

Call for papers: "Intelligenza
Artificiale: prospettive bioetiche,
bio giuridiche e sociali"

L'IA nell'azione bellica: problemi e
sfide per il controllo umano della
guerra

*AI in warfare: problems and
challenges for human control of
war*

GUGLIELMO TAMBURRINI
guglielmo.tamburrini@unina.it

AFFILIAZIONE
Università di Napoli Federico II

SOMMARIO

Conservare il controllo umano sull'azione bellica è una questione di grande rilevanza etica e giuridica. Non ricadono sulle macchine le responsabilità per lo svolgimento di un'azione bellica e per violazioni del Diritto internazionale umanitario. La vigilanza dei controllori umani può inoltre contribuire a evitare errori nei quali la macchina potrebbe incorrere operando su campi di battaglia dinamici e poco prevedibili. Questa esigenza normativa si scontra con i rapidi sviluppi tecnologici nel settore delle armi autonome e con le condizioni materiali di impiego dei sistemi dell'IA per il supporto alle decisioni belliche. Lo scarto crescente tra richiesta normativa e realtà dello sviluppo tecnologico mette impietosamente in luce la lentezza delle attuali iniziative diplomatiche, intergovernative e della società civile sulla regolamentazione dell'IA in ambito bellico, insieme all'urgenza di affrontare con maggiore efficacia le sfide etiche e giuridiche poste dalle applicazioni militari dell'IA.

PAROLE CHIAVE

Armi autonome

Sistemi di supporto alle decisioni belliche

Teoria della guerra giusta

Diritto internazionale umanitario

Intelligenza artificiale.

ABSTRACT

Preserving human control on warfare is a major ethically and legally motivated demand. Machines cannot be held responsible for the outcomes of warfare actions or for violations of moral and legal norms that combatants are required to respect. Human control may additionally help one to avoid errors made by machines in the fog of war. However, these normative needs increasingly jar with both developments of AI technologies enabling autonomous weapons and circumstances of use for AI decision-support systems in warfare scenarios. These tensions shed light on the inadequacy of current actions on the regulation of AI systems in warfare, and urgently call for more effective societal, diplomatic and intergovernmental initiatives.

KEYWORDS

Autonomous weapons systems

Decision-support systems in warfare

Just war theory

International humanitarian law

Artificial intelligence.

DOI: 10.53267/20240102



1. IA, ARMI AUTONOME E SUPPORTO ALLE DECISIONI IN AMBITO BELLICO

Le tecnologie dell'Intelligenza Artificiale (IA) si stanno ampiamente diffondendo nel settore militare. In questo ambito, le applicazioni dell'IA sono utilizzate anche per lo svolgimento di funzioni specifiche delle forze armate in ambito bellico. Vi sono sistemi dell'IA che forniscono un supporto alle decisioni necessarie per pianificare e svolgere le azioni belliche. E i cosiddetti sistemi di arma *autonomi* (che indicheremo qui più brevemente come 'armi autonome') sono in grado di individuare un obiettivo e di attaccarlo senza che sia necessario l'intervento di un operatore dopo la loro attivazione. Lo sviluppo delle armi autonome più avanzate si basa su tecnologie dell'IA per la percezione artificiale, la pianificazione e l'esecuzione delle azioni¹.

Lo sviluppo delle armi autonome si trova all'origine di un problema nuovo nella storia dell'interazione tra esseri umani e macchine in ambito bellico. La diffusione delle armi autonome potrebbe rendere il controllo umano sull'azione bellica tecnicamente superfluo su una scala difficilmente concepibile in precedenza, perché nessun operatore umano è inserito nel processo di elaborazione e decisione che conduce dalla selezione di un possibile obiettivo, alla pianificazione dell'attacco contro di esso e infine alla messa in atto dell'attacco. A quali condizioni sia possibile mantenere nella mano dell'uomo il controllo sulle armi autonome e perché dovrebbe essere mantenuto è una questione ampiamente dibattuta in ambito militare, accademico, diplomatico e intergovernativo. Ma il problema del controllo umano sull'IA impiegata a scopi bellici è balzato più recentemente alla ribalta in un altro contesto di interazione umano-macchina, che non riguarda l'autonomia operativa nell'azione bellica, bensì il supporto che l'IA offre alla scelta delle azioni da compiere sul campo di battaglia. Il problema è emerso in tutta evidenza da un'inchiesta giornalistica condotta dalle riviste "972" e "Local Call" a proposito dei sistemi informatici Habsora e Lavender, che sono stati sviluppati utilizzando anche tecniche di IA. Questi sistemi sono stati utilizzati nel 2023 dalle Forze armate israeliane (IDF) per la pianificazione di bombardamenti aerei da mettere in atto nella Striscia di Gaza².

