

# «Terminator Scenario?» Intelligenza artificiale nel conflitto armato: *Lethal Autonomous Weapons Systems* e le risposte del diritto internazionale umanitario

Federico Chesini\*

«TERMINATOR SCENARIO?» INTELLIGENZA ARTIFICIALE NEL CONFLITTO ARMATO: *LETHAL AUTONOMOUS WEAPONS SYSTEMS* E LE RISPOSTE DEL DIRITTO INTERNAZIONALE UMANITARIO

ABSTRACT: Among the many fields of foreseeable application of artificial intelligence, warfare must certainly be included. Operational and economic considerations argue in favor of algorithmic decision-making in said context. This raises a vast number of questions, both ethical and legal, peculiarly relevant to lethal autonomous weapons systems, defined as weapons that can select and engage their targets without human intervention, once activated. This contribution aims to illustrate and address the questions raised by such weapons, ranging from compliance with international humanitarian law to accountability and human dignity, and argues that human control over autonomous weapons can and must be retained.

KEYWORDS: Artificial intelligence; armed conflict; international humanitarian law; autonomous weapons; human control

SOMMARIO: 1. Introduzione e definizioni: per un inquadramento generale – 2. LAWS e regolazione internazionale degli armamenti nel diritto umanitario – 3. La selezione di (legittimi) obiettivi: il regime di *targeting law* – 3.1. Il principio di distinzione – 3.2. Il principio di proporzionalità – 3.3. Il principio di precauzione – 3.4. Il principio di necessità militare – 4. Questioni tecniche e questioni «ontologiche» - 4.1. Responsabilità: ipotesi ed interrogativi – 4.2. Decisione algoritmica e dignità umana – 5. Una possibile soluzione? La prospettiva del controllo umano significativo – 6. Conclusioni.

## 1. Introduzione e definizioni: per un inquadramento generale

Sembra sempre meno dubbio che l'intelligenza artificiale (*artificial intelligence*, d'ora in poi AI) possa rappresentare una nuova rivoluzione tecnologica per il genere umano, o quanto meno giocare un ruolo essenziale nella stessa. Il suo potenziale applicativo si estende a numerosi settori dell'attività umana: salute, circolazione di veicoli, informazione, sicurezza, traduzione ed amministrazione della giustizia costituiscono solo alcuni degli esempi che potrebbero essere suggeriti al fine di illustrarne la pervasività. In questi ed altri campi, i benefici che l'AI può apportare (ed apporta) all'uomo sembrano indubbi. Ciò specie in considerazione degli sviluppi che le tecniche di apprendimento automatico (*machine learning*), ivi comprese quelle di apprendimento profondo

\* Dottore in Giurisprudenza presso il Dipartimento Facoltà di Giurisprudenza dell'Università degli Studi di Trento. Mail: [federico.chesini-1@alumni.unitn.it](mailto:federico.chesini-1@alumni.unitn.it). Contributo sottoposto a doppio referaggio anonimo.



(*deep-learning*), hanno avuto nel corso degli ultimi anni: per cogliere il potenziale delle prospettive da tali innovazioni aperte, basti pensare al *buzz* creato, nel mondo scientifico, dalla notizia della vittoria di un sistema dotato di AI contro uno tra i migliori giocatori al mondo di Go<sup>1</sup>.

Al contempo, le innovazioni nel campo dell'AI sollevano importanti questioni di carattere tanto etico, quanto giuridico<sup>2</sup>. In considerazione dei beni giuridici interessati, tali questioni appaiono particolarmente rilevanti nel contesto del conflitto armato. Storicamente, il settore bellico ha sempre attinto risorse dallo sviluppo tecnologico e non sembrano sussistere ragioni per asserire che l'AI possa fare eccezione<sup>3</sup>. Ciò specie in ragione del fatto che l'allontanamento dell'essere umano dal campo di battaglia sembra costituire un obiettivo di lungo corso nella storia dell'industria militare<sup>4</sup>. Questo in primis al fine di ridurre il numero di caduti, tutelando, almeno parzialmente, la società dalle conseguenze negative del conflitto, ma anche per motivazioni di stampo prettamente utilitaristico: i tempi di reazione degli esseri umani e le variazioni negli stati d'animo cui possono andare soggetti possono compromettere il successo in battaglia<sup>5</sup>; inoltre, essendo gli uomini dotati di autonoma volontà, le loro azioni possono sfociare in episodi di disobbedienza ed ammutinamento<sup>6</sup>. Testimonianza della tendenza all'allontanamento dell'uomo dal teatro degli scontri sembra essere il crescente impiego, negli ultimi anni, di droni e dispositivi a controllo remoto nelle operazioni militari<sup>7</sup>. L'implementazione dell'AI in questo settore potrebbe permettere di compiere un passo ulteriore, in questa prospettiva, consentendo la delegazione, da parte dell'uomo, anche dei processi decisionali. Ciò, è stato argomentato, comporterebbe rilevanti vantaggi operativi: un sistema dotato di AI potrebbe raccogliere e processare informazioni molto più agevolmente e rapidamente rispetto ad un essere umano, risul-

<sup>1</sup> I. ANTONOGLIOU, *et al.*, *Mastering the game of Go with deep neural networks and tree search*, in *Nature*, 529, 2016, 484 ss; la notizia ha causato grande fermento in quanto il gioco di Go era stato precedentemente considerato, a causa della sua complessità e della logica induttiva su cui si fonda, una sfida difficilmente superabile per l'AI, specifica L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, (trad. di M. NAPPI), Torino, 2018, 26.

<sup>2</sup> Alcune di queste saranno richiamate e discusse nel corso di questo lavoro, quando rilevanti ai fini dell'oggetto. Per una analisi organica delle stesse si rinvia tuttavia, in maniera puramente esemplificativa, a: C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, Fascicolo speciale, 2019, 101; A. D'ALOIA, *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in *BioLaw Journal – Rivista di Biodiritto*, 1, 2019, 3; A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal – Rivista di Biodiritto*, 1, 2019, 63; J.S. AZADIAN, G.M. FAHY, *Artificial Intelligence and the Law: Navigating "Known Unknowns"*, in *The Computer & Internet Lawyer*, 35, 2, 2018, 19; G.S. ALI, R. YU, *What's Inside the Black Box? AI Challenges for Lawyers and Researchers*, in *Legal Information Management*, 19, 1, 2019, 2; W.N. PRICE, *Artificial Intelligence in Health Care: Applications and Legal Issues*, in *The SciTech Lawyer*, 14, 1, 2017, 10.

<sup>3</sup> J.S. THURNHER, *Means and Methods of the Future: Autonomous Systems*, in P.A.L. DUCHEINE, F.P.B. OSINGA, M.N. SCHMITT (a cura di), *Targeting: The Challenges of Modern Warfare*, The Hague, 2016, 182.

<sup>4</sup> E. FUZAYLOVA, *War torts, autonomous weapons systems, and liability: why a limited strict liability tort regime should be implemented*, in *Cardozo Law Review*, 40, 3, 2019, 1335 ss.

<sup>5</sup> C. HEYNS, *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, Christof Heyns, del 9 aprile 2013, Human Rights Council, para. 53-54.

<sup>6</sup> R. FRAU, W. HEINTSCHEL VON HEINEGG, T. SINGER, *Introduction*, in R. FRAU, W. HEINTSCHEL VON HEINEGG, T. SINGER (a cura di), *Dehumanization of warfare: legal implications of new weapon technologies*, Cham, 2018, 3.

<sup>7</sup> HUMAN RIGHTS WATCH, *Losing our Humanity. The Case against Killer Robots*, 2012, 6.



tandone un grado di efficienza maggiore ed una netta riduzione nei tempi di reazione<sup>8</sup>. Inoltre, l'automazione potrebbe implicare l'eliminazione di una componente, la comunicazione tra uomo e macchina, particolarmente vulnerabile ad attacchi elettronici e *cyber*<sup>9</sup>, nonché richiedente un non irrilevante carico di lavoro umano. Anche in ragione della riduzione dello stesso, peraltro, sostengono numerosi autori, l'impiego dell'AI nelle operazioni militari comporterebbe un decremento dei costi, specialmente nel lungo periodo<sup>10</sup>.

