'Italia presso la Conferenza del Disarmo, Ginevra, «Italy's Participation in the Second 2025 Session of the GGE LAWS».
12. Regolamento (UE) 2024/1689, art. 2, par. 3, e considerando 24; Gazzetta Ufficiale dell'Unione Europea, L 2024/1689, 12 luglio 2024, in vigore dal 1° agosto 2024.
13. Geopolitical Monitor, «Europe's Military AI Gap Begins at Fielding», agosto 2026 (Fondo Europeo per la Difesa, EDIRPA, regolamento SAFE, posizione del Belgio in Consiglio).
14. South China Morning Post, «1 soldier, 200 drones: China showcases rapid launch and agility in swarm warfare tactics», 23 gennaio 2026.
15. China Aerospace Studies Institute, dati ripresi da drone-warfare.com, «PLA Drone Swarms: Demonstration vs. Fielding Reality», giugno 2026 (sistema CETC «Atlas»).
16. Ukraine Defense Intelligence (HUR); Kyiv Post, «First AI Lancet Drone Attack Rocks Kyiv?», 16 marzo 2026; The New Voice of Ukraine, «Lancet Drone with Artificial Intelligence — Updates, Western Details, and Risks», 23 marzo 2026.
17. Robert Sparrow, «Killer Robots», Journal of Applied Philosophy, 2007.

Bibliografia essenziale

Riferimenti bibliografici

Schmitt, Michael N. / Thurnher, Jeffrey S., "Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict, in 4 Harvard National Security Journal 231 (2013) — il contributo dottrinale di riferimento sulla tesi opposta a quella "abolizionista": il diritto dei conflitti armati esistente è già in grado di disciplinare le AWS. Utile per dare conto anche della posizione che dissente dal filone Sparrow/Chengeta che hai già citato.

<https://journals.law.harvard.edu/nsj/2013/05/out-of-the-loop-autonomous-weapon-systems-and-the-law-of-armed-conflict>

Stop Killer Robots, Policy Brief: Autonomous Weapons Systems — Key Issues & the Path to a Treaty, 2025 — utile perché riassume in modo aggiornato lo stato del negoziato GGE con apparato di note proprio. https://www.stopkillerrobots.org/wp-content/uploads/2025/05/WEB-Policy-Brief_Autonomous-Weapons-Systems_Key-issues-and-path-to-a-treaty_Stop-Killer-Robots.pdf

Riferimenti sitografici

CBS News — <https://www.cbsnews.com/news/anthropic-claude-ai-iran-war-u-s/>;

Bloomberg Opinion —

<https://www.bloomberg.com/opinion/articles/2026-03-04/iran-strikes-anthropic-claude-ai-helped-us-attack-but-how-exactly>;

The Times of Israel —

<https://www.timesofisrael.com/hours-after-trump-announced-ban-on-claude-ai-us-military-used-it-in-iran-strikes-reports/>;

NBC News (ruolo di Palantir/Maven) —

<https://www.nbcnews.com/tech/tech-news/us-military-using-ai-help-plan-iran-air-attacks-sources-say-lawmakers-rcna262150>;

JURIST — <https://www.jurist.org/commentary/2026/07/autonomous-weapons-are-outrunning-the-law-that-governs-them/>;

CRS, IN12669 (Congress.gov) — <https://www.congress.gov/crs-product/IN12669>

Barchart (ricorso, 9 marzo) — <https://www.barchart.com/story/news/653503/anthropic-sues-pentagon-over-supply-chain-risk-designation-that-blocks-federal-contracts>;

Pearl Cohen (analisi legale) —

<https://www.pearlcohen.com/anthropic-sues-department-of-defense-over-supply-chain-risk-designation/>;

NPR (ricorso) — <https://www.npr.org/2026/03/09/nx-s1-5742548>;

CNN (ordinanza cautelare, 26 marzo) —

<https://www.cnn.com/2026/03/26/business/anthropic-pentagon-injunction-supply-chain-risk>;

The Hill (27 marzo) — <https://thehill.com/policy/technology/5803486-anthropic-lawsuit-pentagon-claude>;

CNBC (diniego sospensione, 8 aprile) —

<https://www.cnbc.com/2026/04/08/anthropic-pentagon-court-ruling-supply-chain-risk.html>;

SecurityWeek (Emil Michael, Golden Dome) — <https://www.securityweek.com/pentagons-chief-tech-officer-says-he-clashed-with-ai-company-anthropic-over-autonomous-warfare/>;

Breaking Defense (Camera, 216-212) —

<https://breakingdefense.com/2026/07/house-passes-2027-ndaa-while-measure-remains-stalled-in-the-senate/>;

Breaking Defense (Senato blocca, 50-46) —

<https://breakingdefense.com/2026/07/senate-democrats-block-ndaa-amid-concerns-on-iran-war-budget-topline/>.

Profilo autrice

Autrice del White Paper

Miriam Foti	Autrice del presente White Paper. Area tematica: Diritto e Legittimità.
--------------------	---

DYNAMES

Difesa • Sicurezza • Resilienza

DYNAMES è una rivista registrata dedicata all'analisi dei temi strategici di difesa, sicurezza, comunicazione, geopolitica, resilienza e interesse nazionale.

Registrata al Tribunale di Busto Arsizio N. 01/2026 del 25 febbraio 2026

Iscrizione ROC n. 44335

Con sede in Via Carlo Noè 45 Gallarate (VA), C.F. 3800470126

Direttore responsabile Fabio Boscacci

Pubblicazione diffusa in formato digitale mediante il sito www.dynames.it e gli altri canali ufficiali della testata.

Tutti i diritti sono riservati. La riproduzione, distribuzione, comunicazione al pubblico, elaborazione o utilizzo, anche parziale, dei contenuti pubblicati è consentita nei limiti previsti dalla legge e previa corretta citazione della fonte, salvo diversa autorizzazione scritta dell'editore o dei titolari dei diritti. Restano consentiti gli utilizzi per finalità di cronaca, studio, ricerca, critica e discussione, nei limiti e secondo le condizioni previste dalla normativa applicabile.

Le immagini, i marchi, i loghi, i segni distintivi, le fonti documentali e gli eventuali materiali di terzi presenti nella pubblicazione appartengono ai rispettivi titolari e sono utilizzati, ove applicabile, per finalità informative, illustrative, scientifiche o di commento. Eventuali richieste di rettifica, segnalazioni relative ai diritti d'autore, comunicazioni legali o richieste di rimozione possono essere inviate a redazione@dynames.it.

Il trattamento dei dati personali eventualmente connesso alle attività editoriali della testata è effettuato nel rispetto della normativa applicabile in materia di protezione dei dati personali. L'informativa privacy aggiornata è disponibile all'indirizzo: redazione@dynames.it.