Le liste di obiettivi da bombardare compilate da questi sistemi devo-

no essere vagliate da un operatore umano, che ha il compito di accettare, modificare oppure rifiutare ciascuna delle proposte in lista. La configurazione dell'interazione umano-macchina prevede costitutivamente la subordinazione decisionale della macchina all'essere umano. Al contrario di quanto accade nel caso delle armi autonome, l'operatore umano svolge una funzione di controllo essenziale, e più specificamente di filtro sugli obiettivi selezionati dalla macchina. E tuttavia, come vedremo, le reali condizioni al contorno per lo svolgimento di una tale funzione di filtraggio non bastano a garantire un controllo umano efficace.

Sotto l'ipotesi che alle testimonianze riportate dall'inchiesta giornalistica corrispondano elementi di verità, il caso di Habsora e Lavender illustra drammaticamente come il problema del controllo umano sulle decisioni e sulle azioni messe in atto da macchine 'intelligenti' dell'IA in ambito bellico abbia una portata più ampia di quanto sia stato comunemente assunto o rilevato nel dibattito etico e giuridico. Infatti, l'attenzione alle questioni normative si è finora concentrata sull'autonomia operativa dei sistemi d'arma che possono selezionare e attaccare un obiettivo senza che sia necessario un intervento umano.

Oltre a determinare un'estensione del problema del controllo umano sull'azione bellica, gli sviluppi tecnologici recenti dell'IA contribuiscono ad aggravarne ulteriormente la portata anche all'interno del suo perimetro originario, e cioè in riferimento alle armi autonome. La recente sperimentazione di un caccia da combattimento manovrato in volo da un sistema dell'IA ci pone di fronte alla concreta realizzazione di un velivolo autonomo da combattimento capace di ingaggiare un confronto ravvicinato con altri velivoli da caccia, senza richiedere in nessuna fase del confronto l'intervento da parte di un pilota o di un qualsiasi altro operatore umano³.

Conservare il controllo umano sull'azione bellica ha grande rilevanza etica e giuridica. Non possono ricadere sulle macchine coinvolte le responsabilità per lo svolgimento di un'azione bellica e per eventuali violazioni delle norme morali e giuridiche che i combattenti devono rispettare.⁴ Gli attuali sistemi dell'IA, per quanto intelligenti possano essere ritenuti, non sono agenti morali e non possono dare conto delle azioni nelle quali sono causalmente coinvolti. La vigi-

lanza dei controllori umani può inoltre contribuire a evitare errori che la macchina potrebbe commettere negli ambienti dinamici e poco prevedibili che spesso caratterizzano i teatri bellici. Bisogna però constatare che questa esigenza normativa non mantiene il passo con i rapidi sviluppi tecnologici nel settore delle armi autonome e con le condizioni materiali di impiego dei sistemi dell'IA per il supporto alle decisioni belliche. Lo scarto crescente tra esigenza normativa e realtà tecnologica mette impietosamente in luce la lentezza e l'inadeguatezza delle attuali iniziative diplomatiche e intergovernative sulla regolamentazione dell'IA in ambito bellico, ponendo con urgenza il problema di mettere rapidamente in campo iniziative più tempestive ed efficaci.

Ma procediamo con ordine, a partire da un esame del problema del controllo umano nel caso di sistemi per il supporto all'azione bellica abilitati dalle tecnologie dell'IA.