Per le motivazioni sopra accennate, l'applicazione dell'AI al conflitto maggiormente presagibile – nonché più controversa – è rappresentata da armamenti capaci di deliberazioni autonome in merito ai propri obiettivi. La denominazione che, al momento, appare trovare maggior fortuna per tali dispositivi, in quanto adottata in occasione di incontri, svoltisi a Ginevra, fra rappresentanti dei Governi nazionali presso la *United Nations Convention on Certain Conventional Weapons* (UN CCW)<sup>11</sup>, è quella di *Lethal Autonomous Weapons Systems* (d'ora in poi LAWS). Alcuni autori hanno già indicato i LAWS come la terza grande rivoluzione nel mondo degli armamenti, dopo l'introduzione della polvere da sparo e delle armi nucleari<sup>12</sup>. Una prima questione attiene al profilo definitorio: in seno al dibattito sovranazionale, ha suscitato discussioni il lemma «autonomous». Alcuni interpreti hanno affermato che, per essere definito autonomo e, pertanto, rientrare nella categoria dei LAWS, un armamento dovrebbe presentare requisiti piuttosto gravosi, tra cui «situational understanding», «consciousness» e «self-determination»<sup>13</sup>. Requisiti siffatti potrebbero essere posseduti soltanto da una AI *generale*, se non *forte*, capace di replicare forme di intelligenza umana – un traguardo che appare, al momento, distante, se presagibile<sup>14</sup>. A tale corrente si contrappone la definizione proposta dal Comitato Internazionale della Croce Rossa (CICR), secondo cui riveste qualifica di LAWS «any weapon system

<sup>8</sup> Prospettive di automazione del conflitto armato hanno portato, d'altra parte, numerosi autori a sostenere che, in futuro, il campo di battaglia potrebbe divenire un contesto eccessivamente complesso per gli esseri umani, a causa delle limitazioni insite al loro processo decisionale; si veda, ad esempio, C. GRUT, *The Challenge of Autonomous Lethal Robotics to International Humanitarian Law*, in *Journal of Conflict and Security Law*, 18, 1, 2013, 8.

<sup>9</sup> J.S. THURNHER, *Feasible Precautions in Attack and Autonomous Weapons*, in R. FRAU, W. HEINTSCHEL VON HEINEGG, T. SINGER (a cura di), *op. cit.*, 106.

<sup>10</sup> I. BODE, H. HUELSS, *Autonomous weapons systems and changing norms in international relations*, in *Review of International Studies*, 44, 3, 2018, 410.

<sup>11</sup> In occasione di tali incontri, tale denominazione ha rappresentato peraltro oggetto di discussione: da più parti si è sostenuto come detta locuzione corra il rischio di delimitare eccessivamente il campo per mezzo dell'aggettivo «lethal», specie in considerazione del fatto che non è il carattere della «letalità», quanto l'uso della forza, ad innescare gli obblighi derivanti dal diritto internazionale. Tuttavia, il termine «lethal» è stato infine riconosciuto come un supplemento puramente convenzionale, privo di finalità esclusive, ed in quanto tale mantenuto; si veda R. ACHESON, *It's time to exercise human control over the CCW*, in *Reaching Critical Will*, 7, 2, CCW Report, del 27 marzo 2019, 2.

<sup>12</sup> D. EDMONDS, *Vindication for Anthropocentrism: The Humanitarian Case for a Preemptive Treaty Ban on Autonomous Weapons*, in *Undergraduate Journal of Politics, Policy and Society*, 1, 1, 2018, 189.

<sup>13</sup> Posizione assunta dal Ministero della Difesa britannico e dal Joint Air Power Competence Centre (JAPCC) della NATO, come riporta D. AMOROSO, *Jus in bello and jus ad bellum arguments against autonomy in weapons systems: A re-appraisal*, in *QIL – Questions of International Law*, <http://www.qil-qdi.org/jus-bello-jus-ad-bellum-arguments-autonomy-weapons-systems-re-appraisal/> (ultima consultazione 31/10/2017).

<sup>14</sup> R. LÓPEZ DE MÁNTARAS, *Towards Artificial Intelligence: advances, challenges and risks*, in *Mètode Science Studies Journal*, 9, 2018, 123-125.



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» in maniera indipendente<sup>15</sup>. Nel tentativo di dirimere tale dibattito, un ulteriore filone propone di spostare l'attenzione dall'elemento dell'autonomia verso quello rappresentato dal livello di controllo umano, ricomprendendo tra i LAWS armamenti in grado di «identify and fire on targets without meaningful human intervention»<sup>16</sup>. Tale formulazione permette di aggirare il problema di specificare i caratteri dell'«autonomia»; essa, nella sostanza, appare sovrapponibile alla seconda definizione riportata. Questa risulta apprezzabile per la sua attinenza alla realtà concreta degli sviluppi tecnologici nel campo, nonchè per la sua valenza sostanziale: essa non si limita, difatti, a sistemi che, una volta dispiegati ed azionati, non ammettono intervento correttivo umano (secondo una fortunata tripartizione, indicati quali armi c.d. *man-out-of-the-loop*), bensì si estende anche a dispositivi sulle cui deliberazioni l'uomo mantenga un potere di veto (armi c.d. *man-on-the-loop*), riconoscendo che, nel caso in cui l'operatore abbia una facoltà teorica di *override*, ma nella pratica difetti il tempo necessario al suo esercizio, la decisione in merito all'obiettivo sarà da attribuire integralmente all'AI<sup>17</sup>. Sono esclusi dalla categoria dei LAWS, invece, i sistemi in grado di selezionare i propri obiettivi, ma necessitanti una autorizzazione umana al fine di aprire il fuoco (armi c.d. *man-in-the-loop*).

Per tali ragioni, la definizione proposta dal CICR sarà adottata ai fini del presente lavoro. Ciò comporta un necessario corollario: il riconoscimento dell'esistenza, già da anni, di sistemi riconducibili alla categoria dei LAWS. Si tratta principalmente di sistemi difensivi, posti, come suggerisce il nome, a presidio di determinati luoghi e programmati per rilevare potenziali minacce e rispondervi autonomamente<sup>18</sup>. Ne costituiscono esempi il *Phalanx Close-In Weapon System*, equipaggiato alle navi della Marina statunitense e capace di individuare, tracciare, valutare ed ingaggiare i propri obiettivi<sup>19</sup>, ed il *Counter-Rocket, Artillery and Mortar (C-RAM)*, che ne rappresenta sostanzialmente la trasposizione sulla terraferma<sup>20</sup>; i sistemi *Iron Dome* israeliano e *NBS Mantis* tedesco, dislocati rispettivamente

<sup>15</sup> INTERNATIONAL COMMITTEE OF THE RED CROSS (ICRC), *Views of the International Committee of the Red Cross (ICRC) on autonomous weapon systems*, Ginevra, 2016, 1. A tale definizione si allineano altre che sono state avanzate, cfr. UNITED STATES OF AMERICA DEPARTMENT OF DEFENSE DIRECTIVE NO. 3000.09 (Autonomy in Weapon Systems), 2012, Part. II, 13-14; IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*, New York, 2017, 116; WORLD COMMISSION ON THE ETHICS OF SCIENTIFIC KNOWLEDGE AND TECHNOLOGY (COMEST), *Report of COMEST on Robotics Ethics*, Parigi, 2017, para. 87.

<sup>16</sup> D. AMOROSO, *Jus in bello and jus ad bellum arguments against autonomy in weapons systems*, cit.

<sup>17</sup> H. OKORIE, *Autonomization of Weapons and the Principles of Humanitarian Law*, in *Abia state University Law Journal*, 2016, 3.

<sup>18</sup> Si segnala peraltro l'opinione contraria del Parlamento Europeo, espressa in una Risoluzione la quale ha dichiarato che «weapons and weapon systems specifically designed to defend own platforms, forces and populations against highly dynamic threats such as hostile missiles, munitions and aircraft are not considered lethal autonomous weapon systems», cfr. EUROPEAN PARLIAMENT, *European Parliament resolution of 12 September 2018 on autonomous weapon systems*, 2018, para. 6. Peraltro, tale affermazione potrebbe apparire strumentale alla precedente statuizione secondo cui «none of the weapons or weapon systems currently operated by EU forces are lethal autonomous weapon systems».

<sup>19</sup> J.M. BEARD, *Autonomous Weapons and Human Responsibilities*, in *Georgetown Journal of International Law*, 45, 617, 2014, 630.

<sup>20</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 10.



presso Striscia di Gaza e Penisola del Sinai ed in Afghanistan<sup>21</sup>; ed il *Samsung SGR-1*, robot-sentinella collocato lungo il confine della zona demilitarizzata che separa Corea del Sud e Corea del Nord<sup>22</sup>. Tali sistemi sembrano godere di un buon grado di accettazione: la loro pericolosità per la popolazione civile appare contenuta in quanto dispositivi statici, programmati per compiere una serie di azioni definite e ripetitive in ambienti tendenzialmente strutturati e prevedibili, in cui è meno verosimile si trovino civili, e diretti tipicamente contro armi (missili, razzi, altri proiettili di artiglieria), non persone<sup>23</sup>. Recentemente, tuttavia, ingenti risorse sembrano essere investite nella progettazione e produzione di LAWS aventi finalità offensive, in grado pertanto di lanciare autonomamente un attacco. Tali sistemi appaiono rivestire, al momento, prevalentemente la forma di prototipo: così è, per esemplificare, per lo statunitense *X-47B* ed il britannico *Taranis*, progettati sul modello dei droni ma indipendenti, nello svolgimento di determinate operazioni, dal controllo di un operatore umano<sup>24</sup>, e per lo statunitense *Sea Hunter*, imbarcazione capace di condurre in autonomia missioni di pattugliamento, localizzare ed ingaggiare sommergibili nemici<sup>25</sup>. Fa eccezione l'israeliano *Harpy*: descritta come una combinazione tra un velivolo privo di pilota ed un missile da crociera, tale arma è programmata per sorvolare determinate aree, individuare segnali radar, determinare (mediante la consultazione di un *database*) se provengano da obiettivi nemici e, in tal caso, lanciare un attacco, senza necessità di approvazione umana<sup>26</sup>. Si tratta di un elenco puramente esemplificativo, che sembra testimoniare una tendenza comune: un nuovo orizzonte è stato aperto anche dalla notizia secondo la quale il Gruppo Kalashnikov avrebbe sviluppato un'arma «completamente autonoma», fondata sulla tecnologia delle reti neurali<sup>27</sup>. Per quanto non si riscontrino conferme della stessa, appare difficile negare che le potenze economiche mondiali stiano destinando rilevanti investimenti nel settore, il che sembra comportare significative implicazioni di carattere geopolitico.

