CENTRO ALTI STUDI
PER LA DIFESA

ISTITUTO DI RICERCA E
ANALISI DELLA DIFESA

Istituto Superiore di Stato Maggiore Interforze

25° Corso – 1ª Sezione – 2° Gruppo di Lavoro

"Applicazione di Artificial Intelligence per fini militari: individuazione dei criteri relativi al passaggio dall'approccio Human in the Loop allo Human on the Loop e definizione delle conseguenti implicazioni sul ciclo di definizione e approvazione delle ROE, con considerazioni sull'adeguamento del quadro normativo in caso di incidenti/eventi avversi."

(AS-SME-01)

ISTITUTO DI RICERCA E ANALISI DELLA DIFESA

L'Istituto di Ricerca e Analisi della Difesa (di seguito IRAD), per le esigenze del Ministero della Difesa, è responsabile di svolgere e coordinare attività di ricerca, alta formazione e analisi a carattere strategico sui fenomeni di natura politica, economica, sociale, culturale, militare e sull'effetto dell'introduzione di nuove tecnologie che determinano apprezzabili cambiamenti dello scenario di difesa e sicurezza, contribuendo allo sviluppo della cultura e della conoscenza a favore della collettività e dell'interesse nazionale.

L'IRAD, su indicazioni del Ministro della difesa, svolge attività di ricerca in accordo con la disciplina di Valutazione della Qualità della Ricerca e sulla base della Programma nazionale per la ricerca, sviluppandone le tematiche in coordinamento con la Direzione di Alta Formazione e Ricerca del CASD.

L'Istituto provvede all'attivazione e al supporto di dottorati di ricerca e contribuisce alle attività di Alta Formazione del CASD nelle materie d'interesse relative alle aree: Sviluppo Organizzativo; Strategia globale e sicurezza/Scienze Strategiche; Innovazione, dimensione digitale, tecnologie e cyber security; Giuridica.

L'Istituto opera in coordinamento con altri organismi della Difesa e in consorzio con Università, imprese e industria del settore difesa e sicurezza; inoltre, agisce in sinergia con le realtà pubbliche e private, in Italia e all'estero, che operano nel campo della ricerca scientifica, dell'analisi e dello studio.

L'Istituto, avvalendosi del supporto consultivo del Comitato scientifico, è responsabile della programmazione, consulenza e supervisione scientifica delle attività accademiche, di ricerca e pubblicistiche.

L'IRAD si avvale altresì per le attività d'istituto di personale qualificato "ricercatore della Difesa, oltre a ricercatori a contratto e assistenti di ricerca, dottorandi e ricercatori post-dottorato.

L'IRAD, situato presso Palazzo Salviati a Roma, è posto alle dipendenze del Presidente del CASD ed è retto da un Ufficiale Generale di Brigata o grado equivalente che svolge il ruolo di Direttore.

Il Ministro della Difesa, sentiti il Capo di Stato Maggiore dalla Difesa, d'intesa con il Segretario Generale della Difesa/Direttore Nazionale degli Armamenti, per gli argomenti di rispettivo interesse, emana le direttive in merito alle attività di ricerca strategica, stabilendo le linee guida per l'attività di analisi e di collaborazione con le istituzioni omologhe e definendo i temi di studio da assegnare all'IRAD.

I ricercatori sono lasciati liberi di esprimere il proprio pensiero sugli argomenti trattati: il contenuto degli studi pubblicati riflette quindi esclusivamente il pensiero dei singoli autori e non quello del Ministero della Difesa né delle eventuali Istituzioni militari e/o civili alle quali i Ricercatori stessi appartengono.

CENTRO ALTI STUDI
PER LA DIFESA

ISTITUTO DI RICERCA E
ANALISI DELLA DIFESA

Istituto Superiore di Stato Maggiore Interforze

25° Corso – 1ª Sezione – 1° Gruppo di Lavoro

"Applicazione di Artificial Intelligence per fini militari: individuazione dei criteri relativi al passaggio dall'approccio Human in the Loop allo Human on the Loop e definizione delle conseguenti implicazioni sul ciclo di definizione e approvazione delle ROE, con considerazioni sull'adeguamento del quadro normativo in caso di incidenti/eventi avversi."

(AS-SME-01)

"Applicazione di Artificial Intelligence per fini militari: individuazione dei criteri relativi al passaggio dall'approccio Human in the Loop allo Human on the Loop e definizione delle conseguenti implicazioni sul ciclo di definizione e approvazione delle ROE, con considerazioni sull'adeguamento del quadro normativo in caso di incidenti/eventi avversi."

NOTA DI SALVAGUARDIA

Quanto contenuto in questo volume riflette esclusivamente il pensiero dell'autore, e non quello del Ministero della Difesa né delle eventuali Istituzioni militari e/o civili alle quali l'autore stesso appartiene.

NOTE

Le analisi sono sviluppate utilizzando informazioni disponibili su fonti aperte.

Questo volume è stato curato dall'**Ufficio Studi, Analisi e Innovazione dell'IRAD**.

Direttore
Col. c. (li) s. SM Gualtiero Iacono

Capo dell'Ufficio Studi, Analisi e Innovazione
Col. AArnn Pil. Loris Tabacchi

Progetto grafico
1° Mar. Massimo Lanfranco – C° 2ª cl. Gianluca Bisanti – Serg. Manuel Santaniello

Revisione e coordinamento
C.V. Massimo GARDINI – S.Ten. Elena Picchi – Funz. Amm. Aurora Buttinelli – Ass. Amm. Anna Rita Marra

Autore
ISSMI – 25º Corso 1ª Sezione 2º Gruppo di Lavoro

Stampato dalla Tipografia del Centro Alti Studi per la Difesa

Istituto di Ricerca e Analisi della Difesa
Ufficio Studi, Analisi e Innovazione
Palazzo Salviati
Piazza della Rovere, 83 - 00165 – Roma
tel. 06 4691 3205
e-mail: irad.usai.capo@casd.difesa.it

chiusa a luglio 2023

ISBN 979-12-5515-057-2

CENTRO ALTI STUDI PER LA DIFESA

ISTITUTO SUPERIORE DI STATO MAGGIORE INTERFORZE

25° CORSO SUPERIORE DI STATO MAGGIORE INTERFORZE

1ª Sezione - 2° GdL

Applicazione di Artificial Intelligence per fini militari: individuazione dei criteri relativi al passaggio dall'approccio Human in the Loop allo Human on the Loop e definizione delle conseguenti implicazioni sul ciclo di definizione e approvazione delle ROE, con considerazioni sull'adeguamento del quadro normativo in caso di incidenti/eventi avversi.

Anno Accademico 2022 - 2023

COMPOSIZIONE DEL GRUPPO DI LAVORO

Dott.	CECCHINI	Andrea	*Presidente*
Ten. Col (CC)	COMPARATO	Marco	*Segretario*
Ten. Col. (AM)	D'AMBROSIO	Pasquale	
Magg. (EI)	AMERICO	Michele	
Magg. (EI)	ARDIZZONE	Francesco	
Magg. (AM)	BALDI	Paolo	
Magg. (EI)	CASTELLANI	Lorenzo	
Magg. (AM)	D'IPPOLITO	Marcelo	
Magg. (EI)	GUERRIERA	Marco	
C.C. (MM)	POLITO	Raffaele	
Magg. (EI)	SCIÒ	Daniela	
C.C. (MM)	VILLANI	Roberto	
Magg.	ABDULLAHI	Bashir	*(Somalia)*
C.C.	LINARES	Victor	*(Perù)*

INDICE

ABSTRACT

Negli ultimi due decenni l'ambito dell'intelligenza artificiale (IA) ha registrato un importante sviluppo su scala globale, caratterizzato da un ritmo di crescita esponenziale. I sistemi che incorporano ed utilizzano attivamente tecnologie "intelligenti" hanno toccato molti aspetti della vita quotidiana dei cittadini, con particolare riferimento ai Paesi più sviluppati; non deve pertanto sorprendere come l'intelligenza artificiale offra grandi aspettative e potenzialità anche in ambito militare. Allo stato attuale, un numero crescente di armi autonome letali operano sia con compiti difensivi che offensivi, fornendo una superiorità strategica e tattico-operativa derivante dalla contrazione dei tempi di reazione da una parte e da una espansione della capacità d'azione dall'altra. Lo sviluppo dell'applicazione dell'IA in campo militare ha condotto ad una ridefinizione del rapporto *human-machine teaming*, concretizzatasi nel passaggio da un approccio *Human in the loop* (HIL) ad uno di tipo *Human on the loop* (HOL). Ciò impone necessariamente di dover esaminare l'adeguatezza delle Regole d'ingaggio (ROE) conseguente alla predetta transizione nonché di analizzare la necessità di adattamento dei riferimenti normativi in considerazione del principio di *accountability*. A tal proposito, anche in virtù dei dibattiti scaturiti nell'ambito della comunità internazionale, è stato introdotto il principio, non ancora codificato, né da tutti condiviso, del "*Controllo umano significativo*" (CUS). L'IA, applicata ai sistemi autonomi in ambito militare, mira a rivoluzionare le tradizionali logiche del *warfare,* ponendo, al contempo, considerevoli rischi derivanti da una possibile proliferazione indiscriminata e dalla mancanza di una regolamentazione condivisa da tutti.

INTRODUZIONE

L'intelligenza artificiale, inquadrata nell'ambito delle *emerging and disruptive technologies*, è divenuta ormai un fattore strategico decisivo in grado di rivoluzionare e ridefinire *tout court* - in una variegata pluralità di ambiti - il modo con cui l'uomo interagisce con sistemi informatici che, di fatto, simulano i processi di intelligenza umana. L'impiego di queste piattaforme, altamente impattanti e performanti, ha conosciuto un largo sviluppo soprattutto in ambito Difesa, dove, sempre più frequentemente, emerge la necessità di reagire rapidamente alle possibili minacce provenienti da entità ostili o asimmetriche in potenziali teatri operativi, all'interno dei quali appare complesso, o proibitivo, proiettare la presenza fisica. Appare evidente, infatti, che l'ampio ricorso a questa tipologia di tecnologie implica l'opportunità per gli operatori di poter intervenire con una minore esposizione al rischio, un elemento questo dirimente quando le criticità operative, connesse alla missione, non consentono di avere né una *situational awareness* chiara sull'ambiente operativo né tantomeno una conseguente *situational understanding* complessiva ed esaustiva dell'intero dominio d'azione. È auspicabile, inoltre, che l'applicazione dell'intelligenza artificiale per i sistemi di Comando e Controllo (C2), utilizzati per supportare i Comandanti nella direzione e nelle differenti fasi di monitoraggio delle operazioni multi-dominio, possa ulteriormente contribuire a velocizzare i cicli di raccolta e analisi del flusso delle informazioni indispensabili per l'attuazione e la sistematizzazione del processo decisionale in situazioni caotiche, caratterizzate da un elevato grado di imprevedibilità (secondo il processo denominato OODA *loop*)[1]. La possibilità di usufruire di dati costantemente aggiornati, classificati e adeguatamente memorizzati si traduce in un evidente nonché cruciale vantaggio strategico per i decisori, i quali potranno sfruttare questo inestimabile supporto informativo nelle varie fasi operative applicando appieno i concetti di *information superiority* e *information dominance*. In un contesto come quello militare, caratterizzato dall'esigenza di rapidità e dalla lenta ma inesorabile proliferazione dei sistemi autonomi, l'intelligenza artificiale costituisce uno strumento tecnologico utile ad orientare e a tracciare le nuove architetture securitarie dei singoli Stati. L'IA, dunque, come già evidenziato da numerosi studiosi ed

[1] L'OODA loop (*Observe, Orient, Decide, Act*), meglio noto come ciclo di Boyd, costituisce un modello di pensiero, applicabile nell'ambito di un processo decisionale, attraverso il quale il decisore può valutare le situazioni in modo razionale anche in situazioni profondamente caotiche. I primi due elementi costitutivi (*Observ, Orient*) fanno riferimento «alla fase di rilevamento e identificazione degli obiettivi, mentre gli altri due (Decide, Act) intervengono al momento dell'ingaggio e della eliminazione degli obiettivi nemici». Fondamentalmente, l'obiettivo del ciclo OODA è quello di anticipare le mosse dell'avversario, o meglio, anticipare e completare il proprio OODA loop eliminando gli obiettivi del rivale prima che questi possa completare il proprio. https://www.agendadigitale.eu/sicurezza/lintelligenza-artificiale-applicata-al-settore-militare-vantaggi-rischi-e-tutele-necessarie/; cfr. Stato maggiore della difesa, L'impatto delle emerging and disruptive tecnologies, edizione 2022, pp. 16-17.

esperti del settore, rappresenta un fattore cruciale in grado di avere un impatto significativo su tutti i domini operativi, poiché consente di identificare obiettivi potenzialmente pericolosi sulla base di schemi di classificazione gerarchizzata per priorità di minaccia. Tuttavia, è proprio attorno al possibile dispiegamento degli apparati autonomi, con particolare riferimento ai sistemi d'arma autonoma (*Autonomous Weapons System* - AWS), che sono emerse le maggiori problematiche giuridiche intrinsecamente connesse alla loro conformità legale con i principi universali di proporzionalità, distinzione e necessità ascrivibili al Diritto Internazionale Umanitario (DIU).

Nel presente elaborato verranno illustrate, nelle loro differenti declinazioni, le varie modalità d'impiego dell'intelligenza artificiale nell'ambito militare, con particolare riferimento alle diverse forme di sistemi autonomi attualmente esistenti e capaci di esprimere, anche in base alle specificità tecnologiche e componentistiche della piattaforma utilizzata, un differente grado di ibridazione *Human-Machine Teaming*. Saranno, inoltre, analizzati i vantaggi strategici, ma anche i potenziali limiti derivanti dal passaggio da un sistema semi-autonomo *Human In the Loop* (HIL), a un sistema autonomo supervisionato *Human on the Loop* (HOL) e come questa transizione tecnologica comporti inevitabilmente la necessità di adeguare e integrare il ciclo di definizione e approvazione delle *Rules of Engagements* (ROE) con l'obiettivo di circoscrivere i parametri, le modalità di dispiegamento e i margini di intervento dei sistemi autonomi IA durante la conduzione delle ostilità all'interno di un determinato scenario operativo. Le ROE, come vedremo, possono contribuire a disciplinare la tipologia e il grado di interazione tra il singolo individuo e i sistemi IA, stabilendo al contempo procedure *standard* attraverso le quali operatori o comandanti possono monitorare tali piattaforme durante le differenti fasi di dispiegamento. È innegabile, poi, che l'impiego di queste tecnologie sollevi rilevanti dilemmi, oltre che etico-morali, di natura legale. In un siffatto contesto, infatti, emerge chiaramente la necessità di stabilire i criteri per l'attribuzione della responsabilità nel caso di incidenti o eventi avversi connaturati all'impiego di sistemi autonomi IA. Saranno analizzati, a tal proposito, i differenti livelli di *accountability* derivanti dall'impiego di tali piattaforme tecnologiche: individuale, *command*, aziendale e statale.

CAPITOLO I: INTELLIGENZA ARTIFICIALE IN AMBITO MILITARE

Grazie ai recenti sviluppi tecnologici, l'Intelligenza Artificiale rappresenta oggi un'applicazione concreta di sistemi in grado di svolgere ragionamenti logici in maniera similare all'architettura neurale umana. Utilizzata ormai in svariati campi, ne ha già in parte modificato tempistiche e abitudini. Le applicazioni di IA coinvolgono anche l'ambito militare con l'intento di migliorare le capacità belliche, grazie all'impiego di sistemi e tecnologie avanzate che sfruttano proprio l'intelligenza artificiale. L'utilizzo di tali sistemi però, oltre a rivoluzionare le modalità di conduzione della guerra moderna in scenari sempre più complessi, solleva anche numerose preoccupazioni in materia di sicurezza e responsabilità.

1. Lo sviluppo dell'intelligenza artificiale e l'interesse in ambito militare

Sviluppatasi nel contesto di quella che è considerata la 4^ rivoluzione industriale, l'intelligenza artificiale è un ambito di ricerca interdisciplinare che si occupa dello sviluppo di sistemi informatici che, attraverso l'utilizzo di algoritmi complessi e modelli matematici, riescono a prendere decisioni e risolvere problemi in maniera similare all'intelligenza umana. La capacità di analisi di grandi quantità di dati è stata possibile grazie all'aumento esponenziale della velocità di elaborazione e dell'ampliamento del volume di archiviazione dei flussi informativi.

Il Parlamento Europeo definisce l'IA come "l'abilità di una macchina di mostrare capacità umane quali il ragionamento, l'apprendimento, la pianificazione e la creatività". L'IA, secondo la stessa istituzione, "permette ai sistemi di capire il proprio ambiente, mettersi in relazione con quello che percepisce e risolvere problemi, e agire verso un obiettivo specifico" ed inoltre "sono capaci di adattare il proprio comportamento analizzando gli effetti delle azioni precedenti, lavorando in autonomia"[2].

Oggetto di studio per esperti e accademici, l'IA è oggi diventata realtà con importanti implicazioni in grado di stravolgere in breve tempo la vita quotidiana e con innovazioni capaci di alterare profondamente il modo di pensare e condurre la guerra. L'IA è già ampiamente utilizzata in diversi settori, quali l'automazione industriale, la medicina, il commercio, i trasporti, l'intrattenimento e, non da ultimo, il settore militare, offrendo molteplici opportunità volte a ottimizzare i processi e a migliorare la società in cui viviamo. D'altra parte, come tutte le tecnologie dirompenti, l'IA può anche presentare una serie di rischi qualora essa non venga compresa, regolamentata o utilizzata in modo inappropriato,

[2] https://www.europarl.europa.eu.

divenendo persino pericolosa in taluni ambiti come quello militare, ove un errore può comportare la perdita di vite umane con ripercussioni difficilmente prevedibili.

Oltre a trasformare il nostro modo di vivere, l'IA viene percepita dalle principali potenze militari come uno strumento attraverso il quale rivoluzionare le logiche del *warfare* per ottenere un vantaggio strategico significativo sugli avversari. Le applicazioni militari dell'IA, in ambito di pianificazione logistica e gestione amministrativa, sono entrate nell'uso quotidiano da almeno due decenni, mentre è ancora in fase di sviluppo l'utilizzo nel settore delle attività cinetiche. Il potenziamento della robotica, infatti, sta indirizzando anche il settore delle armi verso un futuro tecnologico in cui la componente umana presto potrebbe diventare l'anello debole della catena decisionale.

Una prima classificazione dell'IA, spesso utilizzata per fare un grande distinguo, è quella tra *Artificial Narrow Intelligence* (ANI) e *Artificial General Intelligence* (AGI). Le applicazioni attuali sono tutti esempi di ANI, dove le macchine svolgono un compito specifico per uno scopo stabilito. I sistemi di IA operano utilizzando modelli computazionali che consentono loro di apprendere dai dati e di risolvere problemi complessi in modo più efficiente rispetto ai *software* convenzionali, grazie all'individuazione di relazioni statistiche all'interno di enormi *set* di dati. In quest'ottica, sono comunemente impiegati nelle funzioni di automatizzazione di compiti, elaborazione di *set* di dati complessi, previsione comportamentali, individuazione di anomalie, etichettatura dei dati e correzione degli errori.

L'AGI è, invece, la capacità ipotetica di un sistema informatico di eseguire una pluralità di funzioni cognitive e di rispondere a una vasta gamma di *input*, comprendendo e risolvendo problemi che tradizionalmente richiederebbero l'impegno dell'uomo.

A ciò deve aggiungersi una terza tipologia di IA, ancora in fase di sviluppo, definita come "*Artificial Superior Intelligence*" (ASI). Si prevede, infatti, che questa particolare forma di intelligenza sarà in grado di superare gli esseri umani in quasi tutte le aree, ma soprattutto nella creatività scientifica, nella logica e nelle abilità sociali.

Il funzionamento dell'IA, con le sue capacità generalizzanti e predittive, si basa essenzialmente su algoritmi fondati su di un meccanismo di *input-output*, in cui l'*input* è rappresentato dai dati con i quali vengono "nutriti" gli algoritmi. Risulta pertanto di fondamentale importanza la "qualità" dell'*input* inserito, anche in relazione alla capacità di apprendimento dei sistemi di IA (*machine learning*). Per *machine learning* si intende la capacità delle macchine di apprendere senza essere state esplicitamente e preventivamente programmate. Si tratta di programmi già utilizzati in diversi campi dell'informatica, ma che presentano ancora numerose incognite, limiti e rischi, come ad esempio la distinzione tra apprendimento *online* e *offline* e la possibilità di un apprendimento

“selettivo”, nodo fondamentale per comprendere ed analizzare il futuro di questo genere di tecnologie in ambito militare. Occorre poi tener presente che l’IA non opera in una condizione di completo isolamento, ma funziona come una sorta di “architrave” all’interno di una struttura più ampia, caratterizzata da connessioni reciproche concepite appositamente per consentire al sistema di raggiungere il suo obiettivo.

Per quanto riguarda i sistemi di IA, da un punto di vista prettamente tecnico, è possibile distinguere tre differenti livelli di autonomia: automatico, autonomizzato e autonomo. Con il primo si intendono quei sistemi che si limitano soltanto a eseguire comandi e operazioni predefinite tramite un *input* o un sensore (ne sono un esempio bracci robotici presenti nell’industria da almeno due decenni). Risulta invece più complessa la distinzione tra sistema autonomo e autonomizzato: le macchine capaci di far fronte a eventuali variabili ed esercitare azioni proprie possono essere descritte sia come automatizzate sia come autonome. I sistemi autonomizzati si limiterebbero ad eseguire in maniera logica funzioni già prestabilite, mentre i sistemi autonomi, che rappresentano una sorta di evoluzione dei primi, sono macchine in grado di svolgere un compito senza l’*input* umano. Tali assetti si basano su sistemi informatici che impiegano l’IA per interpretare informazioni ricevute da sensori che attivano dei sistemi attuatori come motori, sistemi d’allarme o armi. Tale attivazione risponde a una variazione di una situazione definita come “normale”, quale, ad esempio, la comparsa di un individuo o di un mezzo nel campo visivo o di intervento di un sistema d’arma. Secondo un approccio “funzionale”, basato sui compiti e sulle decisioni eseguite dalla macchina, riferirsi ad un sistema o arma autonomo/a in termini generici risulta inesatto in quanto è proprio la natura dei singoli *task* che esso svolge ciò che realmente conta e non il suo livello di autonomia. Quest’ultima distinzione deve essere analizzata in relazione ai tipi di attività che vengono eseguite a livello di sottosistemi e funzioni della macchina/sistema. Infatti, alcuni compiti nei sistemi d’arma possono essere autonomi senza presentare elevati rischi di tipo etico, legale o strategico (ad esempio la navigazione), mentre altri, come il *targeting* e l’eliminazione di elementi ostili attraverso uccisioni mirate, al contrario, possono essere fonte di una grandissima preoccupazione[3].

I sistemi autonomi possono funzionare attraverso ciò che viene talvolta descritto come “autonomia a riposo” e “autonomia in movimento”. L’autonomia a riposo descrive i sistemi che operano nel *software*, o nel mondo virtuale, mentre l’autonomia in movimento si riferisce ai sistemi che interagiscono con il mondo fisico. In un contesto militare, entrambe le tipologie possono causare molteplici preoccupazioni: i sistemi di autonomia a riposo potrebbero

[3] LAWS Intelligenza Artificiale e Robotica alla guerra, IRIAD Review, 5/2019, pag. 5.

adottare decisioni critiche sull'impiego della forza, anche letale, con impatti profondamente significativi. I sistemi di autonomia in movimento, invece, potrebbero includere sistemi d'arma autonomi letali (LAWS – *Lethal Autonomous Weapons Systems*) – cosiddetti "*robot killer*" in grado di assumere decisioni autonome su *targeting* ed *engagement* senza alcun tipo di *input* umano.

