

L'uso dell'intelligenza artificiale nei conflitti armati contemporanei: appunti essenziali

Marco Pertile

1. Alcune distinzioni fondamentali

La relazione tra intelligenza artificiale (IA) e conflitti armati interessa da tempo il dibattito pubblico. Per ragioni più che comprensibili, l'aspetto che ha maggiormente colpito l'immaginario collettivo riguarda la questione delle armi autonome letali (LAWS) o dei cosiddetti *robot killer*. Si tratta di quei sistemi di arma che, dopo l'attivazione da parte dell'essere umano, sono in grado di selezionare e colpire in modo autonomo gli obiettivi di attacco. È importante chiarire preliminarmente però che nella condotta delle ostilità i sistemi di intelligenza artificiale non si integrano soltanto nelle armi autonome. Recentemente è giunta all'attenzione dei media anche l'esistenza di sistemi di intelligenza artificiale che, nell'ambito dei conflitti armati, si limitano – per così dire – a sostenere e facilitare l'assunzione delle decisioni da parte dei decisori politici e dei comandanti militari. Si tratta di sistemi che facilitano il lavoro degli analisti o lo sostituiscono nel trattare enormi quantità di dati classificando persone e luoghi, determinando la selezione dei bersagli o orientando le decisioni strategiche sull'uso della forza armata. Nella conduzione delle ostilità una distinzione

decisiva intercorre quindi tra armi autonome e sistemi di sostegno decisionale¹. Un'arma autonoma seleziona e ingaggia il bersaglio; un sistema di sostegno decisionale aiuta invece l'essere umano a individuare e a classificare dei bersagli.

Dal punto di vista giuridico, una buona parte dei problemi più seri nelle operazioni di *targeting* contemporanee riguarda i sistemi di sostegno decisionale. Rispetto a essi può accadere infatti che la decisione sia predisposta dal *software* in tempi troppo rapidi e per una troppo ampia quantità di casi impedendo un controllo umano effettivo. A questo proposito, anche nel contesto bellico, è essenziale distinguere tra *big data* e intelligenza artificiale. I *big data* costituiscono la materia prima del processo decisionale: immagini, metadati, geolocalizzazioni, intercettazioni, reti di contatto². L'IA è invece lo strumento analitico che trasforma questa massa di informazioni in classificazioni e raccomandazioni operative. Nei conflitti armati, i problemi giuridici possono derivare da entrambi i piani: dalla scarsa qualità e dal mancato aggiornamento dei dati, ma anche dal modo in cui gli algoritmi li correlano e li traducono in risultati decisionali.

Questa distinzione richiede però un'ulteriore precisazione. Quando l'IA si fonda su tecniche di auto apprendimento (*machine learning*), il problema può riguardare anche l'eventuale opacità del processo decisionale. In questi casi il sistema non si limita a eseguire regole predefinite, ma apprende correlazioni e criteri di classificazione

¹ A. BLANCHARD, L. BRUUN, *Autonomous Weapon Systems and AI-enabled Decision Support Systems in Military Targeting – A Comparison and Recommended Policy Responses*, SIPRI, 2025, https://www.sipri.org/sites/default/files/2025-06/laws_v_0.pdf. Nel linguaggio tecnico si distingue tra armi autonome assistite dall'IA (*AI-enabled AWS*) e sistemi decisionali assistiti dall'IA (*AI-enabled DSS*). Cfr. ICRC and Geneva Academy, *Artificial Intelligence*

and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts, Expert Consultation Report, ICRC and Geneva Academy, Geneva, 2024, 8-9.

² L. DICKINSON, *The Rise of Big Data and the Law of Armed Conflict*, in *Chicago Journal of International Law*, 2025, <https://chicagounbound.uchicago.edu/cjil/vol26/iss1/13>.

a partire dai dati di addestramento. Proprio per questo, il percorso che conduce dai dati al risultato può risultare difficilmente comprensibile per i suoi utilizzatori dando vita al problema della cosiddetta *black box*: l'essere umano che interagisce con l'IA non è in grado di comprenderne pienamente i processi analitici³. Sul piano giuridico, questa opacità nel processo di analisi⁴ è molto rilevante, perché può impedire di verificare la correttezza della classificazione del bersaglio, di prevedere gli effetti del sistema e, più in generale, di esercitare un controllo umano effettivo sulle decisioni operative.