Già questi cenni permettono di afferrare la rilevanza del tema in questione. Come forse lecito aspettarsi, l'argomento si è mostrato estremamente divisivo: numerose voci si sono già espresse in favore di un bando nei confronti dei LAWS e campagne pubbliche sono nate allo scopo; al contempo, altra letteratura enfatizza non solo i vantaggi operativi, sopra brevemente riportati, da tali sistemi promessi, ma anche i benefici che potrebbero derivare dall'implementazione nel conflitto dell'AI in termini umanitari. Gli argomenti spaziano dal rischio di una nuova corsa agli armamenti (potenzialmente coinvolgente anche attori non statali e cellule terroristiche) a quello connesso ad una minore attenzione dell'opinione pubblica nei confronti della guerra, dal venir meno di stati d'umore e sensazioni umane sul campo di battaglia (in accezione tanto positiva, quanto negativa) ad istanze di sicu-

<sup>21</sup> J.C. ROSSI, *La guerra che verrà: le armi autonome*, in *Sistema Informativo a Schede*, 11, 2016, 7-8.

<sup>22</sup> J. KUMAGAI, *A robotic Sentry for Korea's Demilitarized Zone*, in *IEEE Spectrum*, 3, 2007, 16-17.

<sup>23</sup> HEINRICH BÖLL FOUNDATION (a cura di), *Autonomy in Weapon Systems: The Military Application of Artificial Intelligence as a Litmus Test for Germany's New Foreign and Security Policy*, in *Publication Series on Democracy*, 49, 21.

<sup>24</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 16-17.

<sup>25</sup> G. LUCAS, R. SPARROW, *When robots rule the waves?*, in *Naval War College Review*, 69, 4, 2016, 53-54.

<sup>26</sup> N.E. SHARKEY, *The inevitability of autonomous robot warfare*, in *International Review of the Red Cross*, 94, 886, 2012, 788.

<sup>27</sup> R. HARIDY, *Kalashnikov's new autonomous weapons and the "Terminator conundrum"*, in *New Atlas*, <https://newatlas.com/kalashnikov-ai-weapon-terminator-conundrum/50576/> (ultima consultazione 21/07/2017).



rezza. Nel contesto di tale dibattito, il diritto non sembra poter esimersi dal prendere posizione in merito.

Questo lavoro, pertanto, cercherà di interrogarsi in merito alla conformità dei LAWS alla luce della branca del diritto internazionale che regola la conduzione delle ostilità – il diritto umanitario. Oggetto del secondo e del terzo paragrafo saranno le questioni che i principi fondamentali di tale corpo di diritto pongono ad armamenti dotati di autonomia. Si rivela a questo punto necessaria una premessa metodologica: sussiste una *summa divisio*, all'interno del diritto umanitario, tra conflitti armati internazionali (*international armed conflicts*, o IAC) e conflitti armati di carattere non-internazionale (*non-international armed conflicts*, NIAC). Solamente i primi vedono contrapposti (almeno) due Stati sovrani. Tale dicotomia si riflette nell'ordinamento giuridico internazionale: IACs e NIACs sono soggetti a regimi distinti, i cui caratteri fondamentali sono contenuti, rispettivamente, nel I e nel II Protocollo Aggiuntivo alle Convenzioni di Ginevra. I principi che saranno affrontati nel secondo e nel terzo paragrafo del presente lavoro, tuttavia, quantomeno nel loro nucleo fondamentale, sono assurti a norme di diritto internazionale consuetudinario, applicabili a qualsiasi genere di conflitto<sup>28</sup>. Pertanto, anche per ragioni di spazio, la trattazione sarà nel presente lavoro univoca – seppur mai dimentica della fondamentale distinzione riportata, che sarà richiamata quando rilevante. Dopo aver argomentato che gli obblighi di diritto umanitario non postulano una incompatibilità assoluta, in astratto, per i LAWS, nel quarto paragrafo si tenterà di analizzare alcune questioni che si definiranno «ontologiche» in merito al rapporto fra LAWS e diritto internazionale. Infine, si cercherà di interrogarsi circa possibilità (ed opportunità) di individuare una soluzione alle stesse mediante il concetto – emerso, in seno alla UN CCW, nel dibattito sovranazionale – di controllo umano; e, conseguentemente, di incentrare attorno ad esso una regolazione della materia, tanto *de iure condito* quanto, soprattutto, *de iure condendo*.

## 2. LAWS e regolazione internazionale degli armamenti nel diritto umanitario

Affermare che i LAWS possano trovare disciplina in un solo ambito del diritto sembra eccessivamente limitante. L'implementazione dell'AI nelle operazioni militari può avere rilevanza per varie branche del diritto internazionale, nei confini dei rispettivi ambiti di applicazione (diritto internazionale dei diritti umani, diritto del mare, diritto aerospaziale, diritto penale internazionale)<sup>29</sup>, né sembra scevra da implicazioni di carattere costituzionale<sup>30</sup>. In ogni caso, l'ambito che appare interessato in maniera più

<sup>28</sup> Aderisce a tale interpretazione assolutamente maggioritaria lo studio commissionato dal Comitato Internazionale della Croce Rossa sull'estensione del diritto umanitario consuetudinario, cfr. ICRC (a cura di L. DOSWALD-BECK, J. HENCKAERTS), *Customary International Humanitarian Law*, Cambridge, 2005, Rules 1, 7, 11, 14, 15, 70, 71.

<sup>29</sup> R. CROTOF, *The Varied Law of Autonomous Weapon Systems*, in P.D. SCHARRE, A.P. WILLIAMS (a cura di), *Autonomous Systems: Issues for Defence Policymakers*, Norfolk (Virginia), 2015, 107-115.

<sup>30</sup> Tra le più rilevanti appare il rischio di una diminuzione del controllo democratico sulla guerra. Come sopra accennato, la riduzione dell'attenzione dell'opinione pubblica sul tema costituisce una possibile conseguenza dell'impiego di LAWS. Qualora il costo di un conflitto per l'attaccante, in termini di vite umane, diminuisse radicalmente, la sensibilità sociale sul tema potrebbe sperimentare un netto decremento e la decisione in merito all'entrata in guerra assumere il carattere di una determinazione meramente tecnica ed economica, con conseguente disincentivo al controllo delle assemblee rappresentative nei confronti dell'esecutivo, in



diretta dalla novità portata dai LAWS – e che appare meglio in grado di fornire una visione sistematica – è rappresentato dal diritto internazionale umanitario<sup>31</sup>. Esso rappresenta la branca del diritto internazionale pubblico composta dai principi e dalle norme che disciplinano la conduzione delle ostilità una volta che un conflitto ha avuto inizio. Al suo interno, particolare rilievo assume, per il tema in oggetto, il ramo relativo alla selezione degli armamenti.

In qualsiasi conflitto armato, il diritto delle parti belligeranti di scegliere mezzi e metodi di guerra non è illimitato<sup>32</sup>. Alla necessità di garantire l'effettività di tale principio risponde l'art. 36 del I Protocollo Aggiuntivo alle Convenzioni di Ginevra: «[i]n the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party<sup>33</sup>».

La valutazione imposta dalla norma menzionata non è soggetta a modelli predefiniti o a soluzioni *one-size-fits-all*<sup>34</sup>, deve essere condotta in momenti diversi e susseguenti del processo di sviluppo della nuova arma, deve determinare qualora l'impiego dell'armamento sarà legittimo «in some or all circumstances» e assume quale parametro qualsiasi norma di diritto internazionale ad esso applicabile<sup>35</sup>. Al netto delle norme convenzionali che determinano limitazioni e divieti rivolti a specifiche armi, sussistono in proposito alcune regole rivestenti carattere generale.