2. IA E SELEZIONE DI OBIETTIVI MILITARI

Secondo l'inchiesta giornalistica condotta dalle riviste "+972" e "Local Call", l'IDF avrebbe impiegato due sistemi di supporto alle decisioni basati sull'IA – noti con i nomi di Habsora e Lavender – per generare liste di obiettivi potenziali da bombardare nella striscia di Gaza. Questi sistemi sarebbero stati utilizzati sistematicamente nei primi mesi dopo il massacro di civili innocenti perpetrato il 7 ottobre 2023 in territorio israeliano da Hamas e il concomitante sequestro di ostaggi. Poiché si tratta di sistemi di supporto alle decisioni, le scelte operate da tali sistemi sono solo condizionalmente accettate ed eseguite. Spetta infatti a operatori militari competenti il compito di passare in rassegna le proposte avanzate dalla macchina per poi accettarle, rivederle o respingerle.

Si possono addurre rilevanti motivazioni di carattere etico e giuridico a giustificazione di una siffatta subordinazione della macchina al giudizio di un operatore umano. La teoria della guerra giusta ammette che in determinate situazioni il ricorso alle armi sia moralmente giustificato. Nondimeno, essa prescrive che tutte le parti coinvolte in un conflitto bellico debbano ottemperare a determinati vincoli morali nella conduzione delle operazioni belliche. In particolare, bisogna rispettare l'immunità dei non combattenti e bisogna astenersi dall'infliggere danni sproporzionati

in relazione agli obiettivi militari da conseguire⁵. Questi elementi della teoria della guerra giusta sono incorporati nel principio di distinzione e nel principio di proporzionalità del Diritto Internazionale Umanitario (DIU)⁶. Il giudizio degli operatori deve servire a controllare, tra l'altro, che l'attacco agli obiettivi designati dalla macchina sia compatibile con questi e altri principi del DIU.

Per l'inchiesta di "+972" e "Local Call", il controllo effettuato dagli operatori umani non sarebbe stato all'altezza di questo gravoso compito. Essi avrebbero invece esercitato un controllo affrettato e in definitiva del tutto nominale, svolto in condizioni tali da rendere la loro attività assimilabile a quella di semplici passacarte, che avallano meccanicamente le indicazioni fornite dalla macchina. Fonti anonime dell'IDF menzionate nell'inchiesta giornalistica hanno affermato che l'introduzione dei sistemi di supporto alla scelta degli obiettivi da bombardare ha comportato un chiaro incremento del numero di possibili obiettivi per unità di tempo: «da 50 obiettivi all'anno» elaborati manualmente in situazioni precedenti si è arrivati «fino a 100 obiettivi al giorno». Questo incremento avrebbe innescato una forte pressione psicologica da parte dei superiori militari, con la richiesta rivolta ai controllori di mantenere il passo con la macchina: «Prepariamo gli obiettivi automaticamente e lavoriamo in base a una lista di controllo (checklist) ... È proprio come in una fabbrica. Operiamo con rapidità e non c'è tempo per un approfondimento sull'obiettivo. L'idea è che siamo giudicati in base al numero di obiettivi che riusciamo a generare». Emerge qui chiaramente la pressione psicologica percepita dagli operatori umani di allineare il ritmo delle loro valutazioni all'incremento di 'produttività' impressa dalla macchina e la denuncia della conseguente difficoltà di dare un giudizio ponderato sulle azioni belliche suggerite dal sistema di supporto alle decisioni.

È possibile che la ricerca di un rapido e decisivo vantaggio militare sul campo sia stata la fonte principale della pressione psicologica esperita dalle fonti interne all'IDF citate nell'inchiesta giornalistica. Gli operatori umani sarebbero diventati il principale collo di bottiglia per il conseguimento degli obiettivi militari desiderati. Il tempo richiesto da un controllo scrupoloso sulle liste di obiettivi generati dalla macchina avrebbe indotto una coda sempre più lunga di obiettivi in attesa di convalida. Una lezione fondamentale emerge da questo resoconto,

L'IA nell'azione
bellica

Call for papers:
"Intelligenza
Artificiale:
prospettive
bioetiche,
biogiuridiche e
sociali"