2. Le regole applicabili all'IA nei conflitti armati

Per quanto attiene alle regole giuridiche applicabili all'uso dell'IA nei conflitti armati, è in primo luogo necessario ricordare che l'IA, come fenomeno relativamente recente, non è regolata in modo specifico. La prima cosa da dire con chiarezza però è che l'uso dell'IA nei contesti bellici rientra comunque nell'ambito del diritto dei conflitti armati pur potendo renderne più difficile l'applicazione concreta⁵. In particolare, per quanto riguarda il processo di selezione degli obiettivi e di assunzione delle decisioni di attacco (processo di *targeting*), le parti al conflitto:

- devono distinguere in ogni circostanza tra civili e combattenti, e tra beni civili e obiettivi militari;
- devono rispettare il principio di proporzionalità nell'attacco suscettibile di generare danni incidentali sui civili e/o sui beni civili;
- devono adottare tutte le precauzioni fattibili nell'attacco per minimizzare il danno⁶.

Il diritto internazionale umanitario impone inoltre agli Stati di riesaminare le nuove armi, i nuovi mezzi e i nuovi metodi di guerra prima del loro impiego al fine di valutarne il grado di compatibilità con il quadro giuridico applicabile. Si tratta della cosiddetta "weapons review" che gli Stati sono tenuti a effettuare in base all'articolo 36 del Primo protocollo aggiuntivo alle Convenzioni di Ginevra.

Ciò che cambia con l'IA non è quindi l'esistenza delle regole, ma il modo in cui la conformità a quelle regole viene valutata nella pratica.

Un singolo attacco ad un obiettivo identificato tramite l'IA può essere illecito perché il bersaglio è stato identificato male o perché non sono state adottate precauzioni adeguate per proteggere i civili; ma un sistema può essere problematico anche a monte, se è troppo imprevedibile o troppo difficile da controllare in modo conforme

³ ICRC, *Autonomy, artificial intelligence and robotics: Technical aspects of human control*, August 2019, 13-17, www.icrc.org/sites/default/files/document/file_list/autonomy_artificial_intelligence_and_robotics.pdf; S. SULLIVAN, *Targeting in the Black Box: The Need to Reprioritize AI Explainability, Articles of War*, 28 agosto 2024, <https://lieber.westpoint.edu/targeting-black-box-need-reprioritize-ai-explainability>.

⁴ Per un'analisi delle diverse forme del concetto di opacità (opacità come occultamento intenzionale, opacità come ignoranza tecnologica, opacità come disallineamento cognitivo), vedi M. CARABANTES, *Black-*

box artificial intelligence: an epistemological and critical analysis, in *AI & Society*, 2020, 309-317.

⁵ ICRC, *Submission to the UN Secretary General on the Use of AI in the Military Domain*, RE: ODA/2025-00029/AIMD, 17 aprile 2025, https://www.icrc.org/sites/default/files/2025-04/ICRC_Report_Submission_to_UNSG_on_AI_in_military_domain.pdf.

⁶ Per una succinta descrizione del quadro giuridico, alla luce del diritto pattizio e consuetudinario, vedi ICRC, *Targeting under International Humanitarian Law, How Does Law Protect in War*, <https://casebook.icrc.org/highlight/targeting-under-international-humanitarian-law>.



al diritto internazionale umanitario. Secondo il Comitato internazionale di Croce Rossa⁷, l'integrazione dell'intelligenza artificiale, e in particolare delle tecniche di *machine learning*, nei sistemi d'arma autonomi aggrava i problemi giuridici già posti da questi sistemi. Per usare lecitamente un sistema d'arma autonomo, l'operatore deve poter prevedere con un ragionevole grado di certezza gli effetti dell'arma, così da stabilire se essa possa essere diretta contro un obiettivo militare specifico e se sia possibile limitarne gli effetti, come richiede il diritto internazionale umanitario. Questo diventa problematico per i sistemi che operano in modo opaco, secondo la logica della "scatola nera" descritta in precedenza.