Prima tra esse è costituita dal divieto di armamenti i quali, ad una valutazione *ex ante*, possano causare sofferenze non necessarie (al raggiungimento degli obiettivi militari)<sup>36</sup>. Di conseguenza, non sono proibite *per se* armi che possono arrecare danni particolarmente gravi, persino «horrendous», né armi che non lasciano possibilità di sopravvivenza<sup>37</sup>, bensì quelle che provocano sofferenze eccessive rispetto a quelle, per così dire, *inevitabili* per il raggiungimento degli scopi bellici<sup>38</sup>. Sono altresì vietati gli armamenti «which are intended, or may be expected, to cause widespread, long-term and

---

controtendenza rispetto ai dogmi del costituzionalismo moderno. Purtroppo difetta lo spazio per approfondire la questione, che appare di indubbia rilevanza; si rinvia pertanto a D. AMOROSO, G. TAMBURRINI, *The Ethical and Legal Case Against Autonomy in Weapons Systems*, in *Global Jurist*, 18, 1, 2018.

<sup>31</sup> Ciò sembra essere emerso con evidenza anche negli incontri di Ginevra presso la UN CCW, cfr. GROUP OF GOVERNMENTAL EXPERTS ON EMERGING TECHNOLOGIES IN THE AREA OF LETHAL AUTONOMOUS WEAPONS SYSTEMS (GGE ON LAWS), *Draft Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, Ginevra, 2019, para. 16. Da un punto di vista puramente terminologico, alcuni autori hanno preferito, alla locuzione «diritto internazionale umanitario», espressioni differenti, quali la più tradizionale «ius in bello» ovvero «diritto internazionale dei conflitti armati». In ogni caso, resta ferma la valenza sinonimica fra di esse, che risultano dunque intercambiabili, cfr. Y. DINSTEIN, *The conduct of hostilities under the law of international armed conflict*, Cambridge, 2016, 20.

<sup>32</sup> *Protocollo aggiuntivo alle Convenzioni di Ginevra del 12 agosto 1949 relativo alla protezione delle vittime dei conflitti armati internazionali* (d'ora in poi I Protocollo Aggiuntivo), 1977, art. 35(1).

<sup>33</sup> I Protocollo Aggiuntivo, art. 36.

<sup>34</sup> G. SOLIS, *The law of armed conflict: international humanitarian law in war*, New York, 2016, 752.

<sup>35</sup> D. AKERSON, *The Illegality of Offensive Lethal Autonomy*, in D. SAXON (a cura di), *International Humanitarian Law and the Changing Technology of War*, Leiden, 2013, 95.

<sup>36</sup> I Protocollo Aggiuntivo, art. 35(2).

<sup>37</sup> Y. DINSTEIN, *The conduct of hostilities under the law of international armed conflict*, cit., 74-75.

<sup>38</sup> INTERNATIONAL COURT OF JUSTICE (ICJ), *Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion*, in *I.C.J. Reports 1996*, dell'8 luglio 1996, para. 78.



severe damage» all'ambiente naturale<sup>39</sup>. Anche in tal caso, la valutazione è da effettuarsi a priori e il triplice requisito, la cui formulazione già di per sé sembra imporre una soglia piuttosto alta, deve essere integrato cumulativamente<sup>40</sup>. Questi due divieti non sembrano rilevare, con riferimento ai LAWS: l'elemento di novità da questi ultimi apportato concerne le modalità d'ingaggio, mentre le disposizioni appena accennate appaiono relative agli effetti cui l'impiego di un'arma può dare vita<sup>41</sup>. La differenza fondamentale tra LAWS ed altri armamenti risiede nella loro autonomia, non nel livello di sofferenze o nel tipo di danno che possano infliggere: i divieti in questione potranno riguardare le munizioni con cui un sistema autonomo sia equipaggiato, ma non i LAWS quale categoria generale<sup>42</sup>. Discorso parzialmente diverso va effettuato per il terzo divieto di carattere generale, ossia la proibizione di armi indiscriminate (talora rubricate *blind weapons*). Esse sono definite quali mezzi di combattimento che non possono essere diretti contro obiettivi militari, ovvero i cui effetti non possono essere limitati come richiesto dalle norme rilevanti<sup>43</sup>. Il diritto umanitario conosce solo definizioni negative per civili ed oggetti civili, costruite sulle nozioni di combattenti ed obiettivi militari<sup>44</sup>. Questi ultimi sono indicati quali gli oggetti che, per natura, collocazione, finalità o uso, offrono un effettivo contributo all'azione bellica e la cui cattura, distruzione o neutralizzazione fornisce, nelle circostanze del caso concreto, un vantaggio militare *definito*<sup>45</sup>. Gli attributi che possono rendere tale un legittimo obiettivo militare palesano la natura dinamica e relativa di questo concetto nel contesto del conflitto armato<sup>46</sup>. Uno stesso oggetto, in momenti differenti, può rappresentare un oggetto civile o un obiettivo militare. Questa apertura interpretativa, inevitabilmente, si ripercuote sull'individuazione degli stessi oggetti civili, proponendo quesiti non sempre di immediata soluzione a chi debba giudicare la legittimità dell'obiettivo di un attacco<sup>47</sup>. Se tali quesiti possano trovare adeguata risposta da parte di un sistema dotato di AI è un interrogativo su cui difetta unanimità di vedute. Difatti, parte della letteratura sostiene che il divieto di armi indiscriminate poggi sulla necessaria premessa di un effettivo

<sup>39</sup> I Protocollo Aggiuntivo, art. 35(3).

<sup>40</sup> Per tale ragione, la previsione in questione sembra avere scarsa rilevanza pratica nella tutela dell'ambiente: la norma, difatti, sembra ricomprendere solo casi di danni ambientali estremamente significativi, scolastici o quasi, tali probabilmente da non lasciare dubbio alcuno quanto alla illegittimità dell'arma impiegata, cfr. S. VÖNEKY, R. WOLFRUM, *Environment, Protection in Armed Conflict*, in *Max Planck Encyclopedia of Public International Law*, New York, 2016, para. 27-29. Va inoltre segnalato come spesso la completa estensione degli effetti di un'arma sull'ambiente può diventare apparente solo in maniera graduale, rischiando di essere, quasi inevitabilmente, ignorata da una valutazione *ex ante*, cfr. D. CONSTANTIN, J. HENCKAERTS, *Protection of the Natural Environment*, in A. CLAPHAM, P. GAETA (a cura di), *The Oxford Handbook of International Law in Armed Conflict*, Oxford, 2014, 479.

<sup>41</sup> M.N. SCHMITT, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, in *Harvard National Security Journal Features*, <https://harvardnsj.org/2013/02/autonomous-weapon-systems-and-internationalhumanitarian-law-a-reply-to-the-critics/> (ultima consultazione 05/02/2013).

<sup>42</sup> H. OKORIE, *op. cit.*, 7.

<sup>43</sup> I Protocollo Aggiuntivo, art. 51(4)(b), 51(4)(c).

<sup>44</sup> I Protocollo Aggiuntivo, art. 50(1), 52(1).

<sup>45</sup> I Protocollo Aggiuntivo, art. 52(2).

<sup>46</sup> M. WAGNER, *Autonomy in the Battlespace: Independently operating Weapon Systems and the Law of Armed Conflict*, in D. SAXON (a cura di), *op. cit.*, 112.

<sup>47</sup> G. NOLL, *Sacrificial Violence and Targeting in International Humanitarian Law*, in O. ENGDahl, P. WRANGE (a cura di), *Law at war: the law as it was and the law as it should be, liber amicorum Ove Bring*, Leiden, 2008, 212.





controllo dell'uomo sugli armamenti che utilizza<sup>48</sup>: solo un essere umano potrebbe affidabilmente «dirigere» un'arma contro un obiettivo militare e garantire che i suoi effetti siano adeguatamente limitati. Tuttavia, il diritto umanitario non sembra imporre questo elemento umano, fintantoché il requisito imposto dalla disposizione in analisi possa essere soddisfatto<sup>49</sup>: ove ad un LAWS siano forniti dati sufficientemente precisi ed attendibili, pare difficile che esso non possa essere diretto *in alcun caso* (come l'assolutezza del divieto sembra suggerire) contro un obiettivo militare. Peraltro, vi è chi si è spinto oltre, sostenendo che, anche qualora si rivelasse completamente incapace di distinguere fra civili e combattenti, oggetti civili ed obiettivi militari, un LAWS non sarebbe comunque *per se* illegittimo, allorché fosse impiegato in contesti in cui non sia possibile rinvenire civili ed oggetti civili (ad esempio, nella guerra sottomarina o in ambienti desertici)<sup>50</sup>.

### 3. La selezione di (legittimi) obiettivi: il regime di *targeting law*

Assumendo l'angolo prospettico fornito dalle regole relative agli armamenti, dunque, non sembrano potersi inferire argomenti contrari alla compatibilità dei LAWS con il diritto umanitario. Si rivela tuttavia doveroso estendere il campo di analisi ad un altro ambito: nel caso in cui l'essere umano sia esautorato dalle decisioni più rilevanti nell'utilizzo di un'arma, appare necessario un criterio di legittimità addizionale. Il diritto umanitario contiene una serie di principi e regole che determinano la liceità di uno specifico attacco, disciplinando le operazioni di selezione degli obiettivi. Dove un'arma assuma su di sé tale scelta, nonché quella relativa all'ingaggio, una *legal review* quale quella imposta dall'art. 36 del I Protocollo Aggiuntivo deve accertare anche che essa sia in grado di conformarsi agli stessi<sup>51</sup>.