Volume 9 ■ 2024

theFuture
ofScience
andEthics

25

anche indipendentemente da quale sia realmente la motivazione militare alla base della pressione psicologica esperita dagli operatori e da quali siano le eventuali negligenze dei superiori militari relative al rispetto del DIU. La lezione riguarda il peso da dare al fattore tempo nella progettazione del controllo umano sui sistemi di supporto alle decisioni per l'azione bellica. Una configurazione delle interazioni umano-macchina che preveda la presenza di uno o più operatori come filtro non basta a garantire che questi ultimi riescano a valutare con cura le proposte avanzate dalla macchina. Finestre temporali troppo ristrette per l'esercizio del controllo umano sono una fonte significativa di perturbazioni del processo decisionale affidato agli operatori.

Daniel Kahneman ha esposto in forma semplificata e comunicativamente efficace questo tipo di problema, facendo riferimento a due sistemi cognitivi distinti che generalmente sono all'opera nei processi decisionali⁷. Il Sistema 1 consiste di processi decisionali euristici: veloci, per lo più automatici e connotati emotivamente. Il Sistema 2 consiste di un insieme di processi decisionali più lenti e meno prontamente attivabili, riflessivi e analitici. I due sistemi operano in maniera concorrente, ma non sempre cooperativa. Le opzioni più prontamente indicate dal Sistema 1 vengono spesso accettate per *default*, soprattutto quando bisogna decidere sul da farsi avendo poco tempo a disposizione. Le opzioni di attacco selezionate da Habsora o Lavender sono prontamente disponibili, mentre l'esplorazione di altre opzioni può rivelarsi molto più dispendiosa in termini di tempo e di altre risorse cognitive.

Un procedimento mentale euristico – che induce a preferire opzioni prontamente disponibili ignorando alternative mentalmente più faticose da esplorare, è stato designato da Kahneman e Amos Tversky con l'acronimo WYSIATI (What-You-See-Is-All-There-Is). Questa euristica riflette, per Kahneman, una «grande asimmetria tra i modi in cui la nostra mente tratta le informazioni immediatamente disponibili e quelle che non lo sono». Le persone tendono a focalizzare su ciò che si trova innanzi ai loro occhi, trascurando informazioni rilevanti che potrebbero essere rintracciate o recuperate dalla memoria. WYSIATI accelera il processo decisionale ma porta con sé un pregiudizio (*bias*) che può indurre in errore: «Le informazioni che non sono recuperate (nemmeno incon-

sciamente) dalla memoria potrebbero anche non esistere»⁸. E così la pressione psicologica a prendere decisioni a ritmo serrato può indurre gli operatori ad accettare acriticamente i suggerimenti resi prontamente disponibili dalla macchina.

È stato stimato un margine di errore del 10% per le liste di obiettivi generate dai sistemi di supporto alle decisioni utilizzate dall'IDF. Le implicazioni etiche e giuridiche di una finestra temporale insufficiente ad esprimere un giudizio ponderato risultano ancora più gravi alla luce di questa stima di errore. Un giudizio ponderato su liste di obiettivi gravate da una notevole percentuale di errore è cruciale per risparmiare la vita di persone fuori combattimento, di civili innocenti, del personale sanitario e di altre persone protette dal principio di distinzione codificato nel DIU. Sui comandanti militari ricade la responsabilità di fare il possibile per garantire che ci siano le condizioni per arrivare a un tale giudizio ponderato. E nello scenario descritto tali condizioni evidentemente non sussistono.

Insieme alla disponibilità di tempo sufficiente, altre condizioni devono essere rispettate per esprimere un giudizio umano ponderato sui suggerimenti forniti da una macchina. Evidentemente è necessario che gli operatori abbiano competenze sufficienti e la capacità di resistere al cosiddetto *bias* da automazione («Chi sono io, con le mie limitate capacità cognitive, per contraddire le scelte operate da una macchina che è stata addestrata con quantità gigantesche di dati, dei quali potrò padroneggiare ed elaborare solo una minima frazione nel corso della mia vita?»)⁹. Gli operatori dovranno inoltre convivere con l'opacità dei processi di elaborazione degli attuali sistemi dell'IA; con una mancanza di trasparenza che ostacola l'interpretazione e la spiegazione delle ragioni che si trovano alla base delle decisioni proposte dalla macchina.