Il dibattito giuridico sull'IA militare, e in particolare sulle armi autonome, non si esaurisce però nell'applicazione del diritto vigente. Un processo guidato dagli Stati Uniti ha favorito l'adozione di una dichiarazione politica sull'uso militare "responsabile" dell'IA, strumento di *soft law* che non comporta l'assunzione di impegni vincolanti⁸. Accanto al problema dell'applicazione del diritto esistente, si è quindi aperto anche un dibattito sull'adozione di nuove regole. Il principale foro negoziale è il Gruppo di esperti governativi costituito a Ginevra nell'ambito della Convenzione quadro sulle armi convenzionali che ha il mandato di formulare, per *consensus*, gli elementi essenziali di un futuro strumento giuridico sulle armi autonome letali. Secondo il

documento della presidenza della prima sessione del 2026 il gruppo ha completato una lettura approfondita di una prima bozza negoziale che affronta quattro nuclei centrali: la definizione e le caratteristiche delle armi autonome letali; il ruolo del controllo e del giudizio umano; la distinzione tra casi di proibizione e casi di regolazione; il rapporto tra il diritto internazionale umanitario vigente e il futuro strumento giuridico⁹. Il Gruppo di esperti non ha raggiunto un accordo definitivo su una definizione di LAWS. È però emerso un ampio consenso sul fatto che il diritto internazionale umanitario continui ad applicarsi pienamente ai sistemi d'arma autonomi. Il negoziato si è inoltre ormai strutturato attorno alla distinzione tra norme proibitive e regolatorie, pur restando aperta la questione di quali sistemi debbano rientrare nell'una o nell'altra categoria.

3. La questione dell'empatia e delle emozioni umane nel processo di selezione degli obiettivi e di attacco

Una questione molto rilevante riguarda la possibilità che il ricorso all'IA e alle armi autonome svincoli le decisioni di *targeting* dalla loro dimensione empatica e dal loro collegamento con le emozioni umane. Secondo una prima tesi, liberare la condotta delle ostilità da emozioni come paura, panico, rabbia, spirito di vendetta o stress potrebbe ridurre gli errori umani e le violazioni¹⁰. In senso opposto, altri ritengono che il problema

⁷ ICRC, *Submission to the UN Secretary General on the Use of AI in the Military Domain*, cit.

⁸ Bureau of Arms Control and Non-Proliferation, *Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy*, 9 novembre 2023, <https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy-2>.

⁹ Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons

Systems Geneva, 2-6 March and 31 August-4 September 2026, Chair's summary – First 2026 session of the GGE on LAWS, CCW/GGE.1/2026/WP.2, 1 aprile 2026.

¹⁰ Lethal autonomous weapons systems, Report of the Secretary-General, UN Doc. A/79/88, 1 luglio 2024, posizione della Federazione russa, 96. Sul punto vedi T. VAN BENTHEM, *Battlefield Uncertainties and the Use of AI*, in *International Law Studies*, 2026, 320-350.

non risieda tanto nella scelta “emozioni sì / emozioni no”, ma nel tipo di giudizio che il diritto richiede e nel rischio che le armi autonome vengano trattate come un fattore deresponsabilizzante¹¹. Sotto questo profilo si sostiene la necessità di un “giudizio umano” perché le armi autonome non sarebbero capaci di replicare un giudizio contestuale, valoriale e anche di senso comune tipico della mente umana¹².

Probabilmente è sensato adottare un approccio intermedio: non è necessario ritenere che l’empatia nelle ostilità sia sempre un bene e la sua assenza sempre un male. Questo non significa però che l’eliminazione delle emozioni equivalga automaticamente a un miglior giudizio. Il rischio opposto è che si perda proprio quella sensibilità al contesto e all’ambiguità che nei casi difficili può fare la differenza tra una decisione giuridicamente difendibile e una decisione disumanizzante.