#### 3.1. Il principio di distinzione

La *regola aurea*, in materia, è costituita dal principio di distinzione, fondamentale al punto da non ammettere in alcun caso deroghe in nome di istanze di necessità militare<sup>52</sup>. Esso prescrive che, «in ogni momento», i belligeranti debbano distinguere tra civili (ed oggetti civili) e combattenti (ed obiettivi militari)<sup>53</sup>. Da tale statuizione discendono, per quanto qui interessa<sup>54</sup>, i divieti di attacchi che non

<sup>48</sup> T. CHENGETA, *Defining the Emerging Notion of 'Meaningful Human Control' in Autonomous Weapon Systems*, in *New York University Journal of International Law and Politics*, 49, 3, 2017, 843-844. L'autore porta a dimostrazione della propria tesi il bando, ad opera della Convenzione di Ottawa, di mine antiuomo, armi i cui comportamento e obiettivi non potrebbero essere predetti al momento del sotterramento, salvo il caso di sorveglianza umana sull'area minata ed azionamento a distanza. Peraltro, il Preambolo della stessa Convenzione cita, a sostegno di tale bando, il divieto di sofferenze non necessarie ed il principio di distinzione, non la proibizione di armi indiscriminate, cfr. *Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on Their Destruction*, Ottawa, del 3 dicembre 1997, Preamble.

<sup>49</sup> J.S. THURNHER, *Means and Methods of the Future*, cit., in P.A.L. DUCHEINE, F.P.B. OSINGA, M.N. SCHMITT (a cura di), *op. cit.*, 186.

<sup>50</sup> M.N. SCHMITT, *op. cit.*, 11.

<sup>51</sup> W.H. BOOTHBY, *Weapons and the Law of Armed Conflict*, Oxford, 2016, 348-349.

<sup>52</sup> G. SOLIS, *op. cit.*, 271.

<sup>53</sup> *I Protocollo Aggiuntivo*, art. 48.

<sup>54</sup> Ulteriore corollario del principio di distinzione è costituito dal divieto di atti e minacce di violenza la cui finalità primaria sia diffondere il terrore tra la popolazione civile: è evidente che un sistema programmato per



siano diretti contro obiettivi militari ovvero indiscriminati, ossia quando caratterizzati da disinteresse (o incapacità) dell'attaccante di rivolgerli verso obiettivi militari<sup>55</sup>. Ad un primo sguardo, una questione binaria, quale quella posta dal principio di distinzione, può apparire particolarmente confacente ad un algoritmo. Tuttavia, l'affermazione della binarietà della stessa può apparire ingannevole: come sopra riportato, il diritto umanitario accoglie una definizione meramente negativa per i concetti di «civile» ed «oggetto civile» e ciò sembra rendere impossibile una *codificazione* degli stessi che possa offrire ad un LAWS criteri di valutazione adeguati<sup>56</sup>. Inoltre, la già segnalata natura dinamica e mutevole degli obiettivi militari impone uno scrutinio discrezionale e *contestuale* per stabilire la liceità di un determinato bersaglio: tale valutazione appare più adeguata all'intelligenza umana che a quella artificiale<sup>57</sup>, manchevole di buon senso, nonché nei campi di percezione del contesto e della rilevanza<sup>58</sup>. Ciò sembra particolarmente rilevante alla luce delle forme di conflitto statisticamente più frequenti in età contemporanea, tendenzialmente combattute tra potenze belliche e singoli gruppi armati e, dunque, connotate da una asimmetria crescente, la cui conseguenza è il largo impiego delle tattiche della guerriglia<sup>59</sup>: i combattenti finiscono così spesso per impiegare scudi umani e nascondersi tra i civili, evitando di indossare uniformi ed insegne ed occultando qualsiasi altro elemento distintivo. In considerazione di ciò, pare oggi doversi richiedere una *situational awareness* ulteriormente rafforzata<sup>60</sup>: i LAWS, incapaci, per definizione, di empatia, risulterebbero svantaggiati dalla loro minore capacità di riconoscere intenzioni e stati emotivi umani e di distinguere tra reali e simulati<sup>61</sup>. Molto si è dibattuto – e si dibatte – in merito alla mancanza di emotività nei LAWS. Mentre alcuni autori sottolineano che empatia e compassione possono contribuire a salvare vite, altri richiamano invece l'attenzione sul difetto, nell'AI, di sensazioni, per così dire, *negative*: un LAWS non avrebbe reazioni istintive<sup>62</sup>, non sarebbe soggetto a stress, ansia o altre variazioni psicologiche che possono comportare violazioni del diritto<sup>63</sup>, né agirebbe spinto da paura, odio, sete di sangue, desiderio di vendetta o pulsioni inconsce<sup>64</sup>. Di certo, tuttavia, la mancanza di empatia pare rilevante anche per un'ul-

---

effettuare questo tipo di operazioni sarebbe certamente illecito. Inoltre, dal principio in questione discende anche l'obbligo di non «offuscare» la distinzione, ad esempio collocando installazioni militari in aree densamente popolate, cfr. E. CRAWFORD, A. PERT, *International Humanitarian Law*, Cambridge, 2015, 43.

<sup>55</sup> Y. DINSTEIN, *The conduct of hostilities under the law of international armed conflict*, cit., 146.

<sup>56</sup> N.E. SHARKEY, *op. cit.*, 789.

<sup>57</sup> C. GRUT, *op. cit.*, 11.

<sup>58</sup> È stato affermato che questi caratteri potrebbero costituire l'ostacolo maggiore alla realizzazione di una AI generale «human-like», cfr. M. BODEN, *Artificial Intelligence: a very short introduction*, Oxford, 2018, 136-137; analogamente, si vedano anche L.J. CROCKETT, *The Turing Test and the Frame Problem: AI's mistaken understanding of intelligence*, Norwood (New Jersey), 1994, 62-65; R. LÓPEZ DE MÁNTARAS, *Towards Artificial Intelligence*, cit., 123.

<sup>59</sup> J. CROWE, K. WESTON-SCHUEBER, *Principles of International Humanitarian Law*, Cheltenham, 2013, 41-43; S.C. NEFF, *War and the law of nations: a general history*, Cambridge, 2008, 357-358; C.M. CHINKIN, M. KALDOR, *International Law and new wars*, Cambridge, 2017, 322.

<sup>60</sup> HEINRICH BÖLL FOUNDATION (a cura di), *op. cit.*, 24.

<sup>61</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 31-32.

<sup>62</sup> W.H. BOOTHBY, *Weapons and the Law of Armed Conflict*, cit., 250.

<sup>63</sup> S. HAINES, *The Developing Law of Weapons: Humanity, Distinction, and Precautions in Attack*, in A. CLAPHAM, P. GAETA (a cura di), *op. cit.*, 293.

<sup>64</sup> M. SASSÒLI, *Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified*, in *International Law Studies*, 90, 1, 2014, 310.





riore profilo del principio di distinzione: secondo lo stesso, difatti, i civili non possono mai essere oggetto di un attacco diretto, salvo il caso in cui (e solo per il lasso di tempo in cui) essi partecipino direttamente alle ostilità. La nozione di *direct participation* non riceve una definizione vincolante dal diritto umanitario: tale non è la guida ermeneutica fornita dal CICR, la quale, pur dotata di autorevolezza<sup>65</sup>, ne riflette puramente le opinioni istituzionali. La valutazione in merito, di conseguenza, non può che essere effettuata nel caso concreto ed in maniera contestuale, mediante lettura ed interpretazione delle intenzioni del civile supposto partecipante alle ostilità<sup>66</sup>.

Dall'altro lato, è stato affermato che i critici dei LAWS non sarebbero in grado di apprezzare adeguatamente la rilevanza del contesto di impiego: le capacità dei sistemi e le soglie di ammissibilità degli stessi devono essere, difatti, parametrize all'ambiente in cui saranno utilizzati<sup>67</sup>. Così, in contesti privi di civili ed oggetti civili (quali quelli già accennati discorrendo di armi indiscriminate), un LAWS potrebbe superare il test di distinzione pur fornendo, in ipotesi, un basso (o addirittura nullo) livello di affidabilità<sup>68</sup>. Pertanto, il principio di distinzione non sembra, da solo, poter giustificare una affermazione di incompatibilità dei LAWS con il diritto internazionale. Oltretutto, al di fuori da tali contesti, sono state avanzate argomentazioni e soluzioni alternative per garantirne la conformità al principio di distinzione. È stato sottolineato, in particolare, come tali sistemi, al contrario dei soldati umani, non sarebbero condizionati da necessità di auto-conservazione e pertanto potrebbero essere programmati per attendere più a lungo prima di ingaggiare i propri obiettivi, assumendo un grado di rischio più elevato per dissipare eventuali dubbi<sup>69</sup>. Altri autori, inoltre, propongono di consentire ai LAWS di attaccare solo altri *weapons systems*<sup>70</sup>, ovvero solo oggetti che possono *sempre* essere considerati obiettivi militari (obiettivi militari «per natura»)<sup>71</sup>. Questione successiva è se tale soluzione possa rivelarsi conforme al principio di proporzionalità.