In definitiva, i sistemi di supporto alle decisioni belliche basati sull'IA sono evidentemente dei sistemi socio-tecnici. Bisogna affrontare e coordinare tra loro formidabili sfide tecniche e scientifiche, sociali, organizzative e psicologiche per mitigarne le limitazioni ed evitarne usi eticamente e giuridicamente ingiustificabili. Nelle narrazioni più comuni sull'innovazione tecnologica – in ambito bellico, ma non solo – lo sguardo viene troppo spesso distolto da questi problemi e dalle loro implicazioni per la società.

3. CONTROLLO DELLE ARMI AUTONOME

Nel caso di un'arma autonoma, come è stato sottolineato già all'inizio di questo articolo, non si richiede che le decisioni della macchina siano vagliate da un operatore umano prima di essere messe in pratica. Si pone dunque il problema se sia ancora possibile garantire un controllo umano efficace in queste circostanze e, in caso contrario, se l'uso di un'arma autonoma non soggetta a un controllo umano in corso d'opera sia compatibile con i vincoli etici e giuridici che sovrintendono alla conduzione delle attività belliche.

Una risposta influente a questo problema è stata fornita dal Comitato Internazionale della Croce Rossa (CICR). Nel 2021, il CICR ha elaborato uno schema generale per arrivare a un trattato internazionale vincolante che introduca forti restrizioni sullo sviluppo e l'uso delle armi autonome¹⁰. Lo schema, sostenuto da varie argomentazioni etiche e giuridiche, si basa su due richieste di divieto e su una richiesta di regolamentazione. La proposta del CICR si inserisce in una famiglia più ampia di proposte di regolamentazione delle armi autonome, che sono dette 'a due livelli' perché comprendono sia divieti sia restrizioni. Le proposte di regolamentazione a due livelli riscuotono attualmente un consenso molto ampio nel dibattito internazionale sulle armi autonome. Un consenso molto più ampio rispetto a proposte più restrittive che prevedono la messa al bando di qualsiasi tipo di arma autonoma. Le richieste di divieto avanzate dal CICR riguardano le armi autonome che siano progettate per attaccare direttamente le persone oppure che abbiano effetti imprevedibili rispetto al DIU. Le armi autonome che non sono vietate su queste basi dovranno essere regolamentate nel loro impiego in base a una serie di vincoli che articolano in varie forme la richiesta di una supervisione umana efficace. Ecco, in sintesi, le tipologie di vincoli individuate dal CICR:

1. Bisogna proibire le armi autonome progettate o utilizzate in maniera tale che i loro effetti non possono essere previsti, compresi e spiegati allo scopo di prevenire effetti indiscriminati che possono scaturire dal loro uso e che risultano essere incompatibili con il principio di distinzione del DIU.
2. Bisogna proibire le armi autonome che sono progettate o utilizzate per attaccare direttamente le persone.

Il motivo principale di questa proibizione è la salvaguardia della dignità delle potenziali vittime di un'azione bellica. Queste ultime non dovrebbero mai essere soggette a decisioni da parte di una macchina che riguardano la loro incolumità fisica.

3. Bisogna regolamentare mediante opportuni vincoli la progettazione e l'utilizzazione delle armi autonome che non ricadono nei casi 1-2. In particolare, bisogna limitare a oggetti di natura esclusivamente militare il tipo di obiettivi da attaccare; introdurre limitazioni sulla durata, sull'area geografica e sulla portata dell'azione dell'arma autonoma, al fine di rendere possibile la supervisione umana su ogni specifico attacco; introdurre vincoli sull'interazione umano-macchina, per garantire che vi sia una supervisione umana efficace in corso d'opera, che comprenda la possibilità di intervenire tempestivamente e disattivare l'arma autonoma¹¹.