Occorre inoltre evidenziare che alcune caratteristiche preminenti dell'IA la rendono altamente versatile e adattabile a svariati ambiti: non è influenzata da fattori esterni; è permanente; può essere aggiornata quando necessario; garantisce coerenza logica. Pertanto, sebbene le applicazioni militari dell'IA siano spesso adattamenti delle tecnologie sviluppate nel settore commerciale, che continua a essere il motore dello sviluppo dell'IA, anche grazie a ingenti investimenti di capitali, numerosi governi hanno avviato programmi di sviluppo della stessa in ambito militare. Le organizzazioni governative di ricerca come la *Defense Advanced Research Projects Agency (DARPA) e* l'*Intelligence Advanced Research Projects Agency (IARPA)* negli Stati Uniti *e il Defense Science and Technology Laboratory (DSTL)* nel Regno Unito hanno sviluppato una serie di progetti di IA volti a incoraggiare la collaborazione con i settori commerciali e accademici per adattare, impiegare e applicare tecnologie di IA ai fini militari. Anche altre potenze militari, tra cui la Cina e la Russia, hanno da tempo avviato programmi attivi che prevedono importanti finanziamenti statali per lo sviluppo dell'intelligenza artificiale nel settore militare.

Anche all'interno del contesto NATO l'IA, definita come "*the ability of machines to perform tasks that traditionally require human intelligence*"[4], sta acquisendo sempre maggiore rilevanza nello studio dei possibili impieghi futuri. L'IA possiede alcune caratteristiche che permetterebbero di acquisire un vantaggio strategico rispetto alla controparte, soprattutto in un ambiente complesso e che necessita la valutazione di aspetti molto differenti tra loro, in tempi sempre più ristretti. La capacità di gestire grandi *set* di dati per trovare schemi risolutivi, svolgere compiti ripetitivi con un elevato grado di precisione e senza alcun calo di attenzione, fornire risposte rapide attraverso un'immediata analisi multifattoriale della situazione, rappresentano soltanto alcuni dei principali campi di sviluppo a cui le Forze Armate stanno destinando l'IA. I vantaggi, in tal senso, appaiono evidenti:

[4] "*An Artificial Intelligence Strategy for NATO*", 25 ottobre 2021, www.nato.int.

- Rapidità di analisi e azione;
- Esecuzione di semplici attività automatizzate su larga scala;
- Controllo di sistemi robotici e autonomi;
- Riconoscimento di schemi per prevedere tendenze future o rilevare anomalie;
- Classificazione e riconoscimento di oggetti e segnali;
- Ottimizzazione dei sistemi per raggiungere un obiettivo;
- Miglioramento della qualità delle decisioni.

2. Le applicazioni militari dell'Intelligenza Artificiale

I sistemi militari attualmente in fase di sviluppo contengono funzioni autonome multiple, suddivisibili in cinque macrocategorie: mobilità, *targeting*, intelligenza, interoperabilità e *health management*.

Con "mobilità" si intende la capacità del sistema di eseguire operazioni come la navigazione, il decollo e l'atterraggio in modo indipendente, che richiedono al sistema di determinare con precisione la propria posizione nonché pianificare ed eseguire una determinata rotta. Questi sistemi sono stati sviluppati per svolgere operazioni troppo complesse per un singolo individuo e possono essere utilizzati principalmente per ridurre i rischi in situazioni che richiedono grande precisione.

La seconda area di applicazione riguarda il *targeting*, che comprende l'identificazione, il tracciamento e la selezione del bersaglio, cioè l'insieme delle operazioni necessarie per individuare e colpire uno specifico obiettivo. Le criticità in quest'ambito riguardano la capacità di distinguere se un obiettivo è di tipo militare o civile e i limiti derivanti da cambiamenti improvvisi dello scenario, come quelli causati dalle condizioni meteorologiche. Tali problemi non sono dovuti a limiti negli algoritmi, ma piuttosto al fatto che il processo che parte dall'*input* e arriva all'*output* spesso rimane ancora poco chiaro e complesso da comprendere nella sua interezza. Il ciclo di *targeting* dovrebbe sempre essere supervisionato dall'uomo e solamente in situazioni in cui i tempi sono troppo brevi per l'intervento di un individuo, come nel caso di un attacco imminente, verrebbe utilizzata la modalità completamente autonoma capace di garantire una risposta pressoché immediata.

La terza categoria, "intelligenza", viene utilizzata per eseguire operazioni come la disattivazione di ordigni esplosivi, il rilevamento di intrusi nel perimetro di un'installazione o la localizzazione di postazioni nemiche da cui provengono colpi di arma da fuoco. In sostanza, per operazioni che richiedono una valutazione di dati per individuare eventuali comportamenti anomali rispetto ad una situazione di base, considerata "normale", in un'area

di operazione. Un'altra caratteristica di questa categoria è la capacità dei sistemi di raccogliere ed elaborare informazioni di *intelligence*, quali la *Map Generation*, il *Threat Assessment* e il *Big Data Analytics*.

Una quarta area di utilizzo è l'interoperabilità, ovvero la capacità di attrezzature militari e truppe sul terreno di operare in maniera sincronizzata per raggiungere uno scopo specifico. Questo implica la condivisione di informazioni tra sistemi che hanno obiettivi diversi, come nel caso del sistema *swarm*, che consente il dispiegamento simultaneo di reti cooperative di piattaforme *unmanned*. In tale ambito, il *software* di IA permette ai sistemi autonomi di agire come un intrecciato sciame intelligente: i velivoli, ispirati dagli sciami di insetti, sono in grado di operare in simbiosi per sopraffare i nemici, agendo autonomamente, senza controllo centrale. Le singole unità sono in grado di percepire l'ambiente circostante e le altre componenti del sistema, cooperando tra loro per svolgere un determinato compito. Tra le future applicazioni più attese c'è l'utilizzo di UAS (*Unmanned Aircraft System*) di piccole dimensioni per l'ISR (*Intelligence*, *Surveillance* and *Reconnaissance*) e per la difesa contraerea a basso raggio.

L'ultima area di applicazione, che è anche la meno comune, riguarda l'*health management of systems*, la quale include le operazioni di auto-rifornimento, autoriparazione e diagnosi di eventuali guasti, consentendo alla macchina di gestire alcuni aspetti del proprio funzionamento. L'auto-riparazione, che è una delle operazioni più complesse, consiste attualmente nell'espulsione o nella sostituzione di un modulo difettoso, mentre la capacità di rilevare e diagnosticare eventuali problemi è già in gran parte sviluppata.

Premesso quanto sopra, l'IA in ambito militare viene utilizzata anche per svolgere attività definite "noiose" o "sporche", in sostituzione della componente umana che potrebbe essere impiegata così in compiti più redditizi con una minore esposizione a potenziali situazioni di pericolo. La tecnologia autonoma, infatti, può evitare che gli operatori umani - risorsa sempre più preziosa e il cui "costo etico", oltre che materiale, ha raggiunto un livello molto elevato – siano impiegati in attività rischiose in ambienti non permissivi, come il rilevamento e lo smaltimento di esplosivi (militari o improvvisati – IED), operazioni di bonifica nonché ricognizione ostili. Inoltre, le Forze Armate, come tutte le grandi organizzazioni, si affidano ad un ingente volume di processi organizzativi, amministrativi e di gestione dei dati per raggiungere i propri obiettivi. Si tratta di attività di *routine* spesso monotone che possono rappresentare un carico di lavoro significativo per il personale, il quale, per la natura ripetitiva del compito, può anche essere indotto a errori causati da un calo di concentrazione. L'uso dell'IA per automatizzare queste attività, quindi, produce miglioramenti in termini

organizzativi, consentendo così al personale di occuparsi di questioni più complesse: gestione del personale, logistica e finanziaria/contabile.

Nell'ambito della sicurezza informatica, lo sviluppo di sistemi di IA è già impiegato in maniera concreta ed efficace, come nel caso di attacchi *malware*, che evolvono rapidamente e richiedono, pertanto, una velocità di risposta superiore a quella consentita dalla decisione umana. I sistemi di IA possono identificare in modo proattivo attività sospette e rispondere agli attacchi informatici in tempo reale, scansionando in maniera continuativa schemi di comportamento sospetti o codici potenzialmente maligni. Gli algoritmi utilizzati sono in grado di individuare e correggere le vulnerabilità di sicurezza del sistema in pochi secondi, rispetto ad un approccio convenzionale. D'altra parte, l'IA può essere utilizzata anche in maniera offensiva in una guerra cibernetica, impiegandola per identificare i punti deboli nelle difese delle reti avversarie e per progettare nuovi *malware*, capaci di penetrare il sistema informatico nemico in una pluralità di ambiti.

Inoltre, le operazioni militari e una vasta scala di sistemi d'arma dipendono dallo spettro elettromagnetico per una ampia gamma di funzioni che comportano l'utilizzo di frequenze radio, microonde, radar e comunicazioni satellitari. La guerra elettronica non è altro che l'azione intrapresa, attraverso l'impiego dell'energia elettromagnetica, per controllare lo spettro elettromagnetico, attaccando il nemico e bloccandone le iniziative. In tale contesto, l'IA riveste un ruolo cruciale in tutte le fasi della guerra elettronica attraverso la sua capacità di affrontare potenziali minacce grazie alla straordinaria rapidità di reazione alle avversità, rappresentando un vantaggio strategico rispetto alle tecniche convenzionali.

I sistemi di comando e controllo (C2), utilizzati per supportare i Comandanti nella direzione delle operazioni e nel monitoraggio delle forze assegnate, permettono di usufruire di informazioni in un formato facilmente comprensibile, agevolando e accelerando così l'intero processo decisionale. Lo studio dell'impiego dell'intelligenza artificiale in tale ambito riveste ormai da molto tempo notevole interesse, poiché l'IA potrebbe riuscire a sintetizzare dati provenienti da una serie di sensori molto diversi tra loro in una *Common Operational Picture* (COP) – di altissima valenza per i Comandanti – che mostri in maniera selettiva solamente gli oggetti di interesse sul campo di battaglia. Nei processi di C2, il supporto informativo che precede l'adozione di una decisione è essenziale, poiché permette al Comandante di individuare la soluzione ottimale tra una varietà di possibili opzioni, tenendo in considerazione le restrizioni e i criteri imposti per la precipua missione. Tuttavia, un *surplus* informativo rischierebbe di rendere il processo decisionale eccessivamente lento. Il procedimento di esecuzione deve tener conto di un'organizzazione logica delle attività, che richiede una pianificazione dettagliata, ma flessibile e in grado di adeguarsi alla dinamicità

e complessità dell'area di operazioni. In tale contesto, l'impiego dell'IA dovrebbe garantire sistemi in grado di offrire tutto il supporto decisionale necessario ai Comandanti, permettendo loro di rispondere ad eventi in evoluzione sul campo di battaglia attraverso l'analisi in tempo reale dei dati. Sebbene il giudizio umano sia essenziale nel processo decisionale di C2, la velocità e la capacità degli strumenti basati su IA possono implementare la progressiva estraniazione da attività periferiche, consentendo così al Comandante di poter impiegare le proprie unità in maniera rapida in differenti ambiti di impiego. Inoltre, l'uso dei sensori come mezzi per la raccolta "intelligente" di informazioni, attraverso osservazione e registrazione dei dati, consente la sorveglianza continua del campo di battaglia. Ciò consente di assumere decisioni quasi in tempo reale nel sistema di C2 e di colpire in maniera tempestiva e chirurgica i bersagli nemici, contribuendo alla diminuzione di perdite di vite umane e materiali.

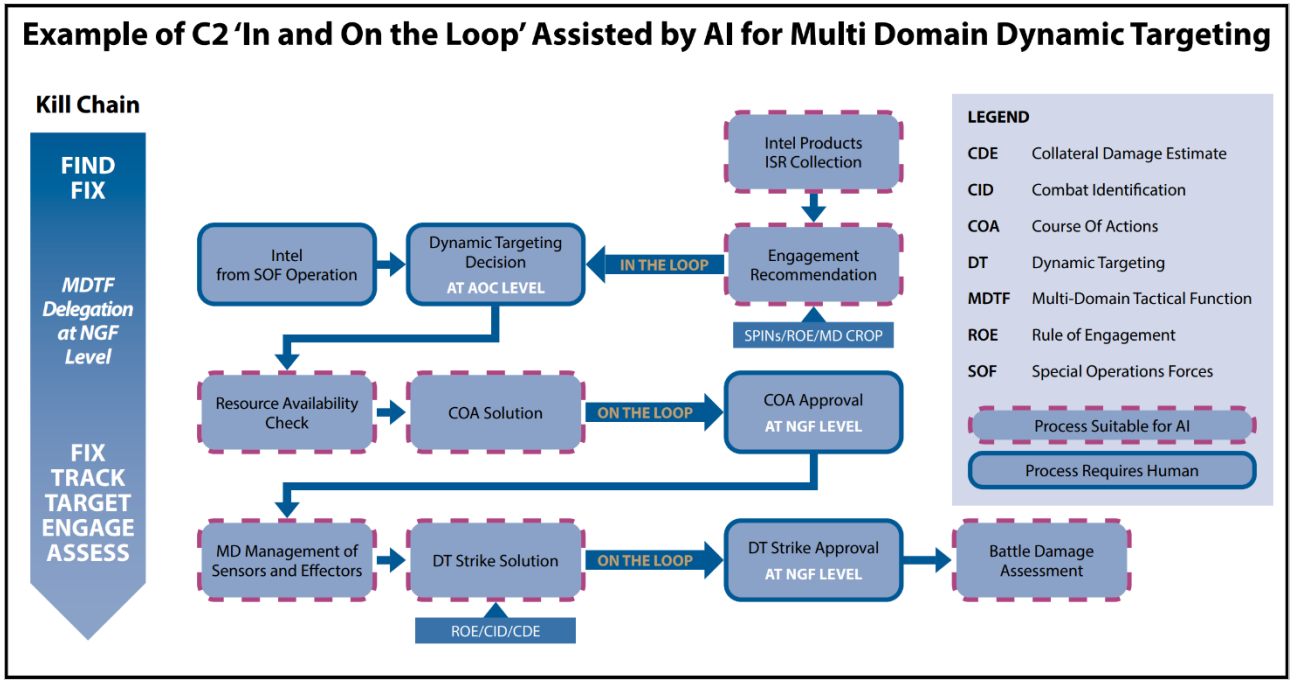

Fig. 1. Esempio di C2 secondo i principi *in/on the loop*.

I sistemi C4ISTAR (*Command, Control, Communications, Computers*, *Intelligence, Surveillance, Target Acquisition, Reconnaissance*) - in cui la moltitudine di sottosistemi componenti richiede una coordinazione ottimale in tempo reale e *online* - sono un elemento fondamentale in un ambiente operativo altamente complesso. La capacità dei *computer* di memorizzare grandi quantità di dati e di elaborarli in maniera estremamente veloce ed organica produce cambiamenti fondamentali nello sviluppo della cibernetica, che utilizza i successi nel campo dell'IA per quanto riguarda l'approccio sistematico alla realtà.

La "cibernetizzazione" della zona di combattimento, attraverso l'impiego di tecniche di alta precisione e attrezzature di armamento moderno, rappresenta un'ulteriore dimensione

in cui l'IA può essere ampiamente sfruttata, portando a significative modifiche del carattere della guerra. La precisione nell'impiego delle forze nella zona di combattimento si basa su un "sistema di sistemi", che consente alle forze di individuare obiettivi in tempo reale, garantendo la distruzione dei bersagli con danni collaterali minimi. Pertanto, le armi ad alta precisione basate sull'impiego dell'IA, della cibernetica e della tecnologia dell'informazione impongono una complessa garanzia informativa, eseguita in tempo reale con vari mezzi elettronici, tra cui radar, sensori, GPS e sistemi ISTAR. Queste tecnologie svolgono un ruolo decisivo nell'armamento moderno, enfatizzando la forza cibernetica delle azioni militari e trasformando le reti informatiche in potenziali obiettivi per il primo attacco.

La complessità del campo di battaglia, con la conseguente necessità di risposte efficaci in tempi rapidi, richiede una rete informativa globale basata su tecnologie avanzate come le reti di sensori e i sistemi d'arma autonomi che utilizzano l'intelligenza artificiale per adottare decisioni e completare missioni in tempi ristretti. L'IA viene utilizzata anche per assistere i piloti da combattimento, monitorare e riconfigurare le reti di comunicazione, e automatizzare la crittografia dei messaggi e delle informazioni provenienti dalle basi di dati. Inoltre, negli algoritmi evoluti, l'IA può essere utilizzata per l'identificazione e il *counter-fire* su obiettivi individuati dai radar, mediante immagini tridimensionali (3D) o bidimensionali (2D). Gli aeromobili e i veicoli controllati a distanza sono stati utilizzati efficacemente in combattimento negli ultimi due decenni ed è in fase di studio l'applicazione progressiva dell'IA a tali sistemi per consentire loro di operare in modo autonomo, così come per i veicoli senza conducente del settore civile, che impiegano sensori e *software* per percepire l'ambiente, riconoscere gli ostacoli, sintetizzare i dati dei sensori, navigare e comunicare con altri mezzi. La tecnologia ha raggiunto un grado di maturazione tale da consentire l'utilizzo di veicoli militari autonomi da parte delle principali potenze: droni capaci di volare in modo autonomo sono ormai in fase di sviluppo da parte dell'*US Air Force* (*Skyborg*) o della *Royal Navy* (*Madfox*).

L'esempio più evoluto di impiego militare dell'IA è rappresentato dai LAWS, definiti dal Comitato Internazionale della Croce Rossa (ICRC) come "*Any weapon system with autonomy in its critical functions. That is, a weapon system that can select (i.e. search for or detect, identify, track, select) and attack (i.e. use force against, neutralize, damage or destroy) targets without human intervention*"[5]. Costituiti da un sistema d'arma gestito attraverso l'IA, questi armamenti sono in grado di identificare, ingaggiare e distruggere, dopo un approfondito ciclo di *targeting*, un bersaglio senza il diretto intervento dell'uomo. Tali

[5] Neil Davison, *"A legal perspective: autonomous weapon systems under international law", International Committee of the Red Cross*, p. 5.

apparati si basano essenzialmente sulla combinazione di un sistema di sensori che monitora l'ambiente circostante e un sistema di IA in grado di identificare un determinato oggetto/individuo/mezzo come potenziale obiettivo o nemico, decidendo se ingaggiarlo o meno e quale arma scegliere per distruggerlo o sortire gli effetti desiderati.

Esistono già sistemi di armamento in cui un algoritmo prende la decisione di sparare, sostituendosi così direttamente all'operatore militare. Tali sistemi in grado di acquisire ed ingaggiare autonomamente bersagli sono, al momento, per lo più concepiti in chiave difensiva. Queste apparecchiature sono destinate ad essere impiegate sotto la supervisione umana e a fare fuoco autonomamente in situazioni in cui il tempo di reazione dell'uomo viene considerato troppo breve per la portata della minaccia che si profila all'orizzonte. Tuttavia, stanno emergendo anche armi offensive in grado di acquisire ed impegnare bersagli autonomamente. Le munizioni *Fire-and-forget*, come il missile MBDA *Brimstone* in dotazione alla RAF[6] e la munizione di crociera IAI *Harop*[7] sono in grado di selezionare i propri bersagli, mostrando per certi versi caratteristiche simili a quelle di un'arma letale autonoma. I sistemi militari sono diventati sempre più complessi, versatili e automatizzati, con la capacità di operare in modo autonomo, raccogliendo ed elaborando informazioni.

L'IA, come analizzato, sta assumendo un ruolo profondamente significativo nello specifico settore. Detto ciò, gli orizzonti futuri risultano ancora incerti, benché appaia evidente che la capacità d'impiego dell'IA nel contesto costituirà l'elemento che garantirà la supremazia nel campo militare.

---

6 Il *Brimstone* è un missile aria-superficie o superficie-superficie a guida radar sviluppato negli anni '90 su richiesta della *Royal Air Force* per l'utilizzo contro formazioni di veicoli blindati. Nel marzo 2019 è stata lanciata la terza versione equipaggiata con un radar ognitempo in grado di resistere alle contromisure dispiegate dal nemico. I missili, del tipo *fire-and-forget,* possono essere programmati per ricercare bersagli solo in una determinata zona e sono in grado di autodistruggersi nel caso in cui non individuino bersagli nella zona richiesta.

7 La IAI *Harop* è una munizione vagante sviluppata dalla divisione MBT di *Israel Aerospace Industries*. Le munizioni vaganti sono progettate per vagare sopra il campo di battaglia e attaccare gli obiettivi schiantandosi contro di essi ed esplodendo. È stato utilizzato per la prima volta in combattimento dall'Azerbaigian nel conflitto del Nagorno-Karabakh nell'aprile 2016, per distruggere gli autobus dei soldati armeni trasportati in prima linea e per distruggere alcuni posti comando armeni.

CAPITOLO II: L'AUTONOMIA NEL RAPPORTO HUMAN-MACHINE TEAMING

1. Classificazione dei diversi approcci: *Human in/on/out of the loop*

L'autonomia è intesa come la capacità di una macchina di eseguire un compito senza l'intervento umano. Essa, quindi, descrive un sistema capace di operare in modo indipendente per un certo periodo di tempo senza il diretto intervento umano. Determinarne il grado o l'entità è rilevante per comprendere le sfide e le opportunità che derivano dai sistemi autonomi. Nell'ottobre del 2016 il *Joint Center for Robotics and Autonomous Systems* (JCRAS) ha definito l'autonomia come "il livello di indipendenza che gli esseri umani concedono a un sistema per eseguire un compito assegnato. È la condizione o la qualità di essere autogovernanti per raggiungere un determinato *task* sulla base della consapevolezza situazionale del sistema (sensazione, percezione, analisi integrate), della pianificazione e della decisione. L'autonomia si riferisce a uno spettro di automazione in cui il processo decisionale può essere adattato ad una missione specifica, al livello di rischio e al grado di interrelazione uomo-macchina"[8].

Esistono tre dimensioni di base dell'autonomia:

- il tipo di compito che la macchina sta svolgendo;
- il rapporto tra l'essere umano e la macchina durante l'esecuzione di quel determinato *task*;
- la sofisticazione della decisione della macchina quando svolge il compito.