<sup>65</sup> Ciò pare riconosciuto anche da alcune voci critiche della guida stessa, cfr. N. MELZER, *The principle of distinction between civilians and combatants*, in A. CLAPHAM, P. GAETA (a cura di), *op. cit.*, 324-325.

<sup>66</sup> Qualora la definizione fornita dal Comitato Internazionale della Croce Rossa fosse adottata quale vincolante, il grado di complessità della valutazione potrebbe essere inferiore, in ragione del fatto che i requisiti da essa indicati (raggiungimento di un determinato livello di danno, sussistenza di un nesso causale sufficiente, obiettiva designazione dell'atto alla provocazione di quel danno) sembrano cercare una oggettività, cfr. ICRC, *Interpretative Guidance on the notion of Direct Participation in Hostilities under International Humanitarian Law*, Ginevra, 2009, 47-64. Peraltro, questi stessi requisiti non appaiono agevolmente convertibili in istruzioni che possano essere trasferite all'interno di un algoritmo, cfr. M. SASSÒLI, *op. cit.*, 328. La complessità della valutazione è ulteriormente inasprita dalla necessità di distinguere, nei NIACs, tra civili direttamente partecipanti alle ostilità e che svolgano *continuous combat function*: questi ultimi sono coinvolti nei gruppi armati parte del conflitto in maniera continuativa, partecipando anche alla fase preparatoria – ovvero comunque antecedente – alle operazioni militari, pertanto la protezione di cui godono è ulteriormente deteriore. Tali soggetti, difatti, potranno costituire legittimo obiettivo militare anche prima di compiere atti ostili, cfr. CRC, *op. cit.*, 33-36. Come un sistema dotato di AI potrà discernere *direct participation* e *continuous combat function* appare questione di difficile risposta.

<sup>67</sup> K. ANDERSON, D. REISNER, M. WAXMAN, *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, in *International Law Studies*, 90, 2014, 401.

<sup>68</sup> H. OKORIE, *op. cit.*, 10.

<sup>69</sup> C. HEYNS, *op. cit.*, para. 69.

<sup>70</sup> C. GRUT, *op. cit.*, 12.

<sup>71</sup> M. SASSÒLI, *op. cit.*, 327.



### 3.2. Il principio di proporzionalità

L'individuazione di un legittimo obiettivo militare non è conclusiva del processo di selezione degli obiettivi. Gli obblighi dell'attaccante si estendono, difatti, alla valutazione degli effetti indiretti che l'offensiva potrà avere nei confronti di civili ed oggetti civili. In ciò si esplica il principio di proporzionalità, il quale richiede che il danno collaterale ragionevolmente attendibile quale conseguenza di un attacco non sia eccessivo rispetto al vantaggio militare che ci si prefigge di ottenere con lo stesso<sup>72</sup>.

Il principio di proporzionalità è stato descritto quale una delle regole più complesse del diritto internazionale umanitario<sup>73</sup>. La valutazione da esso imposta deve essere effettuata *ex ante*, con riferimento alle specifiche circostanze del caso concreto: di conseguenza, si rivela inerentemente contestuale e qualitativa<sup>74</sup>, nonché parametrata su uno standard soggettivo di ragionevolezza<sup>75</sup>. Questi caratteri sembrano allontanare tale scrutinio dalla portata persino dello stato dell'arte nel campo dell'AI. Peraltro, da più parti si è affermato come l'elemento del danno collaterale attendibile possa essere computato da un LAWS: per quanto non sia affatto certo che questa componente si limiti ai danni *materiali* sofferti da civili ed oggetti civili<sup>76</sup>, ciò nondimeno la sua valutazione appare almeno prevalentemente quantitativa e modelli e *software* appositi (tra cui il c.d. CDEM, *Collateral Damage Estimate Methodology*) risultano già esistenti<sup>77</sup>. Analogo discorso, invece, non può essere fatto con riferimento al vantaggio militare attendibile: difettano, difatti, criteri obiettivi per il suo apprezzamento, suscettibile a mutamenti sulla base di considerazioni strategiche tanto di lungo periodo, quanto contingenti. La valutazione di tale elemento deve dunque essere continuativa, ininterrotta, e può variare in ogni momento in base alle decisioni del comandante e agli sviluppi nel contesto del campo di battaglia<sup>78</sup>. La componente di discrezionalità umana, in questo, non appare agevolmente eliminabile.

Peraltro, non va ignorato come il principio di proporzionalità richieda che il danno collaterale prevedibile non sia *eccessivo* rispetto al vantaggio militare atteso: pertanto si avrà una violazione dello stesso nel solo caso di netto squilibrio tra gli interessi in gioco<sup>79</sup>. Tale considerazione potrebbe mitigare il rigore della valutazione da imporsi ai LAWS. Tuttavia, sussiste un ulteriore elemento di complessità: il principio di proporzionalità sembra esigere una comparazione tra «mele e arance»<sup>80</sup>. In altri termini, danno collaterale e vantaggio militare costituiscono valori di diversa natura, non possiedono unità di misura comuni e non sembrano potersi sottoporre ad un confronto «matematizzabile» o codificabile. Tale bilanciamento non dispone, né potrebbe disporre, di parametri predeterminati, essendo in gioco il bene giuridico più rilevante, la vita, e rendendo ciò inaccettabili valutazioni puramente quantitative<sup>81</sup>. Ciò sembra riconosciuto anche dal diritto umanitario stesso, che nega la possi-

<sup>72</sup> I Protocollo Aggiuntivo, art. 51(5)(b), 57(2)(b).

<sup>73</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 32.

<sup>74</sup> E. CRAWFORD, A. PERT, *op. cit.*, 175.

<sup>75</sup> INTERNATIONAL CRIMINAL TRIBUNAL FOR THE FORMER YUGOSLAVIA (ICTY), *Prosecutor v Galić*, Trial Chamber Judgment, del 5 dicembre 2003, para. 58.

<sup>76</sup> WORLD COMMISSION ON THE ETHICS OF SCIENTIFIC KNOWLEDGE AND TECHNOLOGY (COMEST), *op. cit.*, para. 73.

<sup>77</sup> G. SOLIS, *op. cit.*, 540.

<sup>78</sup> D. AKERSON, *op. cit.*, in D. SAXON (a cura di), *op. cit.*, 84; H. OKORIE, *op. cit.*, 16.

<sup>79</sup> R. KOLB, *Advanced Introduction to International Humanitarian Law*, Cheltenham (Regno Unito), 2014, 172.

<sup>80</sup> K. ANDERSON, D. REISNER, M. WAXMAN, *op. cit.*, 403.

<sup>81</sup> A. D'ALOIA, *op. cit.*, 14.

bilità di valutazioni «standardizzate» della vita umana nell'applicazione del principio di proporzionalità (ad esempio, la vita di un adulto e di un bambino potrebbero ricevere considerazione parzialmente differente)<sup>82</sup>. Questo solleva un tema ulteriore: tra le questioni poste dall'AI alla sfera giuridica vi è quella relativa al c.d. *bias*. Lunghi da promesse di assolute imparzialità e neutralità, l'esito delle deliberazioni di un sistema dotato di AI può derivare da, e persino rafforzare, pregiudizi e preconcetti insiti agli algoritmi ed ai dati di *input*<sup>83</sup>. Una questione che non ci si può esimere dal porsi, dunque, è se un LAWS potrebbe assegnare un minor «valore» a determinati soggetti, attribuendo loro meno «peso» nel giudizio di proporzionalità, a causa di differenze, ad esempio, di razza o genere.

Anche le voci più favorevoli ai LAWS sembrano concordare sulla difficoltà, per essi, di conformarsi al principio in esame; tuttavia, non appaiono condividere l'opinione secondo cui ciò ne implicherebbe la mancanza di conformità con il diritto umanitario. In primo luogo, è stata sottolineata anche con riferimento al principio di proporzionalità la rilevanza del contesto di impiego: in ambienti nei quali non si rinvergono civili o oggetti civili, non essendo possibile aversi danno collaterale, la regola non potrà che essere rispettata<sup>84</sup>. In contesti connotati da un più elevato grado di rischio, alcuni autori hanno suggerito che la mancanza di informazioni rilevanti ai fini del vaglio di proporzionalità dovrebbe condurre ad un veto o ad una richiesta di intervento umano<sup>85</sup> – per quanto, in tal caso, secondo la definizione adottata in questo elaborato, non è certo si avrebbe un autentico LAWS. Altri sostengono la necessità, nelle fasi di progettazione e sviluppo, di imporre su tali sistemi correttivi quali limitazioni nel tempo e nello spazio agli attacchi, allo scopo di ridurre il rischio di danni collaterali<sup>86</sup>: vi è chi ritiene che tali correttivi dovrebbero spingersi sino alla programmazione anticipata dei valori massimi di danno collaterale accettabile in diverse situazioni<sup>87</sup>. Tali valori potrebbero essere impostati in maniera conservativa, fermo che i LAWS non sarebbero mossi da istinto di sopravvivenza e, dunque, come

<sup>82</sup> Y. DINSTEIN, *The conduct of hostilities under the law of international armed conflict*, cit., 159-160.