Delle considerazioni etiche e giuridiche a sostegno del punto 1 si è già detto in riferimento alla teoria della guerra giusta e al DIU. Anche il punto 2 si basa sul rispetto del DIU e fa inoltre appello al rispetto della dignità umana¹². A questo proposito, il filosofo Peter Asaro ha sostenuto che dal rispetto della dignità umana discende il diritto di ogni essere umano a non essere privato arbitrariamente della vita. E affinché una decisione di vita o di morte soddisfi questo requisito, ha affermato Asaro, essa deve essere presa da un altro essere umano, da un *qualcuno* che possa provare empatia e avere compassione per chi è oggetto di una decisione di vita o di morte. Un'arma autonoma è invece un *qualcosa*, una macchina che non ha la capacità di apprezzare il valore della vita umana e di valutare adeguatamente il significato della sua perdita. Per questo motivo, affidare a una macchina la decisione di togliere la vita a un essere umano, conclude Asaro, costituisce una violazione della dignità umana¹³.

Il punto 3 si basa sull'ipotesi che armi autonome sufficientemente prevedibili e comprensibili nei loro effetti esistano o possano essere progettate, in modo tale da consentire un'adeguata forma di controllo umano sul loro impiego, da indirizzarne l'uso in conformità al DIU e da permettere l'attribuzione chiara di responsabilità a persone inserite nella catena di comando e controllo, qualora si verificano violazioni di norme morali o

L'IA nell'azione
bellica

Call for papers:
"Intelligenza
Artificiale:
prospettive
bioetiche,
biogiuridiche e
sociali"

giuridiche sulla condotta delle azioni belliche.

Un sistema d'arma autonomo che sembra conformarsi al punto 3 è il sistema mobile antimissile *Iron Dome*, utilizzato in Israele per monitorare e neutralizzare con il lancio di missili intercettori i razzi e altri proiettili balistici diretti verso il territorio israeliano.¹⁴ Gli operatori posizionano sul terreno *Iron Dome* e circoscrivono l'area che esso deve monitorare e proteggere. Dopo aver compiuto queste operazioni preliminari, *Iron Dome* viene abilitato a rispondere in piena autonomia, lasciando però agli operatori la facoltà di disabilitare il sistema in corso d'opera. Dotato di analoghe capacità di intercettazione, anche il sistema NBS (*Nächstbereichschutzsystem*) *MANTIS* è utilizzato dalle forze armate tedesche per proteggere i soldati e le installazioni militari da proiettili balistici in arrivo.¹⁵

Gli attuali sviluppi tecnologici di armi autonome (e di loro precursori) non sembrano accordarsi sempre con il punto 3 e con le altre richieste di proibizione totale avanzate dal CICR. Come è stato già ricordato, nel 2023 l'aviazione militare statunitense ha condotto con successo delle prove sperimentali su un aereo da caccia la cui navigazione aerea è interamente controllata da un sistema dell'IA. Il caccia è stato testato in uno scenario di confronto aereo a distanza ravvicinata (*dogfight*) con un altro caccia, eseguendo manovre di attacco o di evasione. I piloti che erano presenti nella cabina di pilotaggio avevano facoltà di escludere il sistema di IA e di assumere essi stessi il controllo della navigazione aerea. Ma non c'è stato bisogno di fare ciò. Dai test condotti risulta perciò evidente la possibilità tecnologica di trasformare questo aereo in un'arma autonoma, dotandolo delle capacità di selezionare e attaccare un obiettivo, in aggiunta alla capacità acquisita di navigazione autonoma e di *dogfighting* avvalorata dalle prove sperimentali.

Dalla prospettiva assunta nel punto 1 più sopra, è opportuno chiedersi se il comportamento di un caccia di questo genere sia sufficientemente prevedibile, allo scopo di prevenire effetti indiscriminati proibiti dal DIU. Non è ovvio che questa condizione sia invariabilmente soddisfatta in confronti aerei ravvicinati che potrebbero anche coinvolgere un numero elevato di velivoli in interazione cooperativa o competitiva tra loro. Inoltre, i vincoli enunciati al punto 3 prevedono che vi sia la garanzia