Il grado di autonomia utilizzato dai sistemi d'arma, secondo lo stato attuale di maturità tecnologica, dipende dal livello di intervento di un operatore umano nella loro catena di distribuzione e impiego. I sistemi d'arma autonomi possono essere suddivisi, dunque, in tre differenti *cluster* che descrivono e definiscono il grado di ibridazione e interazione *human-machine teaming*:

a) Human in the loop (HIL):

Sistemi d'arma che sfruttano l'autonomia per colpire *target* individuali o gruppi specifici di obiettivi tramite l'intervento di un operatore umano che coordina ogni singola fase del processo decisionale. Un esempio potrebbe essere rappresentato dalle munizioni guidate in cui la tecnologia dell'arma assiste l'operatore durante il ciclo di *targeting*. L'operatore che impiega l'arma, dunque, conosce quali obiettivi devono essere colpiti e compie una scelta consapevole circa il bersaglio da distruggere. Appare evidente che la concettualizzazione

[8] https://sgp.fas.org/crs/weapons/R45392.pdf.

dottrinaria del principio HIL implichi innanzitutto la necessità di contenere il livello di discrezionalità attribuibile a un sistema d'arma autonoma, qualora essa venga dispiegata in funzione offensiva. Alcune tipologie di sistemi d'arma, come ad esempio le munizioni a guida di precisione, anche note con l'acronimo PGM, possono essere incluse nella categoria HIL, poiché l'operatore umano, impegnato nel coordinare la fase di lancio, conosce a priori gli obiettivi da ingaggiare. Alcune *signature-guided munitions*, ad esempio, annoverano nella rispettiva componentistica tecnica sensori passivi in grado di percepire i segnali dell'ambiente circostante per avvicinarsi progressivamente al bersaglio. Altre, invece, sfruttano sensori attivi, come i *radar*, per emettere segnali e tracciarne il ritorno dal *target*. Anche i missili *fire-and-forget*, chiamati così poiché non necessitano del controllo diretto dell'operatore dopo il lancio, hanno raggiunto ormai un tale livello di sofisticazione tecnologica da determinare, proprio in virtù delle informazioni fornite al sistema a premessa del lancio (misurazione *radar*, immagini infrarossi e coordinate GPS), l'estraniazione parziale del singolo individuo. Tali apparati presentano, infatti, peculiarità tecniche che includono, tra l'altro, la capacità di scovare il bersaglio entro una determinata area e di autodistruggersi se incapaci di ingaggiare l'obiettivo nella zona designata. Questa tipologia di armi può operare secondo un *indirect mode*: quando la posizione del *target* non è pienamente visibile al pilota, il missile può ingaggiare il bersaglio sfruttando il proprio sistema di navigazione e i sensori di scoperta. Tuttavia, l'operatore, pur monitorando l'intero ciclo di *targeting* nelle fasi antecedenti il suo lancio, non si trova nella condizione di interagire con l'arma nelle fasi finali dell'ingaggio. Il ruolo dell'operatore umano nell'ambito d'impiego dell'intelligenza artificiale è spesso discusso anche nel contesto della loro relazione in seno al ciclo OODA (Osservare, Orientare, Decidere, Agire). Ad oggi, numerosi studiosi attribuiscono particolare rilevanza al ruolo dell'uomo nel ciclo dell'IA in modalità HIL, al fine di mitigare i rischi etici e operativi intrinsechi al suo impiego in ambito militare. Infatti, in tale architettura concettuale, l'operatore può intervenire direttamente nel ciclo del sistema, garantendo che lo stesso rispetti le leggi e le regole di ingaggio. L'essere umano, in questo caso, funge essenzialmente da garante morale, preservando la dignità umana e compensando la criticità intrinseca all'IA riscontrabile nell'applicazione dei principi di discriminazione, proporzionalità e necessità in ambienti operativi particolarmente complessi. La centralità del decisore umano, che resta in pieno controllo del sistema, riduce o minimizza le possibilità di corruttibilità dello stesso, fornendo un controllo aggiuntivo sul suo corretto impiego.

b) Human on the loop (HOL):

Sistemi d'arma che impiegano l'autonomia per selezionare e colpire obiettivi specifici ed in cui l'operatore osserva il funzionamento della macchina riservandosi la facoltà di intervenire solo nel caso in cui si concretizzino comportamenti avversi da parte del sistema. Attualmente almeno trenta Nazioni utilizzano sistemi difensivi *human-supervised* con un elevato grado di autonomia, dove gli operatori sono *on the loop* per la selezione e la distruzione di obiettivi specifici. Fino ad oggi, questi sistemi d'arma sono stati dispiegati soprattutto in chiave difensiva in situazioni dove il tempo di reazione è così limitato che sarebbe impossibile per gli esseri umani rispondere abbastanza rapidamente per difendersi da una sopraggiunta situazione di pericolo. Gli operatori, che monitorano l'intero ciclo, possono intervenire per disattivare il sistema d'arma, ma non prendono alcuna decisione in merito all'obiettivo da abbattere. L'impiego di sistemi autonomi consente di salvaguardare l'incolumità delle truppe dispiegate nello svolgimento di talune attività tattiche nel teatro operativo, migliorando allo stesso tempo le capacità di reazione, soprattutto in chiave difensiva, di fronte a potenziali minacce. Nel contesto di sistemi d'arma difensivi caratterizzati da minimi tempi d'intervento, quali ad esempio il *Phalanx*[9] o il C-RAM, sviluppati in un'ottica di adattamento e contrasto a minacce emergenti, la necessità dell'autorizzazione al fuoco da parte dell'operatore (modalità semi-automatica) potrebbe rallentare irragionevolmente il sistema, traducendosi in una risposta tardiva. Proprio in quest'ottica l'*Iron Dome*, dispiegato da Israele per prevenire ed intercettare potenziali attacchi missilistici lanciati da Hamas verso il Paese, rappresenta un sistema di difesa mobile capace di abbattere il tempo necessario all'operatore per analizzare i dati e adottare misure consapevoli sull'obiettivo da ingaggiare[10]. Il *software* del computer, attorno al quale ruota l'intero centro di controllo del sistema antimissilistico, modula i parametri del lancio (traiettoria e punto di intercetto) sulla base di un algoritmo programmato direttamente dall'uomo. In aggiunta, laddove la minaccia fosse rappresentata da un avversario che sfrutti i propri sistemi IA alla massima velocità/capacità esprimibile, l'approccio "*in the loop*" si tradurrebbe in un forte svantaggio competitivo. Quanto precede determina particolari

[9] Il *Phalanx*, dispiegato su navi come reazione di difesa rapida contro potenziali minacce, è costituito da una mitragliatrice *radar-guided* montata su una basa girevole capace autonomamente di effettuare la propria ricerca, valutazione, tracciamento, ingaggio e distruzione. Il *Centurion*, invece, rappresenta la variante terrestre del Phalanx ed è essenzialmente equipaggiato con capacità cosiddette *sense*, *warn and response*. Sebbene gli operatori di tali sistemi d'arma difensivi si riservino la facoltà di intervento, adottando la decisione di ingaggiare bersagli, l'alta velocità di missili multipli in arrivo rende inevitabilmente l'operazione autonoma quella preferita.

[10] Il sistema antimissilistico *Iron Dome* è essenzialmente costituito da tre componenti: un radar EL/M-2081 che traccia e rileva le traiettorie dei missili o dei razzi; un sistema di comando e controllo BMC che classifica potenziali minacce in arrivo e calcola il punto d'impatto in aria; l'unità di lancio dei missili è composto da venti missili Tamir dotati di sensori elettro-ottici, guida GPS e otto pinne per virare in aria. Sebbene la Cupola di Ferro non richieda l'intervento dell'operatore, è possibile attivare il suo funzionamento anche manualmente.

benefici derivanti dal collocamento dell'operatore umano in un ruolo di monitoraggio del sistema, determinando al contempo la necessità di garantire un efficace meccanismo d'intervento da parte dello stesso. Proprio questo aspetto dell'architettura con modalità HOL solleva un particolare interrogativo, ovvero come assicurare un ragionevole margine d'azione all'operatore, permettendogli di intervenire abbastanza rapidamente laddove risultasse necessario interrompere l'ingaggio. Un altro elemento critico rilevabile nell'ambito di questo approccio è legato all'impiego dei sistemi autonomi in un contesto elettromagnetico degradato o più in generale non permissivo, nel quale una mancanza di comunicazioni potrebbe negare la capacità di monitoraggio dell'operatore, compromettendone inesorabilmente le possibilità d'intervento.

L'ambito di impiego di sistemi HOL si estende anche alle operazioni di *law enforcement* o alle attività di pattugliamento di ampie aree geografiche che demarcano i confini territoriali tra entità statuali in contrasto tra loro. Proprio la Corea del Sud ha dispiegato, lungo la zona demilitarizzata che attraversa e divide a metà la penisola coreana, l'SGR-A1, sviluppato da Samsung e concepito con l'obiettivo di sorvegliare l'area di confine. Tale sistema, grazie ai suoi sensori, è in grado di accertare la presenza di potenziali incursori, rilevare bersagli fino a 2 Km di distanza, riconoscere movimenti e ingaggiare obiettivi che non rispondono al suo avvertimento vocale. A tal proposito possiamo citare anche il mezzo di superficie a controllo remoto *Protector*, impiegato dalla Marina israeliana per le operazioni di pattugliamento e protezione delle infrastrutture portuali e industriali strategiche disseminate lungo il settore costiero. L'IA, in questo caso, contribuisce ad ampliare le capacità e le prestazioni complessive del sistema difensivo.

Fig. 2. SGR-A1

In chiave offensiva, invece, numerose *loitering munitions*[11], quali *Harpy* o *Harop*, sono dotate di telecamere elettro-ottiche e infrarossi che consentono di individuare l'obiettivo, che viene ingaggiato autonomamente dal sistema. Proprio il drone israeliano, sviluppato dall'*Israel Aerospace Industries* (IAI), è in grado di "*search, find, identify, attack and destroy targets, and performs battle damage assessment as an autonomous platform operations*"[12]. Il sistema americano *Advanced Targeting and Lethality Automated System*, meglio noto con l'acronimo di ATLAS, è stato sviluppato per l'esercito americano con il chiaro obiettivo di integrare ed applicare i progressi ottenuti nel campo del *machine learning* e del *computer vision* ai veicoli terrestri. Attraverso tale sistema il processo di acquisizione, identificazione e ingaggio degli obiettivi può avvenire tre volte più velocemente di quanto attualmente possibile tramite gli operatori umani. Un'ulteriore declinazione di AWS, secondo una prospettiva HOL, ci viene offerta dal settore aerotattico e trova la sua massima espressione nel concetto di *loyal wingman*, attualmente al centro di un'intensa e laboriosa attività di ricerca, sviluppo e sperimentazione da parte delle più importanti industrie internazionali, quali Boeing, Northrop, Lockheed Martin e Sukhoi. Tale sistema entra in azione sotto la supervisione di un velivolo pilotato (*parent ship*), incrementando il volume di fuoco per abbattere unità nemiche con "tattiche più aggressive rispetto a quelle consentite ad un velivolo pilotato"[13]. L'*anti-submarine warfare* (ASW) *Sea Hunter* (figura 3), lanciata nel 2016 dagli USA, è stata concepita con l'obiettivo di tracciare e scovare sottomarini nemici[14].

Fig. 3 Sea Hunter

Questa particolare Nave, inquadrata nel settore degli *Unmanned Surface Vehicles* (USV), può svolgere operazioni di sminamento o attività di *Intelligence e Survaillance*.

---

11 Sistema d'arma in grado di orbitare su un'area assegnata in cerca di un obbiettivo che, una volta individuato, può essere attaccato.

12 "*Autonomous epos system and changing norms in international relations*", *Review of international studies*, Ingvild Bode, Hendrick Huelss, 2018, p. 9.

13 "*Autonomous Weapon System*, possibili sviluppi futuri nell'ambito del potere aerospaziale", Valerio Ceccarelli, Riccardo Musicco, Mattia Nucciarelli, Mirko Tagliamento, Nicola Zanon, Rivista Aeronautica, n. 4/2022, p. 55.

14 L'unità navale *Sea Hunter* è inquadrata nel progetto avviato dal DoD meglio noto come *Ghost Fleet Overlord*. Il progetto, ormai iniziato nel 2018, ambisce a integrare gli *unmanned surface vehicles* nella flotta navale statunitense.

Sebbene l'unità navale sia controllata da remoto "*the progress of the platform's tecnology, and the rapid advancements of algorithms enabling greater levels of autonomy, have inspired the Navy to become thinking about additional missions… so that it can conduct surface warfare missions, fire weapons and launch electronic attack*"[15]. È auspicabile che armi autonome, quali ad esempio la *Sea Hunter*, possano contribuire, nel prossimo futuro, a sconvolgere le consolidate basi della deterrenza nucleare basata sul principio della mutua distruzione assicurata (MAD, *Mutual Assured Destruction*). Ancora, nell'ambito delle operazioni navali appare rilevante notare come la *Royal Navy*, già da qualche anno, stia proficuamente lavorando al programma *Startle*, progettato per fornire raccomandazioni e avvisi nonché generare decisioni immediate contro potenziali minacce. Tale sistema, sfruttando il processo di *deep learning*, consente agli operatori di incrementare la *Maritime Domain Awareness* soprattutto durante il processo di identificazione, tracciamento e ingaggio di possibili obiettivi. Tali iniziative, tra le quali possiamo annoverare anche il *Consolidated Afloat Networks and Enterprise Services* (CANES)[16] sviluppato appositamente per l'US *Navy*, sono state ideate con l'obiettivo di introdurre, proprio attraverso l'impiego dell'IA, *Combat Management system* destinati a migliorare le capacità di *tracking* di una vasta gamma di minacce sulla base di una valutazione che tenga anche conto delle peculiarità di un determinato scenario operativo. Il pionieristico, e forse anche il più noto, programma militare di IA del Pentagono è il progetto *Maven*. L'obiettivo principale di *Project Maven* è quello di sfruttare i vantaggi dell'apprendimento automatico (*machine learning*) e della visione artificiale (*computer vision*) per accelerare il *targeting* e altre decisioni in ambienti di conflitto difficili. *Maven* attinge alla tecnologia dei droni per i suoi *input* visivi e altri sensori al fine di identificare e tracciare potenziali obiettivi in tempo reale e fornire all'operatore informazioni per l'innesto cinetico in velocità. Esistono, poi, una serie di programmi profondamente ambiziosi ma ancora in fase di sviluppo: il *Collaborative Operations in Denied Enviroment* (CODE) ambisce a sviluppare nuovi modelli algoritmici o *software* per i velivoli *unmanned*. Tale programma consentirebbe, infatti, di estendere le abilità operative dei droni e ampliare, conseguentemente, le capacità per le Forze Armate di condurre attacchi in uno spazio aereo ostile. I velivoli equipaggiati con tali sistemi, pur restando sotto il controllo dell'operatore, potrebbero selezionare e ingaggiare bersagli. Il progetto *Global Information Dominance Experiment* (GIDE), invece, sfrutta l'intelligenza artificiale e il *machine learning* per combinare, categorizzare e sintetizzare i dati raccolti da

[15] Bode, Huelss, *Autonomous weapons system*., op. cit., p. 9.

[16] Il programma CANES rappresenta il consolidamento e il miglioramento di cinque programmi di rete per fornire un ambiente informatico comune a unità navali di superficie e subacquee al fine di migliorare l'interoperabilità e diminuire i cicli operativi per difendersi da attacchi *cyber*.

altre fonti con l'obiettivo di elaborare potenziali previsioni capaci di conferire un vantaggio strategico e operativo rilevante alle Forze Armate.

Terrence John O'Shaughnessy – Generale dell'aeronautica statunitense nonché ex Comandante del NORAD[17] e del USNORTHCOM[18] – ha sostenuto la necessità di muovere verso un approccio di tipo HOL, evidenziando così l'esigenza di spingere il controllo umano più distante dal processo decisionale automatizzato. Questo, infatti, permetterebbe comunque all'operatore di supervisionare l'intero sistema, ma l'intelligenza artificiale entrerebbe in funzione senza la previa autorizzazione del singolo individuo.

c) Human out of the loop (HOOL):

Tali sistemi d'arma, secondo la direttiva DoD 3000.09 emanata dal Dipartimento della Difesa statunitense, sfruttano l'autonomia per selezionare e colpire un *target* specifico senza ingerenza alcuna da parte dell'operatore (HOOL). Il Governo olandese, dal canto suo, ha definito AWS come un sistema in grado di "*selects and engages targets* […] *on the understanding that an attack, once launched, cannot be stopped by human operator*"[19]. Quest'ultimo punto, dunque, appare particolarmente rilevante poiché ci consente di distinguere AWS da quei sistemi ancorati al principio del "dualismo dinamico" poiché caratterizzati da un certo grado di interazione e ibridazione *human-machine teaming*. L'impiego di sistemi d'arma autonoma, quindi, implica, da un lato, la completa estraniazione (*out/off the loop*) dell'operatore dall'intero processo decisionale e, dall'altro, la sua definitiva alienazione dal ciclo di *targeting*. Tuttavia, anche in questa circostanza e come evidenziato in talune ricerche scientifiche, possono essere escogitati alcuni meccanismi tecnici che consentono all'individuo di preservare una sorta di capacità di controllo e monitoraggio da remoto. L'algoritmo, ad esempio, potrebbe contemplare un procedimento di autodistruzione qualora la missione non possa essere pienamente completata; quando le condizioni del contesto operativo appaiono particolarmente complesse tanto da alterare o addirittura corrompere l'algoritmo, l'operatore potrebbe disattivare il sistema attraverso un *kill-switch*[20]*.* Un processo di autodistruzione, inoltre, assicurerebbe che un AWS, in caso di eventi avversi, non cada nelle mani delle forze nemiche.

---

17 *North American Aerospace Defence Command*. Struttura di comando congiunta USA-Canada volta a fornire un quadro di situazione comune ad entrambi i Paesi in merito a posizione, direzione, velocità e natura di ogni velivolo nello spazio aereo Nord americano.

18 United States Northern Command.

19 Ibidem.

20 Ibidem.

2. I vantaggi e le vulnerabilità connaturate all'IA nell'ambito delle LAWS

La guerra è un'attività intrinsecamente rischiosa in cui le parti in gioco utilizzano la forza per perseguire obiettivi politico-militari e gli effetti che ne conseguono sono altamente incerti, sia per i combattenti che per i non combattenti. Quanto precede sarà ancor più vero con il progressivo aumento dell'impiego dei sistemi d'arma autonomi, che promettono di concretizzare la *third revolution in warfare*, proponendo nuove sfide ed opportunità nell'ambito del conflitto. Nel presente paragrafo saranno analizzati i vantaggi e le vulnerabilità derivanti dall'applicazione dell'IA nell'ambito militare.

a) Vantaggi

L'impiego di sistemi autonomi promette di fornire notevoli vantaggi in ambito militare che verranno esaminati secondo la seguente suddivisione:

CLASSIFICAZIONE VANTAGGI DELL'INTELLIGENZA ARTIFICIALE		
ETICO/LEGALI	**OPERATIVI**	**STRATEGICI**
• Reazioni cognitive/emotive	• Modellizzazione • *Operational reach* • *Data fusion, sharing & decision making*	• Economia delle forze • Riduzione delle perdite

Questa ripartizione sarà utile per affrontare l'argomento in maniera organica, ma non va intesa in modo eccessivamente rigida in virtù della rapida evoluzione dell'IA nell'ambito degli affari militari. Si procede di seguito all'analisi esplicativa delle categorie precedentemente espresse:

- Etico / Legali – Reazioni cognitive / emotive

  Alcune ricadute positive correlate all'impiego dei sistemi d'arma autonomi possono essere individuate nella sfera etico/legale. In futuro, infatti, i sistemi autonomi potranno ridurre talune criticità sul campo di battaglia, poiché, diversamente dai combattenti, non saranno sottoposti a *stress* psicologico correlato alla necessità di preservare la propria incolumità fisica. Il loro giudizio, inoltre, non risulterà mai legato ad emozioni né tantomeno potrà essere soggetto a distorsioni cognitive tipiche dell'essere umano. Quanto precede potrebbe apparire particolarmente rilevante nel prevenire/ridurre incidenti da fuoco amico, ovvero danni collaterali legati a dinamiche "*fire first and ask later*"[21]. In aggiunta, la ridotta esposizione a situazioni particolarmente logoranti per il personale militare potrebbe implicare la

[21] "*The Case for Ethical Autonomy in Unmanned Systems*", Ronald C. Arkin, *Journal of Military Ethics* 9, no. 4-2010.

riduzione di eventi potenzialmente avversi per la popolazione civile (es. ritorsioni e violenze in generale).

- Operativi

  In questa sezione vengono analizzati i vantaggi connessi alla condotta delle operazioni militari e le ricadute positive da loro determinate.

- Modellizzazione

  Le tecnologie di *Machine* o *Deep learning* alla base dei sistemi di IA sono in grado di individuare schemi o possibili irregolarità nei fenomeni da esse esaminati. L'IA, utilizzando un insieme di dati di addestramento e osservazioni, potrà generare modelli descrittivi in seguito all'analisi delle differenti situazioni osservate e modulare le proprie azioni sulla base di tale conoscenza. Questa capacità e rapidità di calcolo consentirà di fornire abilità analitiche e predittive senza precedenti in ambito militare.

- Operational reach

  L'utilizzo di un sistema d'arma che non richieda all'operatore umano di esporsi direttamente a una minaccia concreta permetterà di estendere la portata delle operazioni militari, conseguendo peraltro risultati significativi in aree che non consentono di operare efficientemente con i sistemi attualmente in uso. Secondo il Dipartimento della Difesa statunitense questi sistemi sono "particolarmente adatti a lavori noiosi (es. sortite di lunga durata o compiti ripetitivi), sporchi (es. attività che espongono ad ambienti contaminati) e pericolosi (es. bonifica di ordigni esplosivi)"[22].

- Data fusion, sharing & decision making

  Uno dei grandi punti di forza dei sistemi in esame è la capacità di condividere informazioni in *real time/near real time* e di processarle in maniera rapida, completa ed efficace. In uno scenario operativo futuro, tutti i sistemi autonomi potrebbero operare in simbiosi, scambiandosi costantemente dati informativi in modo tale da non configurarsi come sistemi isolati, ma integrati in un *system of systems*. Questo incessante scambio di informazioni consentirà di comporre ed aggiornare la *Common Operational Picture* (COP).

- Strategici

  Questa categoria evidenzia i benefici per lo strumento militare a più ampio spettro, incrementandone le capacità nette e determinando quindi vantaggi concreti nei confronti di eventuali avversari.

[22] "*Unmanned Systems Roadmap* 2007-2032" US DoD, James R. Clapper Jr.

- Economia delle forze

  L'impiego dei sistemi autonomi permetterà una riduzione delle risorse umane dispiegate nei teatri operativi, determinando una maggiore efficienza dello strumento militare. In aggiunta, la minor concentrazione di unità implicherà la conseguente riduzione del *footprint* logistico delle unità, traducendosi in un minor impatto finanziario delle stesse.

- Riduzione delle perdite

  Al momento l'uomo viene impiegato nella quasi totalità delle missioni, anche in quelle che comportano un rischio concreto per la propria incolumità. Si pensi a missioni collegate a operazioni di bonifica di ordigni esplosivi (EOD - *Explosive Ordnance Disposal*), alla conduzione di attività in ambienti contaminati o addirittura ad immersioni compiute in condizioni non permissive. L'impiego di sistemi autonomi, dunque, consente di preservare l'integrità fisica dell'operatore in attività ritenute pericolose, riducendo sensibilmente le perdite e tutelando, almeno parzialmente, la società dalle conseguenze negative del conflitto.

b) Rischi

CLASSIFICAZIONE RISCHI DELL'INTELLIGENZA ARTIFICIALE		
ETICO/LEGALI	**OPERATIVI**	**STRATEGICI**
• Conformità al DI ed a*ccountability*	• *Hacking*, inquinamento dei dati e attacchi avversari • Incidenti e rischi emergenti	• Uso della forza su vasta scala • Gestione dell'*escalation* • Proliferazione • Stabilità strategica

I sistemi di IA attualmente in fase di sviluppo possono rappresentare indubbiamente minacce per la vita, i diritti umani e il benessere sociale, con l'effetto finale che una corsa indiscriminata tra gli Stati alle applicazioni militari dell'IA, senza mettere in atto le adeguate precauzioni, potrebbe ridurre, piuttosto che aumentare, la sicurezza. Il loro impiego appare particolarmente controverso e incontra, anche nell'opinione pubblica, numerose opposizioni, molte delle quali connesse al loro potenziale uso in chiave offensiva. Prendendo spunto dallo studio condotto dalla RAND *Corporation* pubblicato nel marzo 2020[23], i rischi connessi all'impiego dei sistemi d'arma autonomi possono essere suddivisi come segue:

[23] *"Military applications of Artificial Intelligence"*, RAND Corporation, Marzo 2020.