<sup>83</sup> Alcuni autori hanno difatti segnalato come, con l'utilizzo dei sistemi di AI, pregiudizi non scompaiono, ma vengono «camuffati»: tanto i parametri presi in considerazione, quanto i dati forniti, possono avere connotazioni tali da sfavorire, statisticamente, determinate categorie sociali, cfr. C. O'NEIL, *Weapons of math destruction: how big data increases inequality and threatens democracy*, Londra, 2016, 7-10. Casi di *bias* sono già stati registrati in vari settori: nell'ambito dell'amministrazione della giustizia, uno studio ha dimostrato che il software COMPAS, impiegato per la stima di pericolosità sociale e rischio di recidiva degli imputati e portato all'attenzione pubblica dal caso *Loomis v. Wisconsin*, soffriva di un margine di errore pari a circa il 30%, mostrando una probabilità doppia di determinazioni erranee di alto rischio di recidiva nel caso di imputati di colore, cfr. M. ANNEROTH, et al., *Sustainable AI: An inventory of the state of knowledge of ethical, social, and legal challenges related to artificial intelligence*, AI Sustainability Center Report, Lund (Sweden), 2019, 13; A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto, Journal of Legal Philosophy*, 1, 2019, 96-97. Ulteriore esempio deriva dall'impiego di processi decisionali automatizzati nel mondo del lavoro: è stato riportato il caso di un sistema di intelligenza artificiale, utilizzato da Amazon per selezionare le candidature, il quale «escludeva le candidature delle donne; il sistema, infatti, aveva "imparato" dalle candidature passate, in larga parte di uomini, a "preferire" gli uomini alle donne», cfr. F. FAINI, *Intelligenza artificiale e diritto: le sfide giuridiche in ambito pubblico*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2019, 153.

<sup>84</sup> W.H. BOOTHBY, *How far will the Law allow unmanned targeting to go?*, in D. SAXON (a cura di), *op. cit.*, 57.

<sup>85</sup> M. WAGNER, *op. cit.*, in D. SAXON (a cura di), *op. cit.*, 121.

<sup>86</sup> W.H. BOOTHBY, *How far will the Law allow, cit.*, in D. SAXON (a cura di), *op. cit.*, pag 57.

<sup>87</sup> J.S. THURNHER, *Means and Methods of the Future, cit.*, in P.A.L. DUCHEINE, F.P.B. OSINGA, M.N. SCHMITT (a cura di), *op. cit.*, 189.

già riportato, potrebbero sopportare un grado di rischio ben maggiore rispetto agli umani. Tuttavia, questa non appare una soluzione praticabile – non sembra possibile, difatti, prevedere ogni situazione che un LAWS possa ritrovarsi ad affrontare – né legittima: la valutazione di proporzionalità deve sempre rimanere contingente e continuativa.

### 3.3. Il principio di precauzione

Il principio di precauzione si pone l'obiettivo di minimizzare il danno collaterale, imponendo ai belligeranti un dovere di diligenza nelle fasi di pianificazione ed esecuzione di un attacco e fungendo, così, da complemento alle regole di distinzione e proporzionalità<sup>88</sup>. Il dovere generale che discende da tale principio, difatti, è quello di prestare costante attenzione allo scopo di risparmiare civili ed oggetti civili<sup>89</sup>. Tale obbligo, pur estremamente relativo (il danno collaterale pare una conseguenza ineluttabile, in un conflitto), rafforza la statuizione, derivante dal principio di proporzionalità, della necessità di valutazioni *costanti* e continuative: ciò si rivela particolarmente rilevante nel caso dei LAWS, uno dei cui elementi distintivi rispetto ad armi prive di autonomia è costituito dalla possibilità di lanciare un attacco molto tempo dopo il loro dispiegamento<sup>90</sup>.

Il generale dovere di *constant care* si specifica in una serie di obblighi distinti. Alcuni di essi non sembrano rappresentare ostacoli insuperabili per i LAWS: così è per l'obbligo di prendere ogni precauzione possibile, nella scelta degli armamenti, al fine di minimizzare il danno collaterale<sup>91</sup>. Qualora equipaggiato con differenti armi o munizioni, un sistema dovrà essere in grado di determinare quali siano necessarie e sufficienti per il completamento della missione, quale il danno collaterale attendibile dall'utilizzo di ciascuna di esse ed operare una selezione. Trattandosi di un giudizio prevalentemente quantitativo, tale operazione sembra possibile<sup>92</sup>. Maggiori difficoltà potrebbero derivare da altri obblighi, ad esempio quello di fornire avvertimenti efficaci alla popolazione, «unless circumstances do not permit»<sup>93</sup>. Tanto l'efficacia degli avvertimenti, quanto il verificarsi dell'eccezione contenuta nella formula sembrano potersi apprezzare solo con un giudizio contestuale e qualitativo. Tuttavia, il primo problema potrà probabilmente essere risolto in fase di programmazione del sistema, predisponendo gli avvertimenti che potranno rivelarsi più efficaci nelle diverse condizioni in cui presumibilmente il LAWS opererà<sup>94</sup>. Per quanto riguarda l'eccezione, essa non risponde solo ad istanze di segretezza, ma anche e soprattutto di salvaguardia dell'incolumità delle forze attaccanti<sup>95</sup>: stante la mancata necessità di autoconservazione dei LAWS, questa esigenza potrebbe venir meno – garantendo una maggiore tutela, potenzialmente, alla popolazione civile.

<sup>88</sup> R. KOLB, *op. cit.*, 166.

<sup>89</sup> *I Protocollo Aggiuntivo*, art. 57(1).

<sup>90</sup> Per descrivere tale fenomeno sono state impiegate le locuzioni «loitering» e «persistent stare», cfr. D. AKERSON, *op. cit.*, in D. SAXON (a cura di), *op. cit.*, 86.

<sup>91</sup> *I Protocollo Aggiuntivo*, art. 57(2)(a)(ii).

<sup>92</sup> J.S. THURNHER, *Feasible Precautions in Attack*, *cit.*, in R. FRAU, W. HEINTSCHEL VON HEINEGG, T. SINGER (a cura di), *op. cit.*, 113.

<sup>93</sup> *I Protocollo Aggiuntivo*, art. 57(2)(c).

<sup>94</sup> J.S. THURNHER, *Feasible Precautions in Attack*, *cit.*, in R. FRAU, W. HEINTSCHEL VON HEINEGG, T. SINGER (a cura di), *op. cit.*, 112.

<sup>95</sup> R. KOLB, *op. cit.*, 170.

Altri obblighi derivanti dal principio di precauzione sembrano presentare, tuttavia, sfide eccessivamente complesse. Per conformarsi a tale principio, un LAWS dovrebbe essere in grado di sospendere o cancellare un attacco nel caso in cui, ad una valutazione successiva, sopravvenga la sua mancata conformità ai principi di distinzione e/o proporzionalità<sup>96</sup>. Tale requisito presume la capacità di riconoscere una siffatta violazione: quanto illustrato relativamente a distinzione e proporzionalità sembra sottolineare il difetto di tale capacità nei LAWS<sup>97</sup>. Inoltre, qualora si manifesti la possibilità di scelta tra obiettivi che offrono un pari vantaggio militare, l'attaccante deve optare per l'offensiva che, presumibilmente, causerà il minor danno collaterale<sup>98</sup>: come sopra esposto, le caratteristiche dell'elemento del vantaggio militare attendibile appaiono renderlo difficilmente apprezzabile da parte di un sistema dotato di AI.

Così come quelli di distinzione e proporzionalità, anche il principio di precauzione è teso alla tutela di civili ed oggetti civili: pertanto, in contesti dove essi non possano essere rintracciati, non sembrano porsi questioni di conformità. Peraltro, l'esigenza di apprestare, nel conflitto, misure precauzionali è stata indicata anche quale un argomento a favore dell'introduzione dei LAWS. Ciò non soltanto poiché essi, in virtù della propria natura, potrebbero sopportare rischi più elevati e rendere possibili maggiori precauzioni, ma anche per un altro ordine di ragioni: stante l'obbligo di impiegare armamenti che minimizzino il danno ai civili, una declaratoria di illegittimità nei confronti dei LAWS non solo non sarebbe necessaria, ma nemmeno opportuna; qualora maggiori precauzioni possano derivare dall'utilizzo di un sistema a guida umana (o comunque di altra natura), sussisterà il *dovere* di non impiegare uno autonomo, che invece potrà essere dispiegato quando rappresenti la migliore scelta possibile<sup>99</sup>. Tuttavia, la vera questione è relativa alla possibilità, per un'AI, di effettuare la valutazione richiesta dal principio di precauzione: essa pone peculiare attenzione sulla praticabilità della messa in atto di misure precauzionali, concetto che appare richiedere uno scrutinio di ragionevolezza, connesso a variabili contestuali<sup>100</sup>. Vi è chi si è persino chiesto se l'eliminazione stessa della supervisione umana sia in radice compatibile con l'obbligo di prendere tutte le precauzioni possibili<sup>101</sup>. Ed in effetti, in particolare in considerazione dei problemi di imprevedibilità presentati dalle tecniche di *machine learning*<sup>102</sup>, i quali possono condurre a risultati inattesi e contro-intuitivi, non sembra certo potersi escludere a priori che un monitoraggio effettivo ed efficace sulle attività dei LAWS possa costituire un elemento precauzionale a tutela dei civili.