Dal punto di vista teorico può essere importante, appunto, distinguere tra casi “facili” e i cosiddetti “hard cases”. Nei casi relativamente chiari, una maggiore freddezza cognitiva può apparire vantaggiosa garantendo meno reattività emotiva e meno distorsioni dovute allo stress. Nei casi-limite, invece, cioè quando il diritto richiede valutazioni sofisticate e contestuali, l’assenza di una reale comprensione umana può diventare un problema. Si pensi per esempio alle valutazioni attinenti alla genuinità di una dichiarazione di resa (che nel diritto dei conflitti armati deve essere sempre accettata), oppure alla decisione sul

risparmiare un combattente che non si è formalmente arreso ma è di fatto privo di concreta capacità offensiva. Al di là dell’importanza teorica di queste distinzioni, resta il fatto che, in termini pratici, è difficilissimo prevedere quando in un contesto bellico si presenteranno casi facili o difficili rendendo quindi il ricorso alla scelta autonoma guidata dall’IA particolarmente controverso.

L’empatia non va intesa comunque in senso sentimentalistico, ma come parte di una più ampia sensibilità pratica che consente di percepire l’altro non solo come un fattore di rischio, ma anche come un essere umano vulnerabile. È in questo spazio che il giudizio umano può fare la differenza¹³.

4. La questione delle responsabilità per l’uso dell’IA nei conflitti armati

Sullo sfondo di tutte queste valutazioni si colloca il tema della responsabilità statale e individuale per il ricorso all’IA. Dal punto di vista giuridico, l’intelligenza artificiale non è né può diventare un soggetto responsabile. La responsabilità per le condotte poste in essere nei conflitti armati – incluse le violazioni gravi del diritto internazionale umanitario e gli eventuali crimini di guerra – continua a gravare su soggetti umani e istituzionali: in primo luogo sullo Stato che si serve dell’IA, ma anche, nei casi e alle condizioni previste dal diritto applicabile, sugli individui che hanno progettato, autorizzato, impiegato o ommesso di controllare il sistema¹⁴.

¹¹ JLD WILSON, *AI, war and (in)humanity: the role of human emotions in military decision-making*, in *Humanitarian Law & Policy*, 20 febbraio 2025, <https://blogs.icrc.org/law-and-policy/2025/02/20/ai-war-and-in-humanity-the-role-of-human-emotions-in-military-decision-making>.

¹² *Lethal autonomous weapons systems, Report of the Secretary-General*, cit., posizione dell’Irlanda, 58-59. Cfr. J. DORSEY, *The erosion of human(e) judgment in*

targeting? Quantification logics, AI-enabled decision support systems and proportionality assessments in IHL, *International Review of the Red Cross*, 2025, 1041-1071.

¹³ S. SULLIVAN, *Targeting in the Black Box: The Need to Reprioritize AI Explainability*, cit.

¹⁴ B. BOUTIN, *State responsibility in relation to military applications of artificial intelligence*, in *Leiden Journal of International Law*, 2023, 133-150.