È importate notare che questa classificazione dei rischi è da intendersi come una panoramica concettuale in continua evoluzione e non costituisce, dunque, una rigida ripartizione dottrinaria. Si procede di seguito all'analisi esplicativa delle categorie sopra indicate:

- Etico / Legali – Conformità al DI ed *accountability*

  Nel corso dei conflitti i Comandanti che pianificano e decidono un attacco hanno l'obbligo di distinguere tra combattenti e non combattenti secondo quanto previsto dalle Convenzioni di Ginevra del 1949 e dai relativi Protocolli addizionali del 1977. Devono altresì garantire che il danno causato sia proporzionale agli obiettivi militari da conseguire e adottare le necessarie precauzioni per proteggere i civili, scegliendo il mezzo di guerra che possa arrecare il minor danno possibile. La perdita di vite civili per danni collaterali è legittima solo se proporzionata all'obiettivo militare, secondo il principio di necessità e i Comandanti possono essere chiamati a rispondere di eventuali violazioni del Diritto Internazionale Umanitario (DIU) e del Diritto umanitario (DU) di fronte ad apposite Corti di giustizia che ne accerteranno la condotta.

  L'aspetto più difficile da sostenere è la certezza che le armi autonome possano agire sempre in conformità "agli usi stabiliti fra le nazioni civili, dalle leggi d'umanità e dalle esigenze della coscienza pubblica", previsti dalla "clausola di Martens" inclusa nel preambolo delle Convenzioni dell'Aja del 1899 e valida in ogni contesto. Sono state sollevate numerose preoccupazioni riguardo l'impiego di sistemi di IA che potrebbero portare ad una "lacuna di responsabilità" nel caso in cui si verifichino violazioni del DIU. Non è chiaro, infatti, chi sarebbe considerato responsabile qualora si verificassero eventi avversi o imprevisti rispetto a quanto pianificato, colpendo *target* errati o coinvolgendo i civili. Siffatte condotte potrebbero prefigurarsi, dunque, come crimini di guerra. I problemi relativi all'*accountability gap*, esaminati in questa sezione, verranno affrontati in maniera più approfondita e analitica nel corso del presente lavoro, al fine di delineare più dettagliatamente i vari aspetti della tematica ed analizzare le possibili soluzioni.

- Operativi

  A questa categoria si riferiscono tutti i rischi ascrivibili all'impiego di sistemi IA nelle loro applicazioni militari, determinando un funzionamento inatteso degli stessi.

- Fiducia ed affidabilità

  In questa sottocategoria rientra sia il rischio di accordare eccessiva fiducia in tali tecnologie che quello di non confidare affatto sull'efficienza di questi sistemi. Tale sentimento di sfiducia può essere spiegato con il cosiddetto paradigma della "scatola nera" secondo cui l'operatore non sarebbe in grado di comprendere come o perché l'IA sia giunta a certe conclusioni. Questo implica addirittura il rischio di sfociare in usi impropri o addirittura sub-ottimali. Sussistono, infatti, diverse ambiguità che non consentono di manifestare un certo grado di fiducia nei sistemi abilitati all'IA: l'eccessiva confidenza, meglio nota come a*utomation bias*, comporta un'accettazione pressoché incondizionata della soluzione o del piano d'azione offerti dal sistema senza prima verificarne l'effettiva bontà. Nelle successive figure 4, 5, 6A e 6B si riportano alcuni dati relativi alla percezione relativa all'impiego di AWS sviluppata dal *Journal of Indo-Pacific Affairs* nel 2020[24].

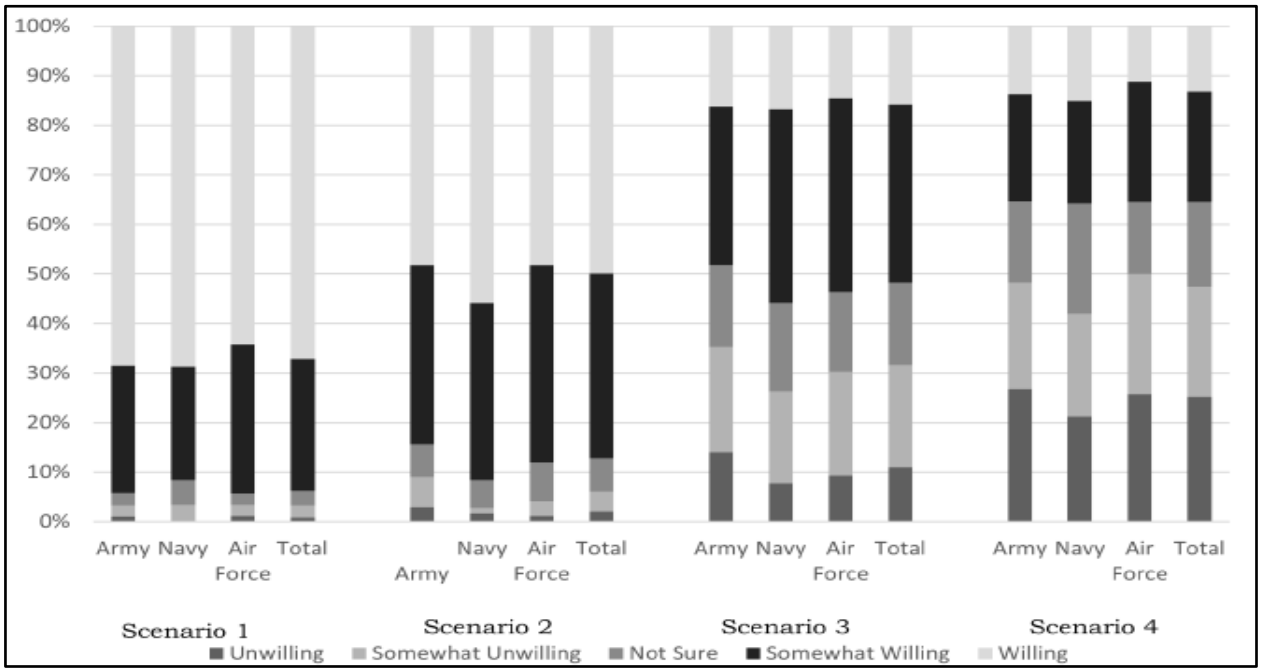

Fig. 4 Willingness to deploy alongside autonomous systems.

[24] *Journal of Indo-Pacific affairs*, Dr. Jai Galliott and Dr. Austin Wyatt, 2020.

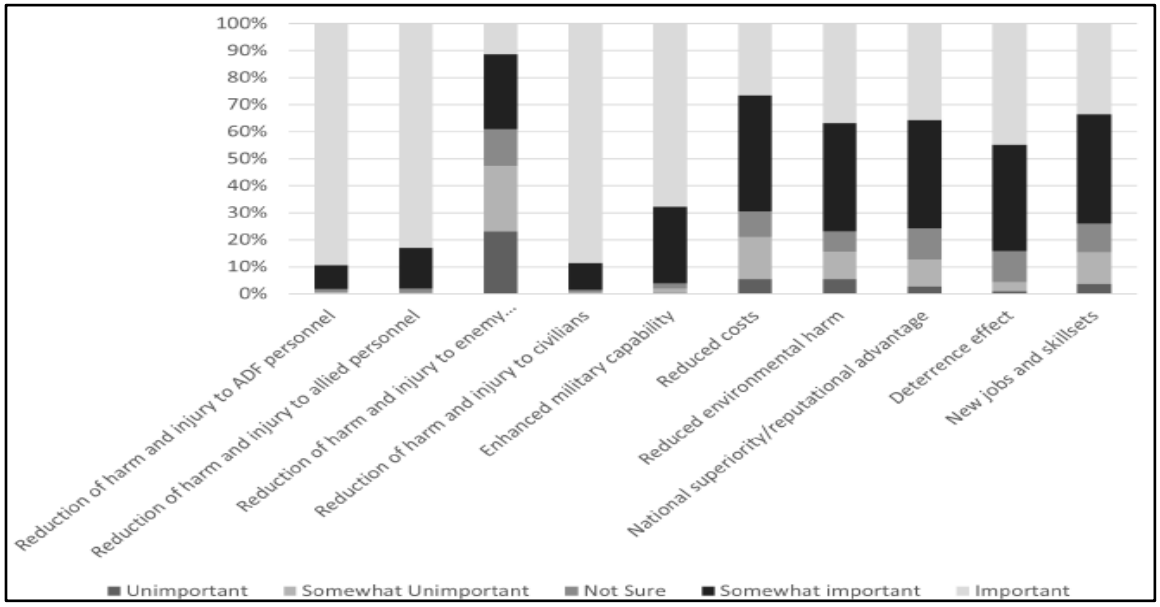

Fig. 5. Importance of perceived benefits of AWS

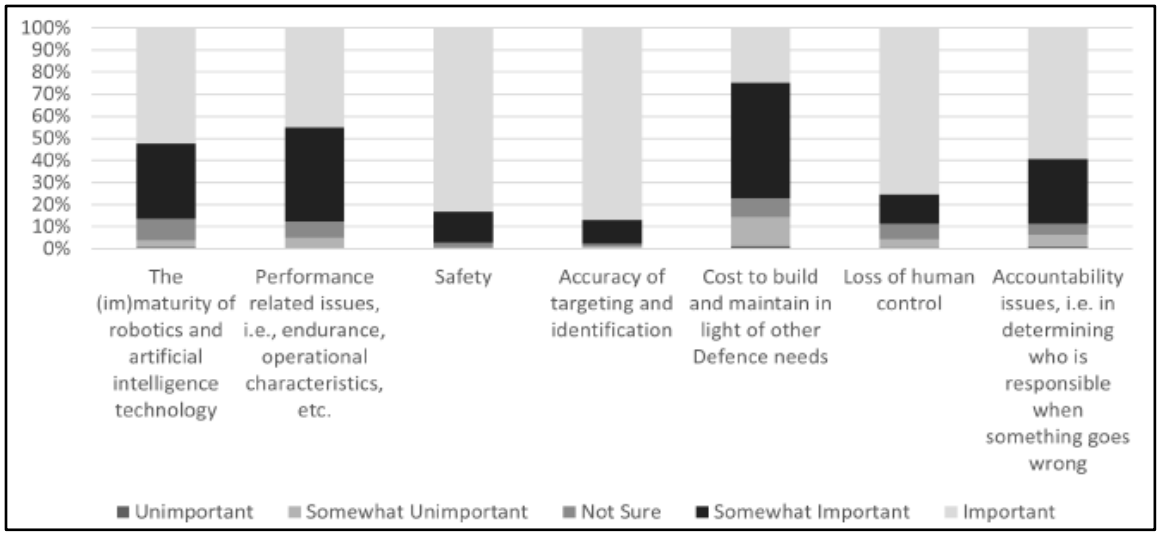

Fig. 6A. Importance of perceived risks of AWS

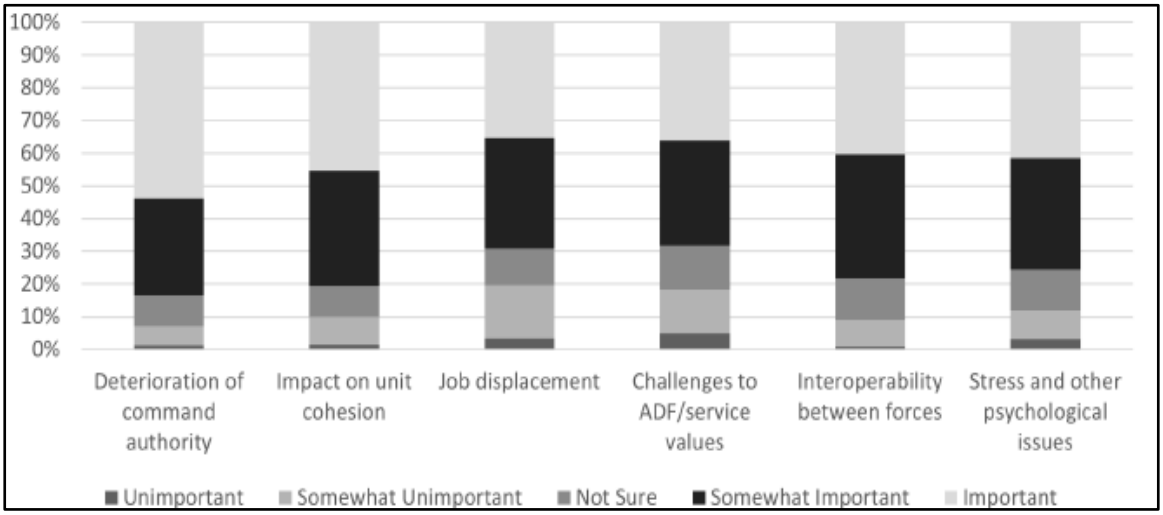

Fig. 6B. Importance of perceived risks of AWS

- *Hacking*, inquinamento dei dati ed attacchi avversari

  Un altro rischio legato all'impiego di sistemi IA risulta inestricabilmente correlato alle vulnerabilità degli stessi ad attacchi nel dominio *cyber* (*hacking*) tramite l'inoculazione di segmenti di codice vulnerabile, volti a renderlo sensibile a successivi attacchi (es. installazione di una *back door*). Inoltre, un sistema basato sull'apprendimento - una caratteristica questa preminente dell'IA - è esposto ad azioni o attività critiche (*data poisoning)* che comportano l'invio di informazioni contradditorie all'algoritmo. In aggiunta, laddove l'avversario sia in grado di individuare degli schemi di comportamento del sistema autonomo, quest'ultimo potrà tentare di ingannarlo o confonderlo fornendo allo stesso dati opportunamente manipolati e riconfigurati per indurlo in errore (es. falso segnale GPS), determinando così una condotta imprevedibile. In aggiunta a quanto precede, i sistemi autonomi, soprattutto laddove essi operino in sciami, risultano particolarmente sensibili agli strumenti di guerra elettronica avversari (*jamming* e *spoofing*) capaci di comprometterne inesorabilmente le funzionalità. In futuro le vulnerabilità analizzate potrebbero essere in parte mitigate dalla stessa IA, ma ad oggi appare necessario adottare misure di contrasto volte a minimizzare i rischi derivanti da potenziali azioni malevole intraprese da attori ostili.
- Incidenti e rischi emergenti

  La possibilità che si verifichino inconvenienti o incidenti fortuiti non può essere completamente evitata. Più i sistemi diventano complessi e interdipendenti, più la possibilità che l'incidente si verifichi aumenta inesorabilmente e drasticamente. In tale ambito, l'incremento della complessità e dell'interconnessione derivante dall'impiego dell'IA comporterà necessariamente maggiori possibilità di incidente, nonché ulteriori rischi causati dalla minore capacità di monitoraggio e supervisione dei sistemi da parte dell'operatore umano.
- Strategici

  I rischi strategici possono rappresentare una sfida significativa a livello globale. Il grande potenziale (ancora solo parzialmente espresso) ed i relativi benefici dell'IA nell'ambito militare hanno spinto i principali attori geopolitici mondiali a stanziare importanti risorse economiche per gli investimenti in tale settore, alimentando ulteriormente l'instabilità del quadro internazionale. Il dispiegamento di sistemi autonomi nel corso di conflitti armati potrebbe accrescere notevolmente il pericolo di *escalation* attraverso vari meccanismi. La guerra informatica potrebbe vedere sviluppi simili: i *software* IA potrebbero non solo costituire un valido strumento difensivo contro azioni malevoli condotte nel dominio *cyber*, ma sarebbero anche in

grado di lanciare un attacco di rappresaglia automatizzato contro le reti informatiche avversarie. Inoltre, se i sistemi avanzati di intelligenza artificiale dovessero svilupparsi fino al punto da poter addirittura prevedere le tattiche o il dispiegamento delle forze nemiche, questo comporterebbe conseguenze inevitabili e altamente destabilizzanti, soprattutto nel campo della dissuasione nucleare. Se l'IA fosse in grado di prevedere la posizione delle forze avversarie, un aggressore potrebbe essere incoraggiato a lanciare un attacco preventivo con l'intenzione di distruggerle senza temere rappresaglie. La potenziale vulnerabilità dei sistemi difensivi potrebbe, quindi, spingere a intraprendere un'azione militare ancor prima di subirla da parte dell'avversario. Il dilemma irrompe inevitabilmente: meglio usare subito le proprie unità per sferrare un attacco o perdere anticipatamente non avendo il tempo necessario per poter dispiegare le proprie forze prima che queste siano annientate dall'avversario?

Data la velocità di risposta delle tecnologie IA, per gli operatori diventa sempre più difficile monitorare i sistemi autonomi e prevederne eventuali malfunzionamenti anche a causa della limitazione dei *test* effettuabili prima del loro impiego in virtù della necessità di segretezza. Un esempio potrebbe essere quello legato all'uso di sciami di UAS, nei quali l'errore compiuto da una singola unità viene condiviso con altre, dando vita a un procedimento che non può più essere controllato o fermato.

- Uso della forza su vasta scala

  Tra i rischi strategici individuati possiamo annoverare l'uso della forza su vasta scala. Infatti, anziché perseguire opzioni non militari per risolvere i conflitti, abbassando la soglia per l'uso della forza, i *leader* politici potrebbero ricorrere all'uso di sistemi militari autonomi in un conflitto armato poiché il loro dispiegamento comporterebbe un minor rischio di danni al proprio personale. In tale ottica, occorre comunque tener presente che, sebbene meno soldati potrebbero trovarsi in zona di guerra con una minore esposizione a situazioni di estremo pericolo, le popolazioni civili, concentrate nelle zone delle operazioni militari, rimangono comunque esposte ai rischi connessi al conflitto.

- Gestione dell'*escalation*

  In aggiunta a quanto già argomentato, bisogna considerare la possibilità che un impiego sempre più significativo e massiccio di sistemi d'arma autonomi può comportare l'innesco di conflitti tra due attori oltre la volontà degli stessi. In tale contesto lo spazio per il negoziato in ambito diplomatico diminuirebbe, anche in virtù dei tempi di reazione dei sistemi dotati di IA mentre il rischio di *escalation* aumenterebbe in maniera significativa.

Se la guerra diventasse più probabile a causa dell'impiego di tecnologie IA, l'incremento della frequenza delle azioni militari stesse comporterebbe inevitabili rischi di *escalation*. La velocità con la quale l'azione militare potrebbe essere eseguita (a velocità di macchina, invece che entro i tempi umani) produrrebbe il rischio di una "guerra lampo" involontaria capace di svilupparsi improvvisamente con conseguenze imprevedibili e con danni potenzialmente incalcolabili. La combinazione di tecnologie militari emergenti e dirompenti potrebbe, poi, alimentare ulteriormente i rischi di *escalation*.

- Proliferazione

  Gli sviluppi della tecnologia informatica hanno consentito il raggiungimento di capacità offensive da parte di attori statali minori (ma anche organizzazioni criminali e terroristiche) difficilmente tracciabili e attivi al di sotto della soglia del conflitto, determinando il riconoscimento di un dominio operativo completamente nuovo (*Cyber*). Questo sviluppo tecnologico-capacitivo sembra seguire lo stesso schema anche nel campo IA a causa della sua connotazione *dual use*, innescando una proliferazione incontrollata di tale tecnologia, alimentata tra l'altro dalla volontà di evitare svantaggi competitivi nei confronti dei potenziali avversari. Il risultato, tra le altre cose, si tradurrebbe in un aumento esponenziale della velocità operativa con un conseguente incremento di instabilità e, dunque, di rischio di incidenti. Per quanto precede, sarà necessario adottare un quadro regolatorio pienamente condiviso a livello internazionale al fine di garantire e implementare le necessarie misure di prevenzione.

- Stabilità strategica

  Secondo un rapporto della RAND *Corporation*[25], un ulteriore rischio connaturato all'impiego di sistemi d'arma autonomi potrebbe essere rappresentato dall'indebolimento della deterrenza strategica, garantita dal principio di mutua distruzione assicurata. In particolare, l'impiego di sistemi basati sull'IA potrebbe permettere ad un attore di individuare tutti i vettori di ordigni nucleari dell'avversario, consentendo allo stesso di eliminare dal suo calcolo strategico la *second strike option* della controparte. Questo potrebbe dunque produrre un sistema pericolosamente instabile a livello globale. Infatti, la valutazione sulla convenienza di un primo attacco spingerebbe il soggetto "a rischio" ad impiegare il suo arsenale strategico per primo a causa della percepita minaccia nei confronti dei vettori "*second strike*".

[25] "*How Might Artificial Intelligence Affect the Risk of Nuclear War?"*, Edward Geist and Andrew J. Lohn, RAND Corporation, 2018.

CAPITOLO III: REGOLE D'INGAGGIO

1. Criteri per la mitigazione dei rischi

È opportuno identificare dei criteri di base attraverso i quali massimizzare i risultati nel processo di mitigazione dei rischi. Tali parametri esaminano le maggiori aree di criticità da considerare in sede di progettazione dei sistemi e di prefigurazione del loro impiego operativo:

a) Vincoli sul time frame

La Direttiva Statunitense DoD 3000.09 evidenzia che le armi autonome dovrebbero completare i compiti assegnati entro un periodo di tempo coerente all'intento del Comandante che le impiega, interrompendo l'ingaggio o richiedendo un ulteriore *input* da parte dell'operatore laddove non risultassero in grado di farlo. Emerge, dunque, la necessità di imporre vincoli in merito al tempo entro cui i sistemi autonomi possono operare in assenza di *input* dell'operatore. Tuttavia, i dettagli di questo vincolo dipenderanno dalle caratteristiche del sistema, dall'intento del Comandante, dall'ambiente operativo e da ulteriori fattori contestuali.

b) Vincoli geografici

Alcuni sistemi autonomi potrebbero avere ampie capacità di navigazione spaziale. Un vincolo geografico atto a limitarne il movimento potrebbe contribuire a mitigare i rischi connessi alle specificità ambientali di taluni teatri operativi, quali ad esempio la presenza di aree urbanizzate dove si registra un'alta commistione tra gli obiettivi legittimi e le proprietà civili. L'imposizione di tali criteri implica la necessità di definire e stabilire i parametri geografici entro i quali il sistema può operare.

c) Vincoli sui tipi di task

La maggior parte dei sistemi autonomi attualmente in uso tende ad avere solo un numero di funzioni limitato, ma non si può escludere che successive evoluzioni tecnologiche possano condurre allo sviluppo di sistemi in grado di svolgere una vasta gamma di compiti. Un vincolo sulle tipologie di *tasks* eseguibili offrirebbe l'opportunità di circoscriverne l'utilizzo in attuazione dell'intento del Comandante, per imporre maggiori o minori controlli in funzione del livello di uso della forza richiesto dal compito assegnato ad ognuno di essi. Ad esempio, qualora il sistema avesse le capacità di ingaggiare singoli individui anziché esclusivamente mezzi militari, questo richiederebbe vincoli più stringenti nonché rigidi *test* di verifica.

d) Affidabilità e prevedibilità

Al fine di garantire che il sistema operi in conformità con le specificità progettuali è necessario stabilire *standard* di affidabilità e prevedibilità, che devono essere specificati fin dalla fase di concettualizzazione e il loro soddisfacimento dovrebbe costituire condizione essenziale per il raggiungimento della piena capacità operativa dei sistemi stessi. L'aderenza a tali *standard* dovrà, in seguito, essere costantemente monitorata sulla base della valutazione dei dati raccolti. In ultima analisi risulta chiaro che ad una maggiore importanza/incidenza del sistema dovranno corrispondere altrettanti livelli di affidabilità e prevedibilità.

e) Informazioni accessibili

Un ulteriore criterio riguarda il grado di trasparenza richiesto al sistema, che si traduce di fatto nella facilità d'interpretazione dell'interfaccia e nella preparazione dell'operatore. Al fine di mitigare alcuni rischi connessi all'IA, l'operatore deve avere piena conoscenza di ciò che il sistema è destinato a compiere in uno specifico contesto operativo e deve essere in grado di determinare, entro un lasso di tempo ragionevole, come il sistema sia giunto a decisioni critiche o perché abbia intrapreso talune azioni. Secondo un *white paper* presentato dalla delegazione degli Stati Uniti nel 2017 all'UN CCW GGE (*Certain Conventional Weapons – Group of Governmental Experts*), l'interfaccia uomo/macchina dovrebbe: essere facilmente comprensibile per gli operatori addestrati; fornire un *feedback* (registrato) sullo stato del sistema; fornire procedure chiare per gli operatori per attivare e disattivare le funzioni del sistema.