di un intervento umano tempestivo, per disattivare l'arma autonoma o correggerne l'azione. A causa della latenza prolungata di segnali di controllo impartiti a grande distanza dal velivolo – come accade nel caso di droni che sono pilotati da una stazione a terra situata a migliaia di chilometri di distanza – questa richiesta impone che i segnali di controllo per correggere l'azione del caccia autonomo siano impartiti da una postazione collocata a distanza più ravvicinata. Nelle missioni che comportano il sorvolo di territorio nemico o di zone militarmente contestate, questa possibilità di controllo sembra restringersi a postazioni collocate su altri velivoli sufficientemente vicini, con tutte le limitazioni del caso, che comprendono sia la perturbazione delle comunicazioni nello spettro elettromagnetico da parte del nemico sia la neutralizzazione delle postazioni ravvicinate di controllo. Come verrà interpretato il principio di necessità militare incorporato nel DIU se il controllo ravvicinato verrà meno? Si rinuncerà a concludere la missione oppure si lascerà piena autonomia all'aereo da caccia senza pilota a bordo? È difficile conciliare la seconda opzione con la proposta di regolamentazione a due livelli avanzata dal CICR.

Anche senza elaborare ulteriormente sui problemi tecnologici collegati all'esercizio del controllo umano su un aereo da caccia autonomo abilitato a confronti ravvicinati, la sperimentazione di un tale velivolo entra in tensione con il rispetto di vincoli eticamente e giuridicamente motivati che sono contenuti nella piattaforma 'a due livelli' per la regolamentazione delle armi autonome proposta dal CICR. La possibilità tecnologica di un *dogfight* 'autonomo' non era stata ancora avvalorata sperimentalmente nel 2021, all'epoca in cui fu resa pubblica la proposta del CICR. Ma tale proposta è stata reiterata pressoché *verbatim* anche nel 2024 nel contributo del CICR al Rapporto del Segretario Generale dell'ONU sulle cosiddette 'armi autonome letali'.¹⁶ La persistente assenza di una regolamentazione internazionale non può che aggravare il problema posto da uno sviluppo incalzante di applicazioni militari dell'IA che erodono il controllo degli esseri umani sulle attività belliche, con le conseguenti minacce per il rispetto del DIU e per la protezione della pace. Lo scarto tra esigenza normativa e realtà tecnologica in evoluzione mette impietosamente in luce i ritardi delle attuali iniziative della società civile, degli organismi intergovernativi e della

diplomazia sulla regolamentazione dell'IA in ambito bellico, la persistente assenza di concreti esiti normativi, insieme alla necessità, sempre più pressante, di fronteggiare i rischi crescenti derivanti da una perdita di controllo umano sull'azione bellica.

NOTE

1. Per una analisi critica recente del dibattito sulle armi autonome in lingua italiana e per un approfondimento sui principi etici e giuridici sottiacenti, si veda F. Farruggia (a cura di), *Dai droni alle armi autonome*, Roma: FrancoAngeli, 2023. Il volume *open access* è liberamente scaricabile all'indirizzo <https://www.francoangeli.it/Libro/Dai-droni-alle-armi-autonome?Id=28524>. Si vedano inoltre i capp. 4 e 5 di G. Tamburrini, *Etica delle macchine. Dilemmi morali per robotica e intelligenza artificiale*, Roma: Carocci, 2020, e, in inglese, D. Amoroso, *Autonomous weapons systems and international law. A study on human-machine interactions in ethically and legally sensitive domains*, Napoli / Baden-Baden: Edizioni scientifiche italiane / Nomos, 2020, nonché i capp. 9 e 13 di G. Mecacci et al. (a cura di) *Research handbook on meaningful human control of artificial intelligence systems*, Cheltenham, UK: Edward Elgar Publishing, 2024.