L'uso dell'IA, tuttavia, rende molto più complesso il problema dell'attribuzione della responsabilità¹⁵. Il punto cruciale è che, nei sistemi basati sull'IA, le scelte giuridicamente rilevanti non si concentrano soltanto nel momento terminale della decisione di attacco, ma si distribuiscono lungo l'intero ciclo di vita del sistema¹⁶: nella selezione e nell'aggiornamento dei dati, nella definizione delle categorie operative, nella fissazione delle soglie di rischio, nella progettazione dell'interfaccia, nella validazione del modello e, infine, nella decisione di farvi affidamento in un contesto concreto. I profili di responsabilità nell'uso dell'IA a fini militari devono essere valutati dalla progettazione e acquisizione dei sistemi fino al momento dell'impiego operativo¹⁷. Occorre ricordare che ogni risultato algoritmico è il risultato di una catena di scelte umane e istituzionali pregresse: qualcuno ha selezionato i dati, qualcuno ha deciso quali correlazioni fossero rilevanti, qualcuno ha stabilito le soglie di allarme o di intervento, qualcuno ha ritenuto sufficientemente affidabile il sistema da impiegarlo in un determinato scenario operativo. Sotto questo profilo il ricorso all'IA può costituire un pericoloso strumento di amplificazione dei pregiudizi attraverso l'uso di indicatori *proxy* che possono riprodurre caratterizzazioni stereotipiche degli individui sulla base dell'appartenenza a determinate categorie sociali¹⁸. Un altro rischio è che il risultato dell'analisi condotta tramite IA venga interpretato come una scelta tecnica

insindacabile (il problema dell'IA oracolo). La questione della *black box* e del livello di trasparenza o comprensibilità del processo decisionale guidato dall'IA, inoltre, può costituire un problema giuridico di primaria importanza nella determinazione dell'elemento soggettivo dell'illecito. Il ricorso all'IA, inoltre, può frammentare e distribuire i comportamenti in modo tale da rendere molto difficile l'accertamento anche sotto il profilo probatorio quando è necessario giungere a una determinazione in ordine a chi sapesse cosa, chi avesse il controllo effettivo, chi dovesse verificare l'attendibilità del risultato e chi fosse in posizione di interrompere il processo. È proprio per evitare questa deresponsabilizzazione che il dibattito contemporaneo insiste non solo sulla supervisione umana del singolo attacco, ma anche sulla necessità di mantenere tracciabilità e controllo umano sull'intero processo decisionale¹⁹.

5. Considerazioni conclusive

Il ricorso all'IA nei conflitti armati non ha di per sé effetti necessariamente negativi. In linea di principio, l'IA può aiutare a elaborare meglio i dati, a migliorare la consapevolezza delle situazioni, a rafforzare la verifica dei bersagli. È opportuno però considerare che questi benefici non sono automatici, ma dipendono da come il sistema è progettato e utilizzato. Se l'IA è usata per accelerare e moltiplicare l'adozione delle decisioni, allora il potenziale vantaggio può

¹⁵ ICRC and Geneva Academy, *Expert Consultation Report on AI and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts*, Geneva, ICRC, March 2024.

¹⁶ N. GOUSSAC, V. BOULANIN, *Responsible Procurement of Military Artificial Intelligence*, SIPRI, February 2026, https://www.sipri.org/sites/default/files/2026-02/0226_milai_procurement.pdf.

¹⁷ A. CONN, I. BODE, *Establishing Human Responsibility and Accountability at Early Stages of the Lifecycle for*

AI-based Defence Systems, in *Ethics and Information Technology*, 2025, 51.

¹⁸ L. BRUUN, M. BO, *Bias in Military Artificial Intelligence and Compliance with International Humanitarian Law*, SIPRI, August 2025, https://www.sipri.org/sites/default/files/2025-08/0825_ai_military_bias.pdf.

¹⁹ ICRC, *Artificial intelligence and machine learning in armed conflict: A human-centred approach*, *International Review of the Red Cross*, 2020, 463-479.

Editorial

trasformarsi in un fattore di rischio. Pur non essendo ancora regolato da norme specifiche, il ricorso all'IA non deroga alle norme applicabili del diritto dei conflitti armati. L'applicazione del quadro giuridico esistente all'IA presenta però profili di particolare complessità. Oltre a indurre processi di deliberazione umana più rapidi e deboli, l'IA rischia di rendere poco intelleggibili i meccanismi di attribuzione della responsabilità. Appare difficile isolare l'apporto umano, talvolta remoto e indiretto, in processi decisionali meno comprensibili e trasparenti. La sfida, da questo punto di vista, è garantire che l'integrazione dell'intelligenza artificiale nei meccanismi decisionali preservi un controllo umano determinante nelle scelte belliche.