Allo stesso modo, la sintesi del presidente della riunione dell'UN CCW GGE di aprile 2018 afferma che l'utilizzatore dovrebbe "conoscere le caratteristiche del sistema d'arma, essere certo che esso sia appropriato all'ambiente in cui è impiegato e avere informazioni sufficienti e affidabili su di esso al fine di prendere decisioni consapevoli assicurando il rispetto della legalità". Tuttavia, progettare e costruire sistemi di IA con una trasparenza sufficiente rappresenta una sfida molto complessa.

f) Opzioni di intervento

Un ultimo criterio per la progettazione dei sistemi autonomi riguarda la necessità di garantire un certo margine d'intervento all'operatore in modo che egli possa disimpegnarli o riorientarli tempestivamente, mitigando in tal modo i potenziali rischi derivanti da incidenti o altre situazioni non prevedibili. La rapidità con cui tali opzioni di intervento devono essere esercitate dipenderà anche dal contesto. Tuttavia, gli operatori dovrebbero essere in grado di intervenire su sistemi che impiegano la forza letale nel modo più rapido possibile, soprattutto in contesti ove si registra un'alta concentrazione di civili; diversamente,

l'intervento su sistemi che non impiegano la forza (ovvero su quelli operanti in aree lontane dai non combattenti) può anche essere meno reattivo.

2. Impiego dei sistemi autonomi e Regole d'Ingaggio

Le *Rules of Engagement* (ROE), indipendentemente dalla loro forma, rappresentano l'insieme delle direttive tese a specificare le circostanze e i limiti relativi all'impiego della forza. Esse sono generalmente intese come "un mix di requisiti dettati dal livello politico-militare nazionale declinati in osservanza del quadro normativo di riferimento sia nazionale che internazionale"[26]. Come tali, impongono vincoli operativi, legali e politici alla struttura militare[27].

Le ROE fanno parte di un quadro normativo più ampio relativo al dispiegamento delle Forze Armate e all'impiego della forza, interagendo con le disposizioni connesse al ciclo di *targeting* nonché con la formazione delle Tattiche Tecniche e Procedure (TTP). In tale ambito vengono incluse restrizioni relative agli obiettivi ingaggiabili ovvero misure tese a minimizzare i danni collaterali.

Le ROE possono rappresentare lo strumento appropriato per perimetrare l'impiego dei sistemi autonomi basati sull'impiego dell'IA, delineandone le condizioni di utilizzo relativamente ad uno specifico contesto. Esse possono rappresentare la base attorno alla quale stabilire i parametri per le diverse applicazioni militari dell'IA, traducendo così in istruzioni concrete le considerazioni e le limitazioni politiche, militari, legali ed etiche dettate dai livelli gerarchicamente sovraordinati. Possono rappresentare, dunque, il quadro d'azione da imprimere nella programmazione del sistema di IA. Ad esempio, le ROE potrebbero delimitare l'ambito geografico d'azione o limitare un certo numero di compiti assolvibili dal sistema nonché limiti temporali all'ingaggio. Inoltre, potrebbero disporre autorizzazioni preventive (o divieti assoluti) ad impegnare specifici obiettivi[28] e, allo stesso modo, potrebbero prevedere che il sistema debba segnalare eventi o problemi imprevisti. In un siffatto contesto, l'IA sarà in grado di scegliere quale ROE applicare in base all'ambiente operativo e alla specifica missione assegnata. Le stesse possono anche definire l'interazione tra esseri umani e sistemi di IA per particolari tipologie di operazioni. Possono stabilire, ad esempio, come un Comandante o un operatore debbano monitorare e controllare il sistema durante l'impiego. Poiché la necessità di controllo umano può variare

---

[26] Ibidem.

[27] Le ROE contengono elementi fondamentali che includono istruzioni generali per il Comandante (con questioni politiche e legali generali pertinenti all'operazione); posizionamento delle forze; imbarco, sequestro, recupero, salvataggio; avvertenze prima dell'uso della forza; deviazioni; individuazione dei bersagli; regolamentazione dell'uso di armi specifiche; restrizioni e autorizzazioni per l'uso della forza per difendere civili/oggetti o attaccare obiettivi militari. J. F. R. Boddens Hosang, "*Rules of Engagement and the International Law of Military Operations*", Oxford University Press 2020.

[28] "*Uses of Lethal Autonomous Weapon Systems", Gérard de Boisboissel, International Conference on Military Technologies* (ICMT), 2015.

a seconda del compito specifico attribuito a un sistema di IA e del rispettivo contesto, le ROE possono definire il livello di autonomia attribuibile al sistema stesso[29].

Tali regole risultano particolarmente rilevanti quando il sistema autonomo viene utilizzato per l'ingaggio di persone e/o oggetti con particolare riferimento al ciclo di *targeting*. In particolare, in considerazione del fatto che l'IA non può elaborare, durante il ciclo decisionale, valutazioni etiche relative al contesto[30], il giudizio umano dovrebbe essere significativo nella fase precipua all'uso della forza letale. A tal fine, potrebbe essere istituito un codice di condotta per gli operatori di sistemi di IA operanti all'interno del ciclo di *targeting* o un catalogo di ROE riferito a tali sistemi. Le applicazioni militari esistenti relative al *targeting* sono il *software* di scoperta del bersaglio, come ad esempio il *Super aEgis II*[31], e i sistemi di ingaggio[32]. In tale ambito si evidenzia il missile anti-nave statunitense a lungo raggio (LRASM), al quale la US *Navy* attribuisce la capacità di selezionare ed attaccare autonomamente i bersagli, anche in assenza di comunicazioni. In tali condizioni un sistema d'arma dotato di spiccata autonomia può determinare notevoli vantaggi competitivi e strategici. Un altro significativo sviluppo nell'ambito dei sistemi autonomi è stato registrato dalle Forze Armate turche con l'impiego del drone *Kargu*-2 (Fig. 7), utilizzato nel marzo 2020 nel teatro operativo libico per l'ingaggio di obiettivi senza prevedere l'autorizzazione da parte dell'operatore. Il suo utilizzo rappresenta un significativo precedente riguardo all'uso di IA per il *targeting* caratterizzato da un limitato controllo umano.

In ultima analisi, l'integrazione delle ROE è particolarmente rilevante nel rapporto uomo-macchina per concretizzare il controllo umano significativo nel contesto del *targeting*, atteso che il personale militare coinvolto debba possedere una conoscenza approfondita della sostanza delle ROE, con particolare riguardo a quale sistema possa usare la forza e in quali situazioni e condizioni specifiche.

Fig. 7. Drone Kargu-2

[29] Le ROE, chiaramente, non possono contraddire alcuna legge/regolamentazione o politica di rango superiore.
[30] Alcuni sostengono che giudicare la proporzionalità di un attacco, ad esempio, richiederebbe più di un bilanciamento dei dati quantitativi. Ciò comporterebbe una valutazione qualitativa ed etica da parte di un essere umano.
[31] "*Mapping the Development in Autonomy in Weapon Systems",* Vincent Boulanin and Maaike Verbruggen.
[32] Ibidem.

3. L'emanazione degli ordini e l'IA

Il mezzo concreto per condurre le operazioni in aderenza a quanto pianificato sono gli ordini. La NATO definisce un ordine come "una comunicazione scritta, orale o tramite segnale, che trasmette istruzioni da un superiore a un subordinato"[33]. Nonostante esistano diverse definizioni di ordini, generalmente essi sono brevi e specifici. Gli ordini possono essere emessi verbalmente o in forma scritta e devono essere conformi alla legge e alle disposizioni impartite dai livelli gerarchici sovra ordinati. Un altro termine frequentemente utilizzato per definire gli ordini in ambito militare è quello di "comando", che è definito come "un ordine impartito da un Comandante, ovvero l'espressione della volontà del Comandante di attuare una determinata azione"[34].

Alla luce di quanto espresso, nel valutare l'interazione uomo-macchina si potrebbe sostenere che gli *input* impartiti da un operatore corrispondano a degli ordini e che il sistema autonomo – proprio per le caratteristiche intrinseche legate alla sua programmazione di base – eseguirà tali "ordini" in aderenza ai parametri riguardanti gli obiettivi e i vincoli della missione assegnata. La risposta della macchina, dunque, sarà fortemente legata a quanto acquisito dal sistema in fase di apprendimento.

Per quanto attiene la continua interazione tra l'IA e l'operatore, essa potrebbe subire importanti evoluzioni in futuro. Le Forze Armate statunitensi, ad esempio, hanno progettato un *software* che consente ai sistemi informatici di comprendere istruzioni verbali, eseguire compiti e fornire *report*. Questi scambi consentirebbero al AWS di chiedere chiarimenti, fornire aggiornamenti in maniera rapida e accrescere la *situational awareness* delle unità militari. Tali applicazioni potrebbero facilitare il lavoro del personale militare con l'IA e ridurre la curva di apprendimento degli operatori per quanto riguarda il controllo dell'IA. Per quanto attiene all'interazione tra sistemi di intelligenza artificiale, gli ordini non sono necessari poiché essi scambiano semplicemente informazioni come parte di una rete complessa di applicazioni digitali. Per quanto riguarda gli ordini al personale militare, sembra improbabile allo stato attuale che le Forze Armate accettino che i sistemi di intelligenza artificiale impartiscano istruzioni ai loro membri. Tuttavia, tali strutture complesse in futuro potrebbero formulare raccomandazioni per l'azione che fungano da *input* per la decisione umana a una velocità e complessità sempre maggiori. Il rischio in questo caso è che il personale militare potrebbe non mettere in discussione tali raccomandazioni, non avere il tempo di valutarle criticamente o semplicemente non essere in grado di capire come il sistema sia giunto a tali conclusioni.

---

33 AAP-06 (n 13) "NATO *Glossary of terms and definitions*".
34 Ibidem.

È anche possibile ipotizzare che militari, collocati ai livelli gerarchici inferiori, nel ricevere istruzioni tramite la tecnologia dell'informazione possano non essere in grado di determinare se un ordine sia stato impartito da un essere umano o generato dall'IA. In sintesi, è probabile che in futuro gli ordini formali diventino irrilevanti per il controllo dell'IA nelle operazioni militari. Tuttavia, i concetti tradizionali di ordini e comandi possono essere utili per analizzare, categorizzare e sviluppare future interazioni tra i sistemi IA e gli operatori umani. In questo contesto, la distinzione tradizionale tra l'approccio manageriale e l'*Auftragstaktik*[35], come sviluppata da Carl von Clausewitz, suggerisce che l'*input* umano all'IA, ovvero lo sviluppo, la programmazione e il controllo operativo dei sistemi, potrebbe essere categorizzato in base al livello di discrezionalità garantito riguardo all'esecuzione di un compito. Date le qualità dell'IA, è ragionevole presumere che tali strumenti saranno più efficaci quanto più si vedranno attribuiti elevati livelli di autonomia; tuttavia, ciò riporta alla questione fondamentale di quanta autonomia debba essere concessa ai sistemi di IA.

4. *Framing* concettuale in ruoli del sistema d'arma autonomo

Per affrontare i rischi sopra descritti, limitatamente alla fase operativa del sistema d'arma autonomo, si ritiene opportuno impostare un *framework* concettuale volto a individuare e implementare le regole d'ingaggio da assegnare alle diverse situazioni riscontrabili per la specifica missione. In tale contesto si ritiene di identificare quattro diversi "Ruoli"[36] attorno ai quali impostare le regole del sistema, individuando secondo questa struttura il processo di definizione delle stesse e, contestualmente, fornendo all'operatore uno strumento concettuale teso a facilitare la comprensione delle logiche alla base delle azioni del sistema.

a) Ruolo emergenza

In questo ruolo "quadro"[37], il sistema agisce immediatamente per mettere in atto procedure di emergenza in risposta a danni o malfunzionamenti. La logica del sistema in questa modalità potrà comprendere il rispetto di regole di navigazione terrestre, navale ed aerea volte ad evitare collisioni o danni a terzi, nonché procedure per l'impiego di sistemi di guerra elettronica difensivi (attivi e/o passivi) in funzione dello scenario d'impiego. Nell'ambito di questo ruolo l'utilizzo di IA basata sulla filosofia HIL risulta fortemente

[35] Dottrina basata sulla delega decisionale fino al più basso livello ordinativi per il raggiungimento di obbiettivi militari conformemente all'intento del Comandante sovraordinato.

[36] Per ruolo si intende una determinata funzione del sistema, caratterizzata da uno specifico *set* di regole, che lo porta a operare in conformità all'intento del Comandante.

[37] Ruolo sempre attivo, seppur in maniera latente, durante l'impiego del sistema in uno qualsiasi degli altri ruoli.

svantaggioso in termini di rapidità d'azione in virtù del fatto che tale approccio presuppone la piena comprensione della situazione da parte dell'operatore. In effetti, un sistema basato su HOL permetterebbe di sfruttare in pieno la rapidità di reazione dell'IA, abilitando la migliore soluzione applicabile all'emergenza contingente. Per portare un esempio, un sistema UAV potrebbe utilizzare i propri sensori per rilevare ed evitare collisioni con altri velivoli (fortuite o meno) in maniera rapida ed efficace pur continuando ad eseguire la missione assegnata.

b) Ruolo difesa

Trattasi di modalità dalle caratteristiche spiccatamente automatiche in considerazione della necessaria rapidità di risposta che il sistema potrebbe dover mettere in atto per reagire efficacemente alla minaccia. Questa caratteristica fa sì che una filosofia HIL risulti non efficacemente implementabile al ruolo in esame. In tale ambito, i tempi di reazione da garantire all'operatore (HOL) dovranno essere funzionali alla minaccia da fronteggiare rendendo poco realistica la possibilità d'intervento da parte dell'operatore[38]. In questo ruolo risulterà dunque auspicabile che il sistema effettui dei frequenti *built-in* (test automatici), fornendo alla stazione di controllo dei *feedback* in merito al corretto funzionamento.

c) Ruolo combattimento

Questo ruolo prevede l'impiego del sistema in funzione offensiva e dovrà intraprendere azioni con tempi di risposta/ingaggio correlati al tipo di sistema d'arma impiegato ed alle considerazioni operative applicabili allo scenario di riferimento. In tale ambito l'adozione di una filosofia HIL garantisce all'operatore la massima capacità di controllo sull'ingaggio pur dilatando i tempi di realizzazione dello stesso. Diversamente, l'impiego di una tecnologia HOL garantirebbe la massima efficacia dell'ingaggio ma potrebbe risultare difficilmente applicabile a scenari complessi (es. urbani). In aggiunta, sia l'impostazione HIL che HOL dovranno prevedere in questo ruolo un sistema di sicurezza volto a verificare l'esistenza del collegamento (bilaterale) con la stazione di controllo, fondamentale per garantire la capacità d'intervento da parte dell'operatore, prevedendo, in caso contrario, l'attivazione di procedure "*hold*" o "*return to base*". La scelta tra i due sistemi potrà essere basata sulla complessità dello scenario, tuttavia, in un ambiente elettromagnetico degradato anche un HOL potrebbe risultare non impiegabile in virtù dell'impossibilità nell'esercitare la funzione di controllo, favorendo così l'impiego di sistemi HOOL.

[38] Il controllo dell'operatore, seppur effettivo, potrebbe essere non attuabile in una situazione che preveda la reazione del sistema in tempi rapidissimi, come potrebbe avvenire ad esempio per il contrasto di minacce ipersoniche.

I ruoli difesa e combattimento potrebbero essere attivati simultaneamente (ma operando in maniera tra loro indipendente) su un sistema complesso che preveda l'impiego di misure sia offensive che difensive, selezionabili/escludibili dall'operatore al fine di attivare/inibire talune funzionalità.

d) Ruolo addestramento

Questo ruolo può essere progettato per l'utilizzo del sistema in addestramento, consentendo il suo impiego in condizioni similari alla realtà ma senza l'ingaggio a fuoco (a meno di autorizzazione dell'operatore). Nella condizione descritta l'addetto potrà valutare il funzionamento del sistema d'arma, soprattutto nelle prime fasi d'impiego, incrementando la fiducia nello stesso. Inoltre, nell'ambito dell'impiego reale può essere inteso come una "sicura" per sistemi d'arma operanti in ambito HOL e di fatto consentirebbe all'operatore di commutarli in modalità HIL conservando ogni funzionalità della macchina (nell'ambito della sua funzione offensiva/difensiva) ma garantendosi maggiore tempo per il controllo delle azioni da mettere in atto.

CAPITOLO IV: ESIGENZE DI ADATTAMENTO DEL QUADRO NORMATIVO IN CONSIDERAZIONE DEL PRINCIPIO DI ACCOUNTABILITY

L'impiego di LAWS potrebbe comportare, in caso di violazioni, inaccettabili vuoti di responsabilità. Proprio la questione della sua attribuzione, in caso di condotta illecita o di errori, è una delle sfide giuridiche e legali centrali dell'IA, non potendo accettare l'opzione dell'impiego di un'arma senza che venga prima definita una catena di responsabilità[39]. La natura dell'IA sembra poter scardinare il tradizionale impianto, che appare connotato dalla generale *rule of thumb* secondo cui l'imputazione di un danno ricade sul soggetto che ha il controllo su un determinato elemento e, pertanto, egli deve adoperarsi per mitigare e minimizzare i possibili danni derivanti dal loro impiego.

Una prima risposta potrebbe essere rintracciata nella dottrina della responsabilità di comando, secondo la quale colui che ne ha ordinato l'impiego dovrebbe rispondere delle violazioni commesse da una LAWS. Questa strada può effettivamente rappresentare quella più opportunamente percorribile: pare indiscutibile, infatti, che un Comandante debba avere piena cognizione delle forze a propria disposizione. Va ricordato, infatti, quanto il concetto di comando sia fondamentale nel diritto internazionale. Tutta la disciplina che riguarda le operazioni militari è incentrata attorno ad esso al fine di accertare e verificare, con la maggior certezza possibile, l'imputabilità delle azioni al relativo soggetto. Esempio chiave è l'art. 29 della Convenzione ONU sul diritto del mare (UNCLOS), secondo cui una nave da guerra per essere considerata tale e venire riconosciuta dal diritto internazionale, deve essere necessariamente posta sotto il comando di un Ufficiale di Marina al servizio di uno Stato e il suo equipaggio sottoposto alle regole della disciplina militare. Questo rende chiaro come lo sviluppo già in atto di navi da combattimento autonome, in mancanza di una disciplina legale di imputabilità certa agli Stati in caso di eventuali violazioni, fa sorgere in capo a tali unità navali una preoccupante presunzione di illiceità che le colloca al di fuori di ogni copertura legittima. Ai sensi della disciplina internazionale vigente, infatti, queste non sono classificabili come navi da guerra a causa dell'assenza del comando umano che rende pertanto impossibile attribuire la responsabilità delle azioni compiute a qualsivoglia entità statale di riferimento.

Un'ulteriore ipotesi suggerisce di volgere lo sguardo verso la "fonte" della condotta di un'IA, imponendo una responsabilità sul soggetto che ha programmato e sviluppato il *software*. Tuttavia, l'imputazione del comportamento di una LAWS al programmatore non

[39] "*The evitability of autonomous robot warfare, in International Review of the Red Cross*", N.E.Sharkey, 789.

impedirebbe l'utilizzo dell'imprevedibilità di tale sistema quale argomento liberatorio, quando esso è dotato di funzioni di apprendimento.

L'attribuzione della responsabilità, invece, al produttore non diminuirebbe il valore di tale argomento (in fondo, anche questi non sarebbe che un semplice anello di una lunga catena); implicherebbe, ulteriormente, in considerazione del grado e del valore dei beni giuridici a rischio, un disincentivo, con ogni probabilità insuperabile, alla produzione di LAWS.

La difficoltà di individuare un soggetto responsabile e la questione delle *many hands* ha indotto alcuni autori a propendere per forme collettive di imputazione. Il diritto internazionale, in particolare, conosce il modello della *Joint Criminal Enterprise* (JCE), di origine giurisprudenziale[40]; tale dottrina prevede, per sommi capi, che, in presenza di violazioni perpetrate da più individui nella cornice di un programma criminoso comune, tutti possano essere considerati *principal perpetrator*. Tale modello, tuttavia, non sembra applicabile al caso in oggetto, in cui alla maggior parte, se non a tutti i soggetti coinvolti si potrebbe al più ascrivere un dolo eventuale, insufficiente *de iure condito* per una imputazione dinanzi ad un tribunale internazionale. Tale considerazione ha condotto a guardare anche alla responsabilità dello Stato quale possibile soluzione, essendo in tale regime assente – ovvero meno rigoroso – l'elemento della *mens rea*[41]. Non sembra che la responsabilità statale possa integralmente sostituire quella individuale, poiché non potrebbe applicarsi, tra l'altro, alle violazioni commesse da LAWS impiegate da gruppi armati non-statali, i quali, statisticamente, rappresentano i soggetti più coinvolti nei conflitti dell'età contemporanea.

Una soluzione generale potrebbe essere rappresentata dall'adozione di un regime di responsabilità oggettiva in capo alle macchine stesse. Per quanto paradossale possa apparire, questo potrebbe trovare una base giuridica nell'ipotesi di uno *status* specifico per i sistemi di IA. Da un punto di vista giuridico, l'imposizione di una responsabilità sulle LAWS appare pertanto praticabile, ma occorre interrogarsi in merito alla sua applicabilità. La risposta, sotto questo aspetto, pare essere negativa: gli obblighi di diritto umanitario sono pensati esclusivamente in funzione antropocentrica, non in relazione agli armamenti, e nessuna sanzione ipotizzabile (spegnimento, smantellamento, risarcimento ad opera di fondi appositamente istituiti) sembra poter avere effetto alcuno sul piano della deterrenza o della ricerca di una *restorative justice*. È forte, inoltre, il rischio che tale regime possa fornire

[40] ICTY, *Prosecutor* v Tadic, *Appeals Chamber Judgment*, del 15 luglio 1999, para. 220.
[41] L'elemento soggettivo del reato.

agli operatori la possibilità di impiegare le LAWS quali "scudi" giuridici per mitigare le rispettive responsabilità.

Risulta necessario, a questo punto, effettuare le opportune valutazioni sulla necessità di adattamento del riferimento normativo in considerazione del principio di *accountability* in relazione ai differenti livelli di *liability* individuati (individuale, *command*, *corporate* e statale).

1. Il vincolo del "controllo umano significativo" per la definizione della responsabilità individuale

Gran parte degli studi condotti negli ultimi anni in tema di armi autonome focalizzano la propria attenzione sulle problematiche relative alla conformità di tali sistemi ai principi cardine del Diritto Internazionale Umanitario (DIU). Si ravvisa la necessità di trovare un fondamento nel DIU e stabilire se possono essere applicate le regole volte a garantire la copertura necessaria nei casi in cui l'arma autonoma, dotata di "capacità di discernimento" (*deep learning*), modifichi in modo imprevedibile il comportamento inizialmente programmato. In questo ambito non vi è una posizione univoca a livello internazionale; si possono individuare, infatti, due distinte nonché antitetiche correnti interpretative di cui talune entità statuali si sono fatte portavoce: la prima ritiene che l'ordinamento vigente non sia adeguato a disciplinare l'impiego delle armi autonome; il secondo, invece, sostiene che il *carnet* normativo attuale offre sufficiente spazio di adattamento.