2. Si vedano gli articoli "A mass assassination factory: Inside Israel's calculated bombing of Gaza" (accessibile all'indirizzo <https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/>) e "Lavender: The AI machine directing Israel's bombing spree in Gaza" (accessibile all'indirizzo <https://www.972mag.com/lavender-ai-israeli-army-gaza/>) I contenuti di questi articoli sono stati ripresi da autorevoli testate giornalistiche. Si veda per esempio l'articolo apparso su The Guardian il 1° dicembre 2023 (<https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets>). Per un'analisi condotta dalla prospettiva del Diritto Internazionale Umanitario (DIU) delle procedure documentate in questi articoli giornalistici, si vedano D. Amoroso, "Sistemi di supporto alle decisioni basati sull'IA e crimini di guerra: alcune riflessioni alla luce di una recente inchiesta giornalistica", *Diritti Umani e Diritto Internazionale* 2 (2024): 347-368 e D. Mauri, "Numeri, persone, umanità: sistemi di supporto alle de-

cisioni umane in campo militare da parte dell'IDF e diritto internazionale umanitario", *Diritti Umani e Diritto Internazionale* 2, (2024): 329-346.

3. I test sperimentali sono stati condotti nella seconda metà del 2023 e i risultati conseguiti sono stati resi pubblici nell'aprile del 2024. Si veda <https://www.defensenews.com/air/2024/04/19/us-air-force-stages-dogfights-with-ai-flown-fighter-jet/>

4. Daniele Amoroso e Guglielmo Tamburrini, "Toward a Normative Model of Meaningful Human Control over Weapons Systems", *Ethics and International Affairs* 35 (2021): 245-272.

5. Michael Walzer, *Guerre giuste e ingiuste. Un discorso morale con esemplificazioni storiche*, Napoli: Liguori Editore, 1990, 175-216.

6. Entrambi i principi sono codificati nei protocolli aggiuntivi alle Convenzioni di Ginevra del 1949, consultabili in rete all'indirizzo https://www.icrc.org/eng/assets/files/other/icrc_002_0321.pdf

7. Daniel Kahneman, *Pensieri lenti e veloci*, Milano: Mondadori, 2021.

8. Kahneman, *Pensieri lenti e veloci*, 96.

9. Per approfondimenti, si vedano A. Coco, "Exploring the Impact of Automation Bias and Complacency on Individual Criminal Responsibility for War Crimes", *Journal of International Criminal Justice* 21 (2023): 1077-1096; M. L. Cummings, "Revisiting human-systems engineering principles for embedded AI applications", *Frontiers in Neuroergonomics* 4 (2023): 1-6 (doi: 10.3389/fnrgo.2023.1102165); M. L. Cummings, "Rethinking the maturity of artificial intelligence in safety-critical settings", *Artificial Intelligence Magazine* 42 (2021): 6-15.

10. CICR (Comitato Internazionale della Croce Rossa), *ICRC Position on Autonomous Weapons Systems*, Geneva: ICRC, May 12, 2021 <https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems>.

11. CICR, *ICRC Position on Autonomous Weapons Systems*, 2.

12. Daniele Amoroso, Frank Sauer, Noel Sharkey, Lucy Suchman, Guglielmo Tamburrini, *Autonomy in Weapon Systems. The Military Application of Artificial Intelligence as a Lit-*

L'IA nell'azione
bellica

Call for papers:
"Intelligenza
Artificiale:
prospettive
bioetiche,
biogiuridiche e
sociali"

Volume 9 ■ 2024

theFuture
ofScience
andEthics

29

L'IA nell'azione
bellica

Call for papers:
"Intelligenza
Artificiale:
prospettive
bioetiche,
biogiuridiche e
sociali"

mus Test for Germany's New Foreign and Security Policy, Berlino: Heinrich Böll Foundation, 2018.

13. P. Asaro, "On banning autonomous weapon systems: human rights, automation, and the de-humanization of lethal decision-making", *International Review of the Red Cross*, 94 (2012): 687-709.

14. <https://www.army-technology.com/projects/irondomeairdefence-mi/>. Si veda Amoroso e Tamburrini, cit., per un'analisi della compatibilità con il DIU della forma di autonomia "difensiva" esercitata da tali sistemi sotto la supervisione umana.

15. <https://www.army-technology.com/projects/mantis/>

16. UN Secretary General, Lethal Autonomous Weapons Systems, Report to the UN General Assembly, 79th Session, 1 July 2024. <https://digital-library.un.org/record/4059475?v=pdf>