Punto di partenza dell'esame è l'art. 35, par. 1 del I Protocollo addizionale alla Convenzione di Ginevra, il quale stabilisce che la scelta delle armi da impiegare in conflitto non possa essere illimitata e quindi a discrezione delle parti. Le stesse devono poter colpire solo obiettivi militari e la loro potenza deve essere proporzionata all'effettiva capacità di difesa e resistenza dell'avversario in campo. È necessario, dunque, tener conto di questi due aspetti: distinzione e proporzionalità. La maggiore criticità appare correlata proprio alla capacità di distinzione. Per essere considerata lecita, infatti, l'arma autonoma dovrebbe essere in grado di differenziare i singoli obiettivi e altresì essere munita di dispositivi *ad hoc* che ne blocchino l'azione in caso di errore. È evidente che si tratti di un obiettivo assai arduo da raggiungere, ma il cui perseguimento risulta necessario. Appare inequivocabile che l'impiego di armi autonome debba necessariamente rispettare i principi, ormai cristallizzati e codificati nell'articolo 36 del I Protocollo Addizionale alle Convenzioni di Ginevra del 1977, relativi alla conformità di qualsiasi tipologia di arma alla disciplina del diritto internazionale e, in maniera ancora più rigida, alla Clausola Martens[42] che impedisce di considerare lecito

[42] Fornisce una formulazione più ampia del principio di umanità.

tutto ciò che non è previsto dai trattati. La Martens, dunque, si configura come una sorta di "schermo giuridico" che, come sottolineato dalla Corte di Giustizia Internazionale in un parere del 1996, "opera come una vera e propria rete di sicurezza per l'umanità"[43].

Il dibattito è entrato nel vivo nel 2013, quando il relatore speciale delle Nazioni Unite sulle esecuzioni extragiudiziali, sommarie e arbitrarie, Christof Heyns[44] ha evidenziato i molteplici fattori di criticità in gioco. L'approccio allo studio per l'inquadramento normativo dell'utilizzo delle LAWS, secondo lo stesso, deve necessariamente essere multifattoriale, tenendo sì conto delle norme di diritto internazionale a presidio dell'ingaggio degli obiettivi (proporzionalità e distinzione), ma deve anche verificare l'opportunità o meno di rimuovere *tout court* la partecipazione degli operatori umani dal processo decisionale. Nella cornice della Conferenza del disarmo svoltasi a Ginevra e della Conferenza di riesame della Convenzione su certe armi convenzionali (UN-CCW - *Certain Conventional Weapons*), sono stati svolti una serie di lavori per la regolamentazione delle LAWS. I dibattitti che ne sono scaturiti hanno condotto all'individuazione del paradigma del *Controllo Umano Significativo* (CUS) che implica un significativo ruolo di controllo da parte dell'uomo sui sistemi di intelligenza artificiale da un triplice punto di vista: morale, legale e operativo. Secondo Heather Roff e Richard Moyes[45] il CUS non solo è rilevante durante la messa in servizio del sistema, ma deve essere integrato anche in tutte le fasi del ciclo di vita, dalla progettazione, allo sviluppo e all'addestramento, implicando l'assunzione di responsabilità sull'operato dei sistemi stessi. Posto che le conferenze non hanno, per natura, alcun potere decisionale, in seno alla quinta conferenza di riesame della CCW (anno 2016) è stato istituito un gruppo di esperti governativi (UN-GGE) per valutare le tecnologie emergenti nel campo delle LAWS e codificare i risultati raggiunti nelle trattative tra i singoli Stati. Il gruppo, riunitosi a partire dal 2017, si è aggiunto alle Conferenze già organizzate dal Comitato Internazionale Croce Rossa (ICRC) sul tema. Purtroppo, i lavori del gruppo di esperti hanno constatato la contrapposizione dei due orientamenti in merito alla regolamentazione a cui le LAWS dovrebbero ubbidire (richiamare quella esistente o formularne una nuova, *ad hoc*), senza pervenire però ad una soluzione condivisa. Sulla scia di questi studi, nel 2018, al termine di un ulteriore incontro, gli esperti hanno elaborato 10 *Possible Guiding Principles* tra i quali i principi 2 e 3 che sottolineano la necessità di ricondurre l'operato della macchina alla responsabilità umana. A tal proposito, sono stati evidenziati due criteri: conservare, da un

[43] "Legality of the threat or use of nuclear weapons", Corte Internazionale di Giustizia, Advisory Opinion of 8 July 1996, par. 78.

[44] Rapporto sui sistemi d'arma autonomi.

[45] "Meaningful Human Control, Artificial Intelligence, and Autonomous Weapons", Briefing Paper for Delegates at the Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Heather Roff and Richard Moyes, Geneva, April 2016.

lato, la scelta dei sistemi d'arma da impiegare durante il conflitto nonché preservare, dall'altro, la sequenza comando-controllo senza soluzione di continuità.

Questa problematica va analizzata anche sotto il profilo, ancor più estremo, della repressione dei crimini di guerra da parte della Corte Penale Internazionale. Lo scoglio, più difficile da arginare, è l'impossibilità materiale di rinvenire nei LAWS l'elemento soggettivo del reato (*mens rea*), inteso quale caposaldo giuridico volto ad accertare l'evidenza di un delitto e ad assicurare la conseguente attribuzione di responsabilità penale. Alcuni principi cardine sono indicati nelle Conclusioni della sessione di lavori presso la UN CCW di agosto 2019, e in particolare quelli riguardanti il ruolo umano nell'uso della forza. Secondo tali vincoli: "*human responsibility for decisions on the use of lethal force must be retained*", l'operabilità di LAWS è possibile solo "*within a responsible chain of human command and control*" e l'interazione uomo-macchina deve garantire "*that the potential use of weapons systems […] is in compliance with applicable international law, in particular IHL* [*International Human Law*]". Quest'ultima codificazione, nello specifico, è stata indicata dai 320 partecipanti (rappresentanti di 63 Stati parte del CCW) al Forum di Berlino di aprile 2020 come potenziale «pietra angolare» per il prosieguo dei lavori in sede internazionale sul tema.

Il Dipartimento della Difesa USA offre una visione diversa, sintetizzata nel concetto di "giudizio umano appropriato". In particolare, la Direttiva 3000.09 "*Autonomy In Weapon Systems*" del gennaio scorso (riprendendo quanto già stabilito nell'analogo documento del 2012) stabilisce che "i sistemi d'arma autonomi[46] e semi-autonomi[47] devono essere progettati in maniera da garantire ai Comandanti e agli operatori di esercitare un appropriato livello di giudizio umano sull'impiego della forza". In merito, la delegazione degli Stati Uniti in seno alla CCW[48] del 2016[49] ha promosso questa formulazione, preferendola al concetto di CUS, poiché quest'ultimo risulta eccessivamente soggettivo, mentre il "giudizio umano appropriato" definisce meglio l'importanza della relazione uomo-macchina, dal suo sviluppo fino al dispiegamento del sistema. Anche questa prospettiva, tuttavia, risulta piuttosto

---

46 *"Autonomous weapon system: a weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system but can select and engage targets without further human input after activation"*; Department of Defense Directive 3000.09.

47 *"Semi-autonomous weapon system: a weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator. This includes: Semi-autonomous weapon systems that employ autonomy for engagement-related functions including, but not limited to, acquiring, tracking, and identifying potential targets; cueing potential targets to human operators; prioritizing selected targets; timing of when to fire; or providing terminal guidance to home in on selected targets, provided that human control is retained over the decision to select individual targets and specific target groups for engagement"*, Department of Defense Directive 3000.09.

48 *United Nations Convention on Certain Conventional Weapons - Group of Governmental Experts on Lethal Autonomous Systems* (UN CCW-GGE).

49 Intervento effettuato durante una riunione informale del consesso.

nebulosa in merito allo stesso significato di “appropriato”. Nell’agosto 2018 la delegazione degli Stati Uniti nell’ambito dello stesso consesso ONU si è spinta ulteriormente oltre nella critica al CUS, sostenendo che esso “rischia di oscurare le vere sfide nell’interazione uomo-macchina”. La delegazione ha argomentato che in tal senso l’analisi dell’interazione attraverso il paradigma del “giudizio umano appropriato” ha come scopo quello di “contribuire all’attuazione dell’intento dei Comandanti”, sostenendo che un incremento dell’autonomia applicata ai sistemi d’arma potrebbe raggiungere il risultato sperato, promuovendo maggiori tutele nell’ambito del DIU.

L’aspetto relativo all’interazione uomo-macchina, analizzato nelle sue varie declinazioni, riguarda innanzitutto il ruolo assunto dall’uomo nelle differenti fasi del ciclo di vita del sistema stesso. Durante la riunione dell’UN CCW-GGE, svoltasi ad agosto 2022, il Regno Unito ha proposto una distinzione volta ad inquadrare le differenti fasi di detto ciclo, evidenziando la necessità di interazioni e controllo umano in ciascuna di esse. Un concetto questo riassunto nel seguente schema:

Fig. 8. Mitigazione del rischio connessa all’impiego dei sistemi autonomi

Fonte: *Military applications on artificial Intelligence*, Rand Corporation, 2020.

Il grafico in fig. 8 evidenzia come la mitigazione del rischio connessa all’impiego dei sistemi autonomi inizi ben prima dello sviluppo del sistema stesso, partendo dalle decisioni di indirizzo politico nazionale (quale ad esempio la Direttiva 3000.09 del DoD USA), per poi seguirne tutto il ciclo di vita fino all’analisi post-impiego comprensiva di *feedback* sul sistema ed elaborazione delle lezioni identificate/apprese.

Nell'ambito di vari vertici NATO sono stati esaminati alcuni degli aspetti più rilevanti delle Tecnologie Emergenti e Dirompenti (EDT - *Emerging and Disruptive Technologies*), di cui fanno parte le armi autonome e quelle già esistenti. L'intento dei lavori è quello di stabilire alcuni principi, come ad esempio quello di trasparenza, che renda possibile un livello generale di visibilità sugli armamenti utilizzabili in campo, da parte di tutti gli attori. Le tecnologie impiegate, dunque, dovranno essere note non solo allo Stato che le usa ma anche agli avversari. L'idea di una condivisione del "bene comune" è stata avviata in *primis* da Bruxelles, la quale ha sostanzialmente ammonito che l'approccio finora concentrato solo sul concetto di autonomia e controllo, in realtà deve essere, a priori, sostenuto anche da un'accurata analisi del livello di perfezionamento tecnologico che l'IA permette di applicare all'industria militare. È innanzitutto necessario accertare quali siano i difetti intrinseci dell'IA ed esaminare, poi, le effettive possibilità di controllo umano che possono essere applicate alle macchine. Il rapporto dello *Science And Technology Committee* (STC) - *Sub Committee on Technology Trends and Security* (STCTTS) del 2019, ha rimarcato la necessità di affrontare le questioni etiche, legali e sociali connaturate alla salvaguardia della dimensione antropocentrica nell'impiego della Forza Armata.

Il Parlamento Europeo, con l'adozione della risoluzione 2021/C 456/04 del 20 gennaio 2021, ha sottolineato che l'IA, impiegata in un contesto militare, debba essere soggetta ad un "significativo controllo umano" in modo tale che l'operatore disponga di tutti i mezzi necessari per correggerla, bloccarla o disattivarla in caso di comportamento imprevisto, intervento accidentale, attacchi informatici o interferenze malevole di attori terzi. Tale provvedimento, poi, prevede che il rispetto del diritto internazionale, e in particolare di quello umanitario, che si applica inequivocabilmente a tutti i sistemi d'arma e ai relativi operatori, rappresenti un requisito fondamentale per gli Stati membri, i quali devono necessariamente rispettarlo con particolare riguardo alla protezione della popolazione civile o all'adozione di misure precauzionali in caso di attacchi (aggressioni militari e guerra informatica). Si ribadisce che un processo decisionale completamente autonomo non dovrebbe esonerare gli operatori dalla responsabilità che grava sulle persone che adottano la decisione finale, in modo da poter identificare l'essere umano responsabile della stessa. La direttiva sottolinea l'importanza della partecipazione dell'UE alla creazione di un quadro giuridico internazionale per l'uso dell'IA, facendosi promotrice, insieme all'ONU e alla comunità internazionale, dell'approvazione di un quadro globale che disciplini l'uso dell'IA per scopi militari e di altro tipo, garantendo che tale utilizzo rimanga entro i limiti rigorosi stabiliti dal DI e dal DIU. Tale quadro non deve mai violare o contravvenire alle prescrizioni della coscienza pubblica e dell'umanità, come stabilito nella già richiamata clausola Martens. Il

rapporto sottolinea che, nell'uso dei sistemi di IA nel settore della sicurezza e della difesa, la prevedibilità, l'affidabilità e la resilienza del sistema basato sull'IA, nonché la capacità dell'operatore umano di individuare possibili cambiamenti di circostanze e ambiente operativo e la sua possibilità di intervenire in un attacco o di interromperlo sono necessari per garantire che i principi del diritto internazionale umanitario (distinzione, proporzionalità e precauzione) siano pienamente applicati all'intera catena di comando e controllo. Il documento sottolinea ancora che i sistemi basati sull'IA devono consentire agli operatori di esercitare un controllo significativo, assumendo così la piena responsabilità dei sistemi e rispondendo di tutti i loro utilizzi. La risoluzione ritiene inoltre che le LAWS debbano essere utilizzate solo come ultima risorsa e siano lecite solo se soggette a un rigoroso controllo umano, con una persona in grado di assumere il comando in qualsiasi momento, in quanto un intervento e una supervisione umani significativi sono essenziali nell'adottare decisioni letali e gli esseri umani dovrebbero sempre essere responsabili quando decidono tra la vita e la morte. La risoluzione è dell'avviso che i sistemi totalmente privi del controllo umano (*human off the loop*) e della supervisione umana dovrebbero essere vietati senza eccezioni e in qualsiasi circostanza. La Risoluzione, tuttavia, non definisce come dovrebbe funzionare questo meccanismo di imputazione e attribuzione della responsabilità. Potrebbe essere estremamente complesso, se non addirittura impossibile, comparare una violazione commessa da un sistema di IA all'azione perpetrata da un operatore umano. La realizzazione di un sistema di IA richiede "il contributo di molteplici persone, organizzazioni, componenti meccanici, algoritmi, *software* e utenti umani in ambienti spesso complessi e problematici".

Da quanto sopra indicato deriva l'importanza e la necessità di stabilire un meccanismo di attribuzione chiaro ed equo. Un tentativo volto a definire il CUS potrebbe essere rinvenuto nell'AI *Act*[50] che stabilisce differenti gradi di rischio: i) inaccettabile; ii) rischio alto; iii) rischio basso o minimo. Ad ogni livello corrispondono, poi, una serie di requisiti e di obblighi, tra i quali spicca per rilevanza, in relazione ai sistemi ad alto rischio, il vincolo di sorveglianza umana ("*human oversight*"), prevedendo diversi tipi di intervento, tra cui l'azione dell'operatore sul funzionamento del sistema di IA ad alto rischio o l'interruzione del sistema mediante un pulsante di "arresto" o una procedura analoga. L'intervento umano viene dunque interpretato come misura volta a prevenire e a gestire il rischio derivante dalla manifestazione di potenziali effetti pregiudizievoli causati da un sistema IA. Allo stesso

[50] Commissione Europea, Proposta di regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, COM/2021/206, 21 aprile 2021.

tempo l'operatore umano sarà ritenuto uno dei soggetti responsabili qualora tale rischio si concretizzi ed egli risulti inadempiente rispetto ai propri obblighi di sorveglianza.

Rimanendo sempre sulla tematica della definizione del CUS, si è fatta strada, a livello internazionale, l'idea che ci possano essere, in base alle differenti tecnologie IA, vari livelli di controllo umano in grado di fornire adeguate garanzie circa il rispetto dei vincoli normativi ascrivibili al DIU e al DU. Un approccio differenziato al problema del CUS implica inevitabilmente la necessità di distinguere tra vari livelli di controllo condiviso uomo-macchina nello svolgimento dei compiti critici di selezione e ingaggio di un determinato obiettivo, assegnando così a ciascuna arma un livello appropriato, anche in base alla prospettiva normativa etica e giuridica. Sono stati proposti, in tal senso, cinque livelli di controllo condiviso che consentono di fornire un'idea generalizzata di quale sia il problema principale che una soluzione diversificata alla complessa questione del CUS deve necessariamente affrontare. Gli stessi sono classificati in relazione ai diversi gradi di controllo che, *step by step*, si spostano dall'essere umano (L1) alla macchina (L5):

- L1: la selezione dell'obiettivo da attaccare è integralmente effettuata dall'operatore umano;
- L2: la selezione dell'obiettivo da attaccare è effettuata dall'operatore umano in base a un ventaglio di opzioni suggerite dal sistema d'arma;
- L3: l'operatore umano si limita ad approvare o a rifiutare la scelta del sistema d'arma in merito all'obiettivo da attaccare;
- L4: l'operatore umano supervisiona la selezione dell'obiettivo effettuata dal sistema, mantenendo la possibilità di riprendere il controllo e annullare l'attacco;
- L5: l'operatore umano si limita ad attivare il sistema d'arma, definendone la missione nella fase preliminare di pianificazione, senza avere la possibilità di intervenire nella fase operativa.

A partire da questa classificazione si può innanzitutto osservare che L5 appare troppo sbilanciato a favore della macchina, almeno per quanto riguarda la ricerca dinamica di obiettivi (come illustrato dal caso di *Harpy* e di altre munizioni di tipo *loitering*). D'altra parte, L1 e L2, se considerati come i livelli più appropriati per l'esercizio del CUS, risultano eccessivamente inclini ad un approccio antropocentrico. Una soluzione alle numerose difficoltà poste dal CUS è rappresentata dalla necessità di modulare il controllo umano lungo i diversi gradi sulla base delle specificità tecnologiche dei sistemi d'arma e dei rispettivi contesti d'uso. Per garantire che gli operatori esercitino sempre un adeguato livello di CUS, non è necessario escludere l'autonomia dei sistemi d'arma nelle funzioni critiche di

selezione e attacco di un obiettivo, come propone una soluzione ad autonomia zero per qualsiasi sistema. D'altro canto, per mantenere il CUS non è sempre sufficiente prevedere l'esercizio del controllo umano solo nella fase di pianificazione dell'azione militare.

Una soluzione differenziata ma più prudente rispetto al problema del CUS potrebbe articolarsi in due punti principali. In primo luogo, in assenza di informazioni specifiche al contrario, si potrebbero imporre per *default* i livelli L1 o L2 di controllo umano. Potrebbero però essere ammesse eccezioni a questa regola - opportunamente argomentate e sottoscritte dalla comunità internazionale degli Stati - che offrono l'opportunità di andare oltre L2 a patto che il rispetto dei vincoli etici e giuridici sulla condotta delle azioni belliche non sia minacciato spostando in questo modo i privilegi di controllo verso la macchina (aspetto *differenziato* della soluzione).

Le principali motivazioni per questa scelta di *default* derivano dai limiti alla possibilità di prevedere e di controllare il comportamento dei sistemi d'arma basati sulle tecnologie della IA nonché dalle capacità stesse della macchina e dalle sue interazioni con l'ambiente operativo:

- la capacità di apprendere, fondamentale per sviluppare un sistema avanzato della IA e della robotica;
- la capacità di vagare (*loitering*) per un periodo esteso di tempo alla ricerca di un obiettivo da attaccare su un territorio che è soggetto a cambiamenti, anche repentini, delle condizioni osservate al momento di pianificare l'azione dell'arma autonoma;
- la capacità di coordinarsi con altri sistemi in assenza di un sistema centralizzato di controllo che l'operatore umano possa supervisionare (*swarming* o intelligenza da sciame);
- le interazioni e le conseguenti perturbazioni del comportamento atteso che derivano dai tentativi di *jamming*, hackeraggio e altri attacchi cibernetici;
- le complesse e veloci interazioni, di carattere competitivo o cooperativo, con altri sistemi artificiali, che avvengono su un campo di battaglia destinato a divenire sempre più informatizzato e robotizzato.

In definitiva, in assenza di prove convincenti che specifiche armi autonome non sollevino problemi previsionali e di controllo dovute a questi fattori, l'imposizione della regola *default* offre una protezione prudenziale dalle violazioni dei vincoli etici e giuridici che modellano i contenuti del CUS.

I lavori del CCW non hanno sinora registrato progressi significativi né in questa direzione specifica, né nella direzione più generale di un qualsiasi strumento multilaterale in merito al problema del CUS. Ciò, verosimilmente a causa del timore che una definizione troppo restrittiva dello stesso porti ad una proibizione indiscriminata di tutte le armi autonome, nonostante alcuni benefici attesi.

Malgrado gli strumenti indicati non abbiano una particolare forza vincolante, rappresentano senz'altro un indizio di quelli che saranno i prossimi passi, in particolare dell'Unione europea, in questa area di dibattito. Qualsiasi sforzo in questa direzione dovrebbe innanzitutto consentire l'identificazione del soggetto garante, investito di un effettivo potere per mezzo del quale impedire il verificarsi dell'evento dannoso. Questo è reso difficile, se non impossibile, dal già citato "*many hands problem*" e dalle responsabilità che ne conseguono. Peraltro, l'introduzione di una posizione di garanzia "onnicomprensiva", che comporti l'obbligo esteso per un non meglio identificato "sorvegliante umano" di prevenire ogni danno causato da un sistema di IA, sarebbe in palese contrasto con il principio di tassatività. Non si possono ignorare, infatti, le problematicità che i sistemi di IA sollevano per l'accertamento del nesso causale, dovute alla impossibilità per l'uomo di dominare completamente un certo evento, circostanza che in certi sistemi di IA potrebbe non essere applicabile poiché lo stesso, partendo da un insieme di *input*, potrebbe generare *output* considerati illogici e irrazionali per l'operatore. A volte, potremmo non riuscire a comprendere appieno quali *input* hanno avuto un ruolo determinante nell'ottenere *l'output*. In questi casi, dunque, "la distanza tra un'azione umana e le sue conseguenze [dannose] aumenta esponenzialmente". Uno degli snodi più problematici consiste nel livello di attenzione richiesto al potenziale supervisore. Si pensi, ad esempio, all'*automation complacency*, termine coniato in materia di incidenti aerei per far riferimento al fenomeno per cui l'automatizzazione di un qualsiasi compito porta il supervisore umano a confidare che la macchina se ne stia occupando in modo efficace e, di conseguenza, a smettere di prestare attenzione o all'*automation bias*, ossia la tendenza dell'essere umano a riporre fiducia eccessiva nelle raccomandazioni prodotte da un sistema informatico.

2. Responsabilità di comando

Indipendentemente dagli *standard* che potrebbero essere adottati in merito al coinvolgimento umano e alla supervisione sulle LAWS, vale la pena sottolineare che la legge vincola gli esseri umani e non le macchine. I Comandanti, infatti, sono tenuti a selezionare armi appropriate e lecite in base alle singole circostanze. I regolamenti annessi alla IV Convenzione dell'Aia del 1907 vietano l'uso di armi, proiettili o materiale volti a provocare

mali superflui. La legge consuetudinaria, invece, vieta metodi e mezzi di guerra indiscriminati. Tuttavia, la legalità della maggior parte delle armi si basa su come esse vengano impiegate dai combattenti sotto la direzione di un Comandante. Un fucile, ad esempio, è considerato un'arma lecita, ma i soldati possono utilizzarlo in modo illecito o possono ricevere l'ordine (illegittimo) di utilizzarlo in maniera impropria. I Comandanti possono ordinarne un uso improprio attraverso disposizioni deliberatamente illegittime. Il fattore chiave per poter parlare di responsabilità è che esista qualcuno in grado di rispondere per le azioni commesse nel corso delle ostilità: quella persona è proprio il Comandante[51].

Quest'ultimo, guidando le proprie unità, è responsabile della condotta delle forze soggette alla sua autorità, attraverso la quale autorizza l'impiego delle armi e contribuisce a definire regole di ingaggio appropriate per raggiungere l'obiettivo. L'obbedienza agli ordini è una pietra angolare della disciplina militare e sebbene i subordinati debbano rispettare solo quelli leciti, si presuppone che tutti lo siano a meno che ciò non venga confutato. I Comandanti sono responsabili delle violazioni del DI che commettono personalmente o che ordinano di compiere ai loro subordinati, come nel caso di crimini contro la pace (pianificazione e attuazione di guerre di aggressione), crimini di guerra (violazioni delle leggi o delle consuetudini di guerra) e crimini contro l'umanità (omicidio, sterminio, schiavitù). Gli stessi, anche nel caso di impiego di un sistema d'arma IA, hanno l'obbligo di comprendere come una LAWS può operare in un particolare ambiente. Devono avere piena consapevolezza delle funzionalità letali e delle peculiarità tecniche che contraddistinguono tali sistemi, tenendo ben in considerazioni tanto i vantaggi quanto le vulnerabilità ad essi collegati. Nel caso in cui una LAWS si dimostrasse indiscriminata, il Comandante sarebbe ritenuto responsabile. Ogni sistema d'arma nella zona di combattimento e ogni metodo di addestramento - tattiche, tecniche e procedure - rientra nella responsabilità diretta o individuale del Comandante.

Human Rights Watch, in molti dei suoi *report* pubblicati, sostiene che è ingiusto perseguire i Comandanti per l'azione di macchine sulle quali non hanno un controllo sufficiente. L'autonomia di queste armi crea una sorta di *vacuum of liability* che scagiona i Comandanti quando questi ultimi non possono esercitare piena supervisione sulle stesse. L'importanza strategico-militare e l'esposizione a responsabilità, talvolta considerate

[51] Thompson Chengeta, *Accountability gap: autonomous weapons system and modes of responsability in International law e policy*, v. 45 (2016), p. 32. Come evidenziato dall'autore, e anche in ossequio a quelle che sono le disposizioni contenute nello Statuto della Corte di Roma, affinché un Comandante possa essere ritenuto responsabile per le azioni condotte dai suoi subordinati devono essere soddisfatti tre elementi: il Comandante conosce i crimini compiuti dai suoi subordinati; il Comandante responsabile non ha impedito o evitato la perpetrazione della condotta illecita posta in essere dai suoi subordinati; il Comandante non ha punito il/i subordinato/i per la gravità degli atti.

"ingiuste", sono parte dell'*accountability* diretta o individuale del comando ed essa appare pertanto subordinata proprio a quel senso del dovere che i Comandanti hanno nei riguardi delle Forze Armate e della Nazione alle quali prestano il proprio servizio. Come sostiene Marco Sassòli[52]: "è giusto responsabilizzare il Comandante di un *robot* tanto quanto sarebbe giusto responsabilizzare un Comandante che ordina ad un pilota di bombardare un obiettivo da lui descritto come un quartier generale militare, ma che si rivela essere un asilo". L'*accountability* in questo senso è totale, anche se taluni osservatori esterni, come già visto, la ritengono "ingiusta". Nel ribadire che una LAWS non può essere considerata in alcun modo come un agente morale, appare interessante citare le considerazioni formulate da Peter Asaro, il quale sostiene che "la natura della responsabilità di comando non presuppone l'opportunità di tralasciare particolari imposizioni morali e legali tese a disciplinare e a perimetrare l'impiego della forza (proporzionalità e necessità)"[53]. Tale obbligo, continua Asaro, "può essere trasferito a un operatore, ma questo implica la necessità di supervisionare e monitorare la condotta di quel soggetto"[54]. Nel caso delle LAWS, conclude lo studioso, non è possibile delegare a tali sistemi questa peculiarità, poiché essi non si configurano quali agenti morali dotati di una propria dimensione etica.

3. La responsabilità aziendale

Neppure le aziende sono immuni dal rispetto del DI[55]. L'articolo IV della Convenzione per la prevenzione del crimine di genocidio del 1948 prevede pene severe per le persone che perpetrano questi eccidi di massa, "sia che rivestano la qualità di governanti costituzionalmente responsabili o che siano funzionari pubblici o privati". Ralph Steinhardt[56] sostiene che il termine "privati" possa includere anche le società, poiché non esiste alcuna indicazione che tale locuzione possa riferirsi esclusivamente agli uomini. Nonostante la responsabilità di queste ultime venga sancita dal diritto internazionale, date le caratteristiche non umane delle stesse, Ralph Steinhardt evidenzia che sorgono diverse questioni tecniche, anche molto complesse talvolta, che riguardano le aziende coinvolte nella progettazione e produzione di armi. Tuttavia, la responsabilità delle aziende non è universalmente accettata,

---

52 Marco Sassoli, *Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified*, 2014. Lo stesso Sassoli nota che un Comandante che ha potato per il dispiegamento di armi autonome è tenuto a comprendere il loro funzionamento, proprio come per qualsiasi altro mezzo o metodo di guerra. In questo caso, dunque, ci troviamo di fronte a un caso di chiara ed evidente responsabilità diretta, come nel caso di un soldato che spara un mortaio per colpire un carro armato, ma che ucciderà i civili che seguono quel t*ank*.

53 Peter Asaro, *On banning Autonomous weapon system human right, automation and the dehumanization of lethal decision making, International review of the Red Cross*, n. 886 (2012), p. 701.

54 Ibidem.

55 Ralph G. Steinhardt, *Weapons And the Human Rights Responsibilities of Multinational Corporations, in Weapons Under International Human Rights Law* 507, pp. 531-32 (Stuart Casey-Maslen ed., 2014) [hereinafter Steinhardt].

56 Ibidem.

poiché alcune giurisdizioni respingono il fatto che entità "senza anima da condannare e senza corpo da colpire" possano essere penalizzate in modo significativo per atti considerati illeciti. In questo caso, una delle sfide all'*accountability* è rappresentata dal fatto che in alcune giurisdizioni essa è soggetta a importanti limitazioni, mentre in altre viene completamente esclusa qualora il comportamento contestato riguardi sviluppi militari o sviluppi legati a funzioni pubbliche.

Una delle forme di rimedio disponibili per le vittime di LAWS è il risarcimento sotto forma di compensazione. Possono essere citate in giudizio le controparti responsabili, come coloro che hanno impiegato le LAWS o le persone coinvolte nello sviluppo di tali sistemi d'arma, quali produttori e programmatori. Tuttavia, la ricerca della responsabilità del produttore può comportare numerose difficoltà perché lo stesso potrebbe non essere direttamente collegato al danno subito dalla vittima. Le normative relative alla responsabilità del prodotto sono in gran parte inesplorate nella robotica. Ciò significa che per le vittime di LAWS avviare una causa sarà una sfida davvero difficile a meno che non sia chiaro che l'azienda violi apertamente principi giuridici considerati come universali. Sia in una causa civile che in una per responsabilità penale aziendale, la vittima deve essere in grado di dimostrare il nesso di causalità tra la condotta dell'azienda e il danno subito. Questo appare però difficile da dimostrare nel caso di armi completamente autonome. Questa sfida deve essere affrontata dalle giurisdizioni attraverso la formulazione e la codificazione di norme che regolino la responsabilità civile e penale delle aziende coinvolte nella progettazione, produzione e distribuzione di LAWS. Tuttavia, a causa della complessità tecnica delle LAWS e delle difficoltà giuridiche associate alla responsabilità delle aziende, la regolamentazione di tali sistemi d'arma rimane in gran parte incompleta e richiede ulteriori sviluppi. Tuttavia, a causa della complessità tecnica delle armi autonome e delle difficoltà giuridiche associate alla responsabilità delle aziende, la regolamentazione di tali sistemi d'arma rimane in gran parte incompleta e richiede ulteriori sviluppi. La vittima di solito assume l'onere di avviare una richiesta di risarcimento in una giurisdizione straniera con molteplici difficoltà economiche e peripezie burocratiche da affrontare nel corso di questo complesso e articolato *iter*. Christof Heyns ha messo in discussione che un tale approccio possa essere equo per la vittima[57].

Ci sono quattro ambiti di attribuzione della responsabilità delle imprese nel diritto internazionale: progettazione, produzione, vendita/trasferimento e, infine, impiego dell'arma.

[57] *"Report of the Special Rapporteur on Extrajudicial, or Arbitrary Executions, Summary*, Cristof Heyns (*Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*).

Nel primo caso, la responsabilità aziendale è piuttosto evidente quando le LAWS sono progettate per violare il diritto internazionale dei diritti umani (DU) e il diritto internazionale umanitario (DIU). Come sottolineato da Steinhardt, una delle sfide più complesse è rappresentata dal fatto che la maggior parte delle armi potrebbe non essere concepita specificamente per contravvenire ai principi ascrivibili al DIU e al DU. La progettazione di un'arma che viola il diritto internazionale, anche senza la sua effettiva messa in opera, può scatenare inesorabili sanzioni legali. Lo sviluppo di tale arma può essere vietato dal diritto internazionale e coloro che sono coinvolti nelle differenti fasi di progettazione possono essere considerati responsabili per il loro ruolo nella creazione della stessa. La responsabilità aziendale si applicherà chiaramente nel caso in cui le LAWS siano concepite per violare il diritto internazionale umanitario e dei diritti umani.

Tali armi potrebbero avere sufficienti utilizzi duali per renderle lecite nella fase di progettazione. La *mens rea* per una violazione è generalmente una condizione necessaria ma insufficiente per determinare la responsabilità in assenza di un *actus reus*[58]. L'argomento relativo all'impiego duale della tecnologia è stato evidenziato anche in relazione alle LAWS. Vari componenti di questi sistemi presentano utilizzi duali, rendendo così difficile, se non addirittura impossibile, imporre agli Stati l'obbligo di proibire completamente la progettazione di tali componenti.

La responsabilità aziendale è chiaramente delineata nella fase della produzione, durante la quale il produttore sceglie di fabbricare armi che sono illegali in sé, come stabilito dal diritto dei trattati che ne proibiscono la realizzazione. La stessa potrebbe anche essere illegale sulla base del diritto internazionale consuetudinario. Nel caso specifico delle LAWS invece questo appare complicato, perché non sono ancora vietate da alcun trattato e non esiste accordo alcuno che le bandisca ai sensi del diritto consuetudinario internazionale. Nel caso in cui vengano prodotte LAWS che non sono illecite in sé, ma che vengono utilizzate in modo illegale, il produttore non potrà essere mai chiamato in causa.

Ancora, nel caso della vendita e dei trasferimenti vietati, la società è responsabile per la violazione di vincoli internazionali. Tuttavia, è importante notare che, nella pratica, gli Stati possono incontrare delle difficoltà nel far rispettare tali obblighi, soprattutto se la società ha sede in una giurisdizione straniera con leggi e regolamentazioni diverse. Pertanto, vi è la necessità di una maggiore cooperazione e di un più ampio coordinamento internazionale tra le differenti entità statuali per affrontare efficacemente la questione della responsabilità delle società nella vendita e trasferimento di LAWS, tenendo in considerazione i diversi

[58] Fatto illecito.

ordinamenti interni. Quando un'azienda si comporta in modo contrario agli obblighi internazionali dello Stato, come nel caso di un embargo applicato a talune tipologie di armi, possono essere adottate varie forme sanzionatorie nei confronti di queste società produttrici.

Infine, la responsabilità delle imprese non dovrebbe essere confusa con quella del soggetto fisico che impiega o utilizza un'arma nel corso delle ostilità. Non esiste alcuna tipo di arma attualmente in uso in cui l'operatore, dopo aver commesso un crimine di guerra possa dire "non sono stato io, qualcosa è andato storto con la mia arma; chiedi al produttore". Alcuni studiosi[59] hanno messo in discussione se, in termini di Diritto Internazionale Umanitario, sviluppatori, ingegneri e ulteriori attori aziendali possano essere ritenuti responsabili dei crimini di guerra perpetrati da sistemi d'arma autonomi qualora questi ultimi abbiano svolto il rispettivo lavoro prima dell'inizio del conflitto armato. La risposta potrebbe essere positiva nel caso in cui gli stessi sapevano o avrebbero dovuto sapere che il sistema d'arma sarebbe stato utilizzato per commettere crimini di guerra. Affinché il progettista o il produttore possano essere perseguiti, quali autori diretti o co-autori, deve esserci un chiaro legame con il conflitto armato in questione e devono essere soddisfatti i requisiti legali di *mens rea* e *actus reus*. Se un produttore realizza e vende LAWS a un cliente che è parte di un conflitto armato o lo diventa successivamente, senza conoscere che le stesse verranno utilizzate per commettere crimini, non sarà soggetto ad alcuna accusa poiché la *mens rea* deve essere specifica per la particolare violazione contestata. Tuttavia, se le LAWS prodotte sono considerate illegali in sé per sé, il produttore potrebbe non essere perseguito per il crimine di guerra specifico per mancanza di *mens rea*, ma è comunque punibile ai sensi delle leggi penali interne. Anche la fornitura di materiale lecito può costituire un crimine di guerra qualora lo stesso venga fornito con la piena conoscenza che verrà utilizzato per scopi illeciti. In caso contrario, "fornire materiale che ha anche uno scopo del tutto legittimo non è un crimine".

4. La responsabilità dello Stato

Nell'ambito militare la questione della responsabilità statale - in relazione allo sviluppo e all'acquisizione di nuovi sistemi d'arma basati sull'IA - appare particolarmente rilevante, poiché gli Stati costituiscono le strutture primarie all'interno delle quali tali tecnologie vengono sviluppate, regolamentate e dispiegate. Mentre alcuni studiosi hanno sostenuto che la responsabilità statale non sia un modo efficace per affrontare le problematiche intrinseche delle LAWS, altri, viceversa, ritengono che tale *liability* sia preferibile a quella

[59] "*Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical ad Legal Issues to be Clarified*", Marco Sassoli, ed. 2014.

individuale. Tuttavia, tanto la responsabilità individuale, di comando ed aziendale, quanto quella statale operano in modo concomitante e complementare.

La responsabilità statale presenta talune caratteristiche specifiche proprio in relazione all'IA. Innanzitutto, lo Stato ha l'obbligo di riparare in caso di eventi avversi, far cessare la condotta illecita e offrire adeguate assicurazioni e garanzie di non ripetizione. Inoltre, tale *liability* può manifestarsi ancor prima dell'impiego dell'IA militare, poiché emerge anche nelle fasi di sviluppo o acquisizione di tali tecnologie. La responsabilità dello Stato non è, quindi, rilevante solo *ex post* per la condotta illecita posta in essere sul campo di battaglia, ma più in generale per la *governance* internazionale dell'IA durante tutto il suo ciclo di vita.

Quando l'impiego di sistemi di IA sul campo di battaglia comporta gravi violazioni degli obblighi internazionali, la responsabilità dello Stato può essere accertata qualora venga dimostrato che la condotta illecita in questione sia imputabile all'entità statuale medesima. La nozione di attribuzione della condotta rappresenta un caposaldo del diritto e consiste nell'attribuire allo Stato le azioni o le omissioni di persone fisiche o giuridiche che agiscono per suo conto.

Secondo *l'Articles on Responsability of States for Internationally Wrongful Acts* (ARSIWA) la responsabilità statale dipende inequivocabilmente dalle azioni o dalle omissioni dei singoli individui, intesi come soggettività appartenenti a un "gruppo di persone", "società o collettività", o "entità". In tema di responsabilità giuridica dello Stato, l'esistenza della condotta umana rappresenta, dunque, una precondizione essenziale. L'*accountability* stessa, in particolare, può essere chiaramente determinata qualora la condotta umana, imputabile allo Stato, abbia causato o contribuito ad innescare una violazione. Se l'operatore non esercita alcun tipo di controllo diretto sul sistema o se la macchina opera in larga misura autonomamente in modo tale da poter esonerare gli operatori umani da potenziali violazioni, si può sostenere che la condotta umana non può costituire una base solida per l'attribuzione della responsabilità. È dunque importante valutare se le violazioni del DI causate dall'impiego di sistemi di IA siano o meno riconducibili ad azioni e omissioni umane e, a loro volta, allo Stato. La questione di cosa costituisca una "condotta" umana non è esplicitamente definita in relazione al tema della responsabilità giuridica statale. L'identificazione della condotta umana risulta chiaramente correlata all'accertamento di una chiara ed evidente violazione determinata sulla base di specifici fatti criminosi. Nei dibattiti sulle LAWS, tale riflessione si traduce, come già analizzato precedentemente, nella convinzione che tali sistemi debbano rimanere sotto un "controllo umano significativo", poiché i comportamenti derivanti dall'uso di sistemi di IA potrebbero

non configurarsi come azioni od omissioni imputabili a un operatore umano[60]. Posto dunque che la responsabilità statale dipende dalla compresenza di iniziative o omissioni umane comportanti una violazione del DI, si possono identificare tre scenari in base ai quali è possibile o meno ascrivere l'*accountability* allo Stato. In primo luogo, quando tali tecnologie operano sotto il controllo diretto di un operatore, come nel caso in cui un sistema di riconoscimento di oggetti viene utilizzato come ausilio in un contesto in cui l'operatore ha anche una percezione visiva diretta. In tali casi le azioni e le omissioni dell'operatore concorrono direttamente al risultato; pertanto, la condotta in violazione del DI è imputabile allo Stato per conto del quale l'operatore agisce. In secondo luogo, quando sistemi di IA funzionano in modo completamente autonomo una volta attivati. È questo il caso di sistemi di difesa aerea operanti in modalità automatica, che possono, entro parametri limitati, identificare e neutralizzare le minacce in arrivo. In tale scenario, la condotta umana, che contribuisce più chiaramente al conseguente danno, è riconducibile a coloro i quali ne hanno deciso l'attivazione[61]. I comportamenti illeciti, verificatisi in seguito all'impiego di sistemi quasi completamente autonomi, possono essere imputati allo Stato sulla base delle iniziative e delle omissioni dei *decision-makers* militari o politici. Il terzo scenario riguarda i sistemi di IA che operano in una sorta di area grigia, sotto un certo grado di controllo e supervisione umana. In genere, un tale sistema è formalmente sotto il controllo diretto del suo operatore, il quale conserva quella capacità decisionale che gli consente di attenersi scrupolosamente o rifiutare categoricamente le raccomandazioni generate dall'IA[62]. In tali sistemi, un certo livello di discrezionalità viene conferito all'operatore, che valuta costantemente i dati forniti e la "condotta" dell'IA rispetto al proprio giudizio e al grado di consapevolezza della situazione. Tuttavia, in tali ambiti, esiste una linea molto sottile tra il supporto decisionale algoritmico e quello umano, poiché i sistemi di IA funzionano a una velocità tale che rendono difficile, se non addirittura impossibile, per gli operatori valutare realmente se sia necessario o meno seguire determinate raccomandazioni. Di conseguenza, il controllo sui risultati dei sistemi semi-autonomi può diventare superficiale e le azioni o omissioni del soggetto coinvolto potrebbero non costituire motivo sufficiente per

[60] Ad esempio, quando un sistema di intelligenza artificiale opera sotto il controllo formale di supervisione di un operatore umano, senza che l'operatore abbia alcuna capacità significativa di influenzare il risultato, si potrebbe sostenere che non vi è alcuna condotta umana da parte dell'operatore. Più le macchine operano indipendentemente, più è difficile sostenere in modo convincente che la condotta degli operatori umani costituisca motivo di attribuzione.

[61] Infatti, anche se l'operatore del sistema ha la possibilità di interrompere un attacco, ha un tempo limitato per intervenire e una consapevolezza situazionale molto limitata. Di conseguenza, l'influenza che la condotta umana sotto forma di funzioni di *over ride* ha sull'esito diventa priva di significato.

[62] Ad esempio, i sistemi di IA utilizzati a supporto dell'acquisizione del bersaglio possono raccogliere e analizzare dati da vari sensori e fonti e suggerire potenziali obiettivi, mentre l'operatore rimane in ultima analisi responsabile della decisione di lanciare o meno un attacco.

l'attribuzione di responsabilità. Anche in questo caso, dunque, la tesi più rilevante è quella di ricondurre la condotta agli organi statali che ne hanno decretato l'adozione e il dispiegamento[63].

La responsabilità dello Stato può essere accertata anche nelle precedenti fasi di progettazione, sviluppo e acquisizione. Prima della distribuzione dei sistemi e del relativo impiego, infatti, gli Stati possono essere chiamati a rispondere anche se sviluppano o acquisiscono tecnologie di IA in aperta violazione dei loro obblighi internazionali. In tale ambito, lo sviluppo o l'acquisto di tecnologie di IA può di per sé qualificarsi come un atto dello Stato ai sensi degli articoli 4, 5 e 8 ARSIWA[64]. Gli obblighi di diligenza applicabili nella fase di sviluppo dell'IA militare includono il dovere di osservare scrupolosamente i principi ascrivibili al DIU e implicano, inoltre, l'obbligo di adottare misure volte a salvaguardare i Diritti Umani (DU) all'interno della giurisdizione di uno Stato. In ambito militare, come già detto, l'articolo 36 del I Protocollo addizionale alle Convenzioni di Ginevra prevede un vincolo specifico già durante le fasi di sviluppo e acquisizione di una nuova arma, mezzo o metodo di guerra. L'applicabilità degli obblighi esistenti nella fase di progettazione delle tecnologie di IA implica che tali sistemi devono essere sviluppati e concepiti nel pieno rispetto delle norme internazionali. Se si scopre, dunque, che un sistema di IA non può essere tecnicamente conforme a determinati principi, allora il suo sviluppo dovrebbe essere interrotto o riformulato per garantire un coinvolgimento umano sufficiente volto ad assicurare il rispetto di tali vincoli. Continuare a sviluppare un sistema tecnologico quando è stato ampiamente dimostrato che non può rispettare determinate prescrizioni comporterebbe un vuoto di diligenza di cui lo Stato dovrà necessariamente rispondere. Allo stesso modo, nella fase di approvvigionamento da terzi, vi è l'obbligo di verificare che la tecnologia IA sia stata programmata in linea con i criteri già menzionati.

---

[63] In effetti, sono proprio i *decision maker* deputati a valutare le capacità, i limiti ed i rischi di un sistema e giudicare se, in quali circostanze operative e con quale grado di controllo umano, il sistema dovrebbe essere implementato.

[64] Art. 4: comportamenti di organi di uno Stato
1. Il comportamento di un organo dello Stato sarà considerato come un atto dello Stato ai sensi del diritto internazionale, sia che tale organo eserciti funzioni legislative, esecutive, giudiziarie o altre, qualsiasi posizione abbia nell'organizzazione dello Stato e quale che sia la sua natura come organo del governo centrale o di un'unità territoriale dello Stato.
2. Un organo comprende qualsiasi persona o ente che rivesta tale posizione secondo il diritto interno dello Stato.

Art. 5: comportamenti di persone o enti che esercitano prerogative dell'autorità di governo
Il comportamento di una persona o di un ente che non è un organo dello Stato ai sensi dell'articolo 4, ma che è abilitato dal diritto di quello Stato ad esercitare prerogative dell'attività di governo sarà considerato come un atto dello Stato ai sensi del diritto internazionale purché, nel caso in questione, la persona o l'ente abbiano agito in tale qualità.

Art. 8: comportamento sotto la direzione o il controllo di uno Stato.
Il comportamento di una persona o di un gruppo di persone sarà considerato un atto di uno Stato ai sensi del diritto internazionale se la persona o il gruppo di persone di fatto agiscono su istruzione, o sotto la direzione o il controllo di quello Stato nel porre in essere quel comportamento.

Accanto alla responsabilità per la condotta di uno Stato - sia nella fase di sviluppo che di dispiegamento - già ampiamente analizzata nelle sezioni precedenti, la complessa dimensione della *liability* può sorgere anche in relazione alla condotta di altri attori statali, para-statuali, *Rogue States* e privati.

In primo luogo, esistono disposizioni generali e specifiche tese a ostacolare potenziali violazioni del DI da parte di altri Stati. L'art. 16 dell'ARSIWA contempla, infatti, l'obbligo di non prestare alcun tipo di "assistenza consapevole" a un altro Stato nel compimento di un atto illecito. Analogamente, l'art. 1 delle Convenzioni di Ginevra[65] include, da un lato, l'obbligo di non fornire alcun tipo di assistenza ad altri Stati in violazione del DIU, mentre, dall'altro, stabilisce inequivocabilmente che ambo le parti rispettino i principi universali ascrivibili al DIU. Tali obblighi sono particolarmente rilevanti nel contesto di operazioni multinazionali, dove diversi Stati si impegnano congiuntamente e le tecnologie di IA potrebbero essere utilizzate da alcuni dei *partner* della coalizione. Uno Stato che non partecipa attivamente a un conflitto, potrebbe assumere la propria "responsabilità derivata" qualora trasferisse consapevolmente tecnologie di IA[66] ad un'altra entità statuale, la quale le impiega in aperta violazione del DI. A tal proposito, l'art. 6 del Trattato sul commercio di armi (ATT), prevede il divieto di autorizzare trasferimenti di armi, parti o componenti qualora lo Stato sia a conoscenza del fatto che questi potrebbero essere utilizzati per commettere crimini di guerra. Per le esportazioni[67], invece, l'art. 7 dell'ATT impone l'obbligo di valutare se le armi trasferite possano essere utilizzate per perpetrare gravi violazioni del DIU o dei DU.

Per quanto riguarda la condotta degli attori privati, invece, uno Stato può essere considerato responsabile, anche se in maniera indiretta, qualora non riuscisse a garantire che tali attori, all'interno della rispettiva giurisdizione, operino nel rispetto del DI. Nel campo dei DU, il Patto internazionale relativo ai diritti civili e politici (ICCPR) prevede esplicitamente l'obbligo di "prendere tutte le misure necessarie" volte a preservare il rispetto delle clausole del DU. Tale obbligo si configura quale vincolo di *due diligence*, richiedendo agli Stati di adottare misure tese a garantire che gli individui, nell'ambito della propria giurisdizione, non siano soggetti a violazioni dei DU e adottare misure ragionevoli per evitare il rischio di danni arrecati da parte di terzi. Il prefato obbligo è particolarmente rilevante nel contesto dell'IA militare, poiché non si può ignorare o sottovalutare il ruolo che le società private rivestono nello sviluppo e nella vendita di sistemi d'arma IA. È, inoltre, importante tener conto del fatto

[65] Art. 1: Le Alte Parti contraenti s'impegnano a rispettare e a far rispettare la presente Convenzione in ogni circostanza.
[66] Come ad esempio armi, *software* di acquisizione di bersagli o strumenti di sorveglianza.
[67] Non vietate in quanto tali ai sensi dell'art. 6 ATT.

che molte tecnologie di IA con potenziali applicazioni militari sono *dual use*, per cui, come già visto, anche le aziende produttrici dovranno essere soggette ai medesimi vincoli. Uno dei principali strumenti a disposizione degli Stati per garantire che gli attori privati rispettino i DU è rappresentato dalla regolamentazione interna. Quando si tratta di nuove tecnologie, quali ad esempio l'IA, gli Stati devono necessariamente adottare nuove normative per assicurare che i sistemi sviluppati da attori privati non comportino gravi violazioni dei principi connaturati al DU.

CONCLUSIONI

Nel presente elaborato è stato evidenziato il potenziale pervasivo e totalizzante dell'intelligenza artificiale in una vasta pluralità di applicazioni connesse all'ambito militare nonché le sue ambiguità e criticità, connaturate all'assenza di una completa sistematizzazione e concettualizzazione dottrinale in grado di dirimere ed appianare le innumerevoli preoccupazioni di carattere etico-morale e le controversie legali emerse attorno al loro impiego in scenari operativi non permissivi o alla loro possibile proliferazione indiscriminata. Appare tuttavia indubbio che il processo di *weaponization* delle piattaforme algoritmiche costituisca, soprattutto tra le fila delle Forze Armate, una sfida di indiscussa portata tattica e strategica, capace di rivoluzionare, in senso lato, le modalità relative alla conduzione delle operazioni militari in ambienti ostili o proibitivi, dove il dispiegamento di mezzi o unità risente di inesorabili limiti o rischi in termini di capacità operative e resilienza *boots on the ground*. L'IA, dunque, rappresenta uno strumento tecnologico utile a ridefinire l'architettura securitaria delle singole compagini statuali e a strutturare il nuovo concetto di *algorithmic warfare*, cui il dominio Difesa – anche alla luce dei recenti sviluppi che caratterizzano l'attuale scenario geopolitico internazionale – deve necessariamente confrontarsi per vincere le numerose sfide che si prospettano all'orizzonte e imporsi così, nell'intricata e frastagliata trama delle alleanze globali e multipolari, quale attore cruciale di assoluta portata e rilevanza strategica. L'intelligenza artificiale applicata agli *Autonomous Weapon Systems* – pur sollevando tra accademici e addetti ai lavori importanti timori e inquietudini di carattere giuridico connesse alla *compliance* legale con i principi cardine del Diritto Internazionale Umanitario – comporterà una sorta di *reshaping* dello spazio bellico tradizionale e delle tecniche operative. Gli AWS, infatti, consentiranno, da un lato, una minore esposizione del personale a sopraggiunte situazioni di pericolo in contesti operativi contraddistinti da un elevato grado di vulnerabilità, mentre, dall'altro, contribuiranno ad accrescere le capacità di intervento nelle attività operazionali multi-dominio (inclusi quello cibernetico e spaziale), identificando obiettivi potenzialmente pericolosi sulla base di schemi di classificazione per priorità di minaccia. Il perimetro difensivo, come visto, ha sostanzialmente perso – anche in virtù della crescita capillare di *competitors* para-statuali o asimmetrici – la sua tradizionale connotazione spaziale con inevitabili ripercussioni sul processo di selezione e identificazione della minaccia, il quale sta inesorabilmente acquisendo una rinnovata e più complessa connotazione, connessa al suo nuovo carattere ibrido, multidirezionale e sfuggevole. Ne consegue, dunque, che i sistemi autonomi a trazione IA – in virtù delle rispettive peculiarità tecnologiche e specificità componentistiche

che garantiscono alla piattaforma un differente grado di autonomia – rappresentano un innovativo strumento strategico, teso a minimizzare, se non addirittura ad abbattere, soprattutto in chiave difensiva, i tempi di reazione rispetto a pericoli o sopraggiunte minacce provenienti da potenziali entità ostili. In merito al concetto di arma autonoma, è stato evidenziato che, a livello internazionale, non esiste un'univoca definizione condivisa e la stessa per alcuni comprende i sistemi HIL (visione più ampia), per altri no (approccio più restrittivo). Inoltre, va sottolineato che la stessa suddivisione HIL, HOL, HOOL potrebbe essere superata in virtù del concetto di CUS, vincolo dal quale discenderebbe l'esclusione di tutti quei sistemi che non ne possono garantire il pieno rispetto. Va precisato infine che, mentre sistemi d'arma HOL sono già in uso da tempo, sia in ambito difensivo che offensivo, non risultano impieghi di LAWS secondo l'approccio HOOL.

Come evidenziato in più circostanze, l'applicazione dell'intelligenza artificiale per i sistemi di comando e controllo (C2), utilizzati per sostenere il comando nelle differenti fasi di monitoraggio delle operazioni, può contribuire a velocizzare il processo di raccolta, analisi, elaborazione e sistematizzazione del flusso di dati, indispensabili per la razionalizzazione del processo decisionale e acquisire, in tal modo, un vantaggio decisivo sulle *enemy course of action* in un dominio operativo caotico e caratterizzato da un elevato grado di imprevedibilità. La possibilità di usufruire di dati costantemente aggiornati, categorizzati e memorizzati dal sistema IA costituisce un elemento strategico determinante per i *decision-makers*, i quali potranno applicare appieno il concetto di *information superiority* e *information dominance*, connesso ad una gestione ottimale del flusso informativo e alle conoscenze acquisite. Tale supremazia appare pertanto fondamentale per l'applicazione dell'intelligenza artificiale al ciclo decisionale, garantendo in tal modo il «*planning* e *replanning* delle risorse proprie del sistema»[68] con il chiaro obiettivo di assicurarsi la superiorità strategica contro il nemico in qualsiasi dominio. Per ciò che riguarda, invece, le attività C4ISR (*Command*, *Control*, *Communication*, *Computers*, *Intelligence*, *Survaillance* e *Reconnaisance*) o ISR (*Intelligence*, *Surveillance* and *Reconnaissance*), talune unità impiegano già sistemi che sfruttano il potenziale derivante dal dispiegamento dell'intelligenza artificiale per svolgere compiti ritenuti *dull*, *dirty*, *dangerous or dear*. Un ulteriore scenario di particolare interesse per il dispiegamento dei sistemi IA è rappresentato dall'A2/AD (*Anti-access/Area-Denial*), il cui impiego da parte di talune entità statali è volto ad impedire la possibilità di proiettare presenza nemica in alcuni ambiti dei domini terrestre, marittimo e aereo (e in futuro

[68] Stato Maggiore dell'Esercito, "*L'impatto delle emerging and disruptive tecnologies* (EDTs) sulla difesa", ed. 2022, pp. 24-36.

spaziale). I sistemi IA, proprio in questa circostanza, possono contribuire a penetrare, grazie all'azione combinata con le piattaforme ISR, tali aree.

I differenti gradi di interazione e ibridazione *human-machine teaming*, analizzati in relazione ai molteplici livelli di autonomia conferiti agli AWS attualmente conosciuti, ci consentono di attuare alcune importanti considerazioni attorno alle capacità di intervento e controllo da parte dell'operatore, la cui estraniazione, sia essa parziale (HOL) o completa (HOOL), solleva importanti ambiguità etico-morali e problematiche legali nel contesto della tutela di normative cardine del Diritto Internazionale Umanitario e del rispetto del concetto di *accountability*. Risulta necessario evidenziare che un approccio HIL applicato agli AWS, pur garantendo la piena *compliance* al principio chiave di *controllo umano significativo*, non consente una reazione difensiva risoluta ed immediata a un'imminente minaccia. In questo caso, infatti, il tempo di reazione richiesto per l'ingaggio appare così ristretto che potrebbe essere impossibile per l'operatore rimanere *in the loop* e adottare, quindi, un'azione razionalmente consapevole attraverso la quale difendersi efficacemente in situazioni di estremo pericolo. Proprio i sistemi difensivi autonomi HOL, caratterizzati comunque dalla presenza di un soggetto capace di intervenire in caso di eventi avversi o potenziali malfunzionamenti, appaiono ampiamente accettati, poiché essi rappresentano armi stazionarie concepite soprattutto in chiave proattiva per annientare pericoli o minacce incombenti[69]. Tuttavia, l'impiego di tali tecnologie potrebbe comportare alcune problematiche connaturate al principio di *automation bias* che riflette, in sostanza, la progressiva estraniazione dell'operatore rispetto al piano d'azione presentato dal sistema, innescando così un processo teso a trasformare l'addetto stesso in un *automation operator*. Ancor più complesso, invece, appare l'utilizzo di AWS in funzione offensiva, poiché questo aspetto solleva rilevanti questioni legali e numerose controversie etico-morali strettamente interrelate alle fonti e ai principi universali ascrivibili al DIU. L'intero corpo del Diritto Internazionale Umanitario, infatti, è caratterizzato dalla relazione tra due elementi fondamentali e antitetici: la necessità, da un lato, di preservare l'incolumità della popolazione e l'integrità delle infrastrutture civili e il bisogno, dall'altro, di dispiegare armi volte a garantire la supremazia militare contro attori ostili. Tuttavia, l'introduzione di nuovi sistemi d'arma implica la valutazione di criteri generali di *compliance* al DIU al fine di verificare che tali dispositivi rispettino i principi di distinzione, proporzionalità, precauzione e necessità orientati a minimizzare il rischio di potenziali danni inferti ai civili e ad impedire che vengano condotte operazioni non proporzionate rispetto all'obiettivo militare da perseguire.

[69] "*Losing Humanity: The Case against Killer Robots", Human Rights Watch and International Human Rights Clinic* (n. 30), See Chesterman (n. 4) 230.

L'osservanza di questi criteri risulta, poi, strettamente connessa all'attribuzione della responsabilità in caso di eventi avversi. A livello internazionale si è sviluppato un confronto tra coloro che non ritengono necessaria una modifica dell'impianto normativo ed altri che, viceversa, considerano ciò inevitabile. In questo contesto, secondo il principio di *accountability*, si è affermato il già citato principio del CUS, volto a ribadire la centralità di un approccio antropocentrico nell'impiego di armi, anche nell'ambito di IA, in riferimento al quale è necessario individuare gli elementi utili a identificare il soggetto a cui attribuire la responsabilità nel caso di violazione del DIU e dei DU. La stessa definizione di CUS, sviluppata nell'ambito della CCW, non essendo ancora inserita nel quadro normativo internazionale non produce effetti vincolanti. Il Parlamento Europeo ha fatto proprio questo concetto nella citata risoluzione del 2021, dando una chiara linea di condotta da adottare a livello comunitario e auspicando che il principio venga proposto anche in ambito ONU con l'intento di addivenire ad un bando internazionale dei sistemi HOOL. Ad oggi, infatti, non si può garantire con sufficienti margini di sicurezza che tali sistemi possano rispettare appieno i principi del DIU. A tal proposito, la posizione italiana, in linea con questa visione, è stata espressa dall'ex Ministro della Difesa Guerini, il quale ha evidenziato l'importanza di "definire in modo chiaro e condiviso i limiti e le condizioni di autonomia di tali macchine" e aggiungendo che "sistemi autonomi già esistono, ma è evidente che in campo militare esistono significative implicazioni etiche e legali". Esiste pertanto la concreta necessità di "individuare un adeguato sistema giuridico entro il quale poter collocare la robotica autonoma", in accordo con quanto sancito dalla Commissione giuridica del Parlamento Europeo[70]. In conclusione, fermo restando che, ad oggi non esistono LAWS HOOL e che appare chiara la volontà di molti attori internazionali di procedere ad un loro bando, è necessario trovare una definizione condivisa di CUS e in grado di soddisfare le diverse posizioni degli attori nell'ambito della comunità internazionale.

Per quanto concerne le regole d'ingaggio volte a perimetrare l'utilizzo dei sistemi d'arma autonomi in ambito militare, si ritiene che debbano essere considerati i criteri riportati nel Capitolo III, con particolare riguardo ai vincoli sul *time frame*, affidabilità e prevedibilità, opzioni d'intervento, unitamente ai contenuti del catalogo nazionale delle ROE. La valutazione di tali criteri permetterà di utilizzare al meglio le peculiarità proprie degli AWS, sfruttando le opportunità che essi forniscono e, al contempo, temperando le criticità che li caratterizzano. In aggiunta, nel corso delle differenti fasi caratterizzanti il ciclo di *targeting*,

[70] Intervento del Ministro della Difesa Guerini alla "giornata di studio su intelligenza artificiale, sicurezza, responsabilità, etica", riportato su Analisi Difesa 2019. "Dai droni alle armi autonome. Lasciare l'apocalisse alle macchine?", Francesca Farruggia, Franco Angeli S.r.l. 2023.

si dovrà debitamente valutare la complessità del particolare scenario operativo a premessa dell'impiego di un AWS. Infatti, come ampiamente delineato, al crescere della complessità aumentano le criticità in ambito distinzione e proporzionalità attribuite agli AWS, rischiando di produrre danni collaterali indesiderati. Per concludere, si ritiene che l'implementazione di un *framework* concettuale[71] attraverso il quale integrare le regole d'ingaggio nel sistema autonomo potrebbe fornire all'operatore una maggiore confidenza nei confronti del comportamento dell'apparato. Questo consentirebbe all'addetto di operare il controllo sulla macchina con una comprensione del suo funzionamento più profonda, affrontando così il problema dell'*automation bias* per fornirgli maggiori margini di intervento pur conservando i vantaggi capacitivi forniti dalla stessa.

[71] Al precedente Cap. III è stata proposta una schematizzazione su 4 distinte funzioni/ruoli – Ruoli Emergenza, Combattimento, Difesa e Addestramento.

BIBLIOGRAFIA

Libri

- *Dai droni alle armi autonome, lasciare l'Apocalisse alle macchine?,* a cura di Francesca Farrugia, Franco Angeli, Milano, 2023.

Pubblicazioni e documenti

- "*Autonomy in weapon systems*", DoD *directive* 3000.09 *ed. Jan.* 2023.
- "*Artificial intelligence application in the military. The case of United States and China*", Gloria Shkurti Özdemir, SETA, 2019.
- "LAWS Intelligenza Artificiale e robotica alla Guerra", IRIAD REVIEW, ISSN2611-3953, n.5 – Maggio 2019.
- "Artificial Intelligence in military application – opportunities and challenges", István Szabadföldi, Land Forces Academy Review, Vol. XXVI, No. 2(102), 2021.
- "*Proposte per una strategia italiana per l'intelligenza artificiale*", Elaborata dal Gruppo di Esperti MISE sull'intelligenza artificiale, luglio 2019.
- "*None too clever? Military applications of artificial intelligence*", Drone Wars UK, Peace House, dicembre 2021.
- "*Military operations and artificial intelligence*", Tomas Vestner, *Geneva Centre for Security Policy,* 2021.
- *"How Might Artificial Intelligence Affect the Risk of Nuclear War?*", Edward Geist and Andrew J. Lohn, RAND *Corporation*, 2018.
- *"L'intelligenza non è artificiale", Limes, N.12 / 2022*
- "*Military applications of Artificial Intelligence, Ethical concerns in an uncertain world*", Forrest E. Morgan, Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, Derek Grossman, RAND *corporation,* 2020.
- "*Uses of Lethal Autonomous Weapon Systems*", Gérard de Boisboissel, *International Conference on Military Technologies* (ICMT), 2015.
- *"Losing Humanity: The Case Against Killer Robots",* HRW *and International Human Rights Clinic, 2012.*
- "*Mapping the Development in Autonomy in Weapon Systems*", Vincent Boulanin and Maaike Verbruggen.
- "*Rules of Engagement and the International Law of Military Operations",* J. F. R. Boddens Hosang, *Oxford University Press,* 2020.
- AAP-06 (n 13) "NATO *Glossary of terms and definitions*".
- "*Insights for the third offset: Addressing challenges of autonomy and artificial intelligence in military operations*", *CNA Analisys and Solutions*, Larry Lewis, 2017.

- "*Pros and cons of autonomous weapon systems*", *Military review*, Amitai e Oren Etzioni PHD, 2017.
- "*The Case for Ethical Autonomy in Unmanned Systems*", Ronald C. Arkin, *Journal of Military Ethics* 9, no. 4-2010.
- "*Winning the artificial intelligence era – Quantom diplomacy and the power of automation*. Cap. IV. *New warfare*: potenziali rischi e mitigazioni", Enrico Savio ed Enrico Comin.
- "*Risks and Benefits of Autonomous Weapon Systems - Perceptions among Future Australian Defence Force Officers*", *Journal of indo-pacific affairs,* Dr. Jai Galliott and Dr. Austin Wyatt, *2020.*
- "*Unmanned Systems Roadmap* 2007-2032" US DoD, James R. Clapper Jr..
- "Dai droni alle armi autonome. Lasciare l'apocalisse alle macchine?", Francesca Farruggia, Franco Angeli S.r.l. 2023.
- *"Accountability gap: autonomous weapon systems and modes of responsibility in International law", Chengeta Thompson, Denver journal of International law & policy, vol. 45, num. 1 Fall, January 2016.*
- "Intelligenza artificiale, *Human oversight* e responsabilità penale: prove d'impatto a livello europeo", Giannini Alice, Criminalia, annuario di scienze politiche, in disCrimen, 21.11.2011.
- "*Terminator* Scenario? Intelligenza artificiale nel conflitto armato: *Lethal Autonomous Weapons Systems* e le risposte del diritto internazionale umanitario", Chesini Federico, *BioLaw Journal* – Rivista di Bio-Diritto, n. 3/2020.
- "Intelligenza artificiale e armi autonome: criticità giuridiche", Ascani Paola Giorgia, Rivista Marittima, gennaio 2022.
- "Intelligenza artificiale: questioni relative all'interpretazione e applicazione del diritto internazionale", Risoluzione del Parlamento europeo 2021/C 456/04.

Articoli internet

- www.esteri.it. "LAWS", IRIAD, aprile 2020.

Altri siti

- https://discover.dtic.mil/

ACRONIMI

A2/AD	Anti Access / Area Denial
AGI	Artificial General Intelligence
ANI	Artificial Narrow Intelligence
ARSIWA	Articles on Responsibility of States for Internationally Wrongful Acts
ASI	Artificial Superior Intelligence
ASW	Anti-Submarine Warfare
ATLAS	Advance Targeting and Lethality Automated System
ATT	Arms Trade Treaty
AWS	Autonomous Weapons Systems
C2	Comando e Controllo
C4ISTAR	Command Control Communications Computer Intelligence Surveillance Target Acquisition Reconnaissance
CANES	Consolidated Afloat Networks and Enterprise Services
CCW	Certain Conventional Weapons
CEDU	Corte Europea di Diritti dell'Uomo
CIG	Corte Internazionale di Giustizia
CODE	Collaborative Operations in Denied Environment
COP	Common Operational Picture
C-RAM	Counter Rocket Artillery and Mortar
CUS	Controllo Umano Significativo
DARPA	Defense Advanced Research Projects Agency
DI	Diritto Internazionale
DIU	Diritto Internazionale Umanitario
DOD	Department of Defense
DSTL	Defense Science and Technology Laboratory
DU	Diritti Umani
EDT	Emerging and Disruptive Technologies
EOD	Explosive Ordnance Disposal
GIDE	Global Information Dominance Experiment
GGE	Group of Governmental Experts
GPS	Global Positioning System
HIL	Human in the Loop
HOL	Human on the Loop
HOOL	Human out of the Loop

IA	Intelligenza Artificiale
IAI	Israel Aerospace Industries
IARPA	Intelligence Advanced Research Projects Agency
ICCPR	International Covenant on Civil and Political Rights
ICRC	International Committee of the Red Cross
IED	Improvised Explosive Device
IRA	Irish Republican Army
ISR	Intelligence Surveillance Reconnaissance
ISTAR	Intelligence Surveillance Target Acquisition Reconnaissance
JCE	Joint Criminal Enterprise
JCRAS	Joint Center for Robotics and Autonomous Systems
LAWS	Lethal Autonomous Weapons Systems
LOAC	Law of Arm Conflicts
LRASM	Long Range Anti-Ship Missile
MAD	Mutual Assured Destruction
MHC	Meaningful Human Control
NATO	North Atlantic Treaty Organization
NORAD	North American Aerospace Defense Command
OODA	Observe – Orient – Decide – Act
PGM	Precision-Guided Munition
R3	Rapid Relevant Response
ROE	Rules of Engagement
RPM	Revolution per Minute
SCO	Strategic Capabilities Office
STC	Science and Technology Committee
STCTTS	Sub Committee on Technology Trends and Security
TTP	Tactics Technics and Procedures
UAS	Unmanned Aerial Systems
UAS	Unmanned Aircraft System
UAV	Unmanned Aerial Vehicle
UE	Unione Europea
UN	United Nations
UNCLOS	United Nations Convention on the Law of the Sea
USNORTHCOM	United States Northern Command
USV	Unmanned Sea Vehicle


9 791255 150572