<sup>96</sup> I Protocollo Aggiuntivo, art. 57(2)(b).

<sup>97</sup> Analogamente M. SASSÒLI, *op. cit.*, 337.

<sup>98</sup> I Protocollo Aggiuntivo, art. 57(3).

<sup>99</sup> M.N. SCHMITT, *op. cit.*, 24.

<sup>100</sup> D. AMOROSO, G. TAMBURRINI, *The Ethical and Legal Case Against Autonomy in Weapons Systems*, cit., 14.

<sup>101</sup> D. AMOROSO, G. TAMBURRINI, *What makes human control over weapons systems "meaningful"?*, International Committee for Robot Arms Control (ICRAC), Report to the CCW Group of Governmental Experts, 2019, 7.

<sup>102</sup> Si tratta del problema talora descritto anche con l'espressione *black-box*: i sistemi dotati di tecniche di apprendimento automatico sono tendenzialmente rappresentati quali «scatole» chiuse, che ricevono *input* e forniscono *output* senza spiegarne la ragione, cfr. G.S. ALI, R. YU, *op. cit.*, 5. In questo modo, nemmeno programmatori e sviluppatori sono in grado di prevedere con precisione assoluta quale sarà il risultato finale, poiché ricostruire i passaggi compiuti per addivenirvi diviene virtualmente impossibile, cfr. M. KURMAN, H. LIPSON, *Driverless: Intelligent Cars and the Road Ahead*, Cambridge (Massachusetts), 2016, 228.



### 3.4. Il principio di necessità militare

La necessità militare rappresenta una delle *anime* del diritto umanitario, il quale si sviluppa quale sistema di pesi e contrappesi nel bilanciamento tra considerazioni di carattere umanitario ed esigenze dei belligeranti<sup>103</sup>. Di conseguenza, la necessità militare riveste un ruolo, per così dire, «costituzionale» nel diritto umanitario, operando tanto nell'adozione di nuove regole, quanto nell'interpretazione di quelle esistenti<sup>104</sup>. Il principio di necessità militare sembra, tuttavia, agire anche ad un livello ulteriore ed essere dotato di un significato direttamente applicabile nelle operazioni belliche e, in particolare, in quelle di selezione degli obiettivi. Tale significato appare duplice. In primo luogo, riveste un carattere permissivo: istanze di necessità militare possono giustificare deroghe ad alcune regole di diritto umanitario, quando tali eccezioni siano testualmente previste<sup>105</sup>. Al contempo, il principio di necessità militare conserva una finalità restrittiva, proibendo atti di violenza di per sé non vietati, ma da reputarsi *non necessari* al raggiungimento dell'obiettivo<sup>106</sup>.

Quest'ultima accezione del principio di necessità militare concerne indubbiamente gli attacchi diretti contro soggetti feriti in maniera troppo grave o comunque incapaci di proseguire le ostilità (c.d. *hors de combat*), ovvero che abbiano manifestato la propria intenzione di resa, in quanto non rappresentanti (più) una minaccia. L'adeguamento a tale principio richiederebbe pertanto il riconoscimento delle intenzioni umane e la distinzione fra vere e false<sup>107</sup> – altrimenti, un LAWS sarebbe agevolmente ingannabile e, dunque, vulnerabile. Per le ragioni sopra illustrate, dunque, l'AI non sarebbe probabilmente, allo stato attuale, in grado di conformarsi a tale principio<sup>108</sup>.

Tale considerazione comporta un corollario estremamente rilevante agli scopi di questo lavoro. Mentre i principi esaminati nei paragrafi precedenti sono volti alla tutela dei civili, quello di necessità militare cerca di accordare una protezione (anche) ai combattenti. Pertanto, mentre i principi di distinzione, proporzionalità e precauzione possono dirsi rispettati quando le ostilità si svolgano in contesti dove non si trovano civili, né oggetti civili, quello di necessità militare si applica a qualsiasi scenario di battaglia nel quale possano essere coinvolti degli esseri umani<sup>109</sup>. In tal modo, l'ambito di legittimo impiego dei LAWS sembra ridursi ulteriormente, limitandosi a scontri con veicoli a controllo remoto ovvero altri LAWS, aprendo le porte alla prospettiva di conflitti combattuti interamente ed unicamente tra macchine.

Con specifico riferimento alla questione del riconoscimento della resa, sono state ipotizzate alcune soluzioni. La prima è rappresentata dalla istituzione di un segnale elettronico di resa, riconosciuto a

<sup>103</sup> È stato affermato che non si potrebbero avere norme efficaci di diritto umanitario al di fuori di questa dialettica: regole rispondenti ad esigenze puramente militari tenderebbero all'atrofia, perché eccessivamente permissive; un diritto internazionale «meramente umanitario» rischierebbe di non avere efficacia, in quanto considerato, con ogni probabilità, distaccato dalla realtà, inadeguato alle necessità del personale militare, cfr. R. KOLB, *op. cit.*, 77.

<sup>104</sup> *ibid.*, 83.

<sup>105</sup> J. CROWE, K. WESTON-SCHEUBER, *op. cit.*, 53.

<sup>106</sup> D. FLECK, (a cura di), *The Handbook of International Humanitarian Law*, Oxford, 2013, 37.

<sup>107</sup> H. OKORIE, *op. cit.*, 12-13.

<sup>108</sup> W.H. BOOTHBY, *Does the law of targeting meet twenty-first century needs?*, in C. HARVEY, J. SUMMERS, N.D. WHITE (a cura di), *Contemporary challenges to the laws of war: essays in honour of Professor Peter Rowe*, Cambridge, 2014, 226.

<sup>109</sup> HEINRICH BÖLL FOUNDATION (a cura di), *op. cit.*, 25.





livello internazionale ed insegnato alle macchine<sup>110</sup>: tuttavia, sembra difficile immaginare che un dispositivo per riprodurre tale segnale possa essere fornito ad ogni soldato – e, qualora pure fosse possibile, risulta improbabile che milizie e gruppi armati irregolari ne fornirebbero in dotazione a tutti i propri combattenti<sup>111</sup>. Alternativamente, sono stati indicati alcuni ambienti in cui le modalità convenzionali di dichiarazione della resa ne rendono maggiormente praticabile il riconoscimento da parte di sistemi dotati di AI<sup>112</sup>. Anche con tali correttivi, tuttavia, l'ambito di lecita utilizzabilità dei LAWS risulterebbe assai ristretto. Per di più, aperta rimarrebbe la questione relativa agli *hors de combat*: un combattente ferito in maniera troppo grave per continuare sarebbe costretto a segnalare la propria resa per essere riconosciuto come tale da un LAWS? E se le sue condizioni non lo permettessero?

#### 4. Questioni tecniche e questioni «ontologiche»

I principi fondamentali di diritto umanitario sembrano imporre, per le operazioni di *targeting*, valutazioni qualitative e richiedenti capacità di comprensione del contesto: valutazioni, pertanto, per le quali la mente umana sembra particolarmente adatta. Ciò, tuttavia, non implica incompatibilità *per se* dei LAWS con il diritto vigente. Per quanto, allo stato attuale, essi non appaiano in grado di conformarsi agli obblighi accennati, non sembra potersi escludere in radice che il progresso tecnologico – potenzialmente mediante l'adozione di diverse componenti di base, anche di natura biologica, ovvero di innovative tecniche di apprendimento automatico – possa, in futuro, condurre ad un risultato differente<sup>113</sup>. Questo, soprattutto, in quanto il diritto umanitario non richiede – né potrebbe richiedere – assoluta accuratezza nell'adempimento dei suddetti obblighi, bensì impone uno standard di diligenza umano. L'argomento dell'avanzamento tecnologico esponenziale potrebbe indurre ad affermare che, qualora il raggiungimento del livello di *performance* umana sia possibile, esso *inevitabilmente*, in futuro, si verificherà<sup>114</sup>. Certo, legittimo sarebbe chiedersi se, nell'ambito della ricerca, vi

<sup>110</sup> R. SPARROW, *Twenty Seconds to Comply: Autonomous Weapon Systems and the Recognition of Surrender*, in *International Law Studies*, 91, 2015, 715.

<sup>111</sup> *Ibid.*

<sup>112</sup> È stato riportato l'esempio rappresentato dalla guerra sottomarina, in cui una parte può segnalare la propria capitolazione facendo affiorare il mezzo ed alzando bandiera bianca mentre l'equipaggio sale sul ponte o abbandona il sottomarino su scialuppe di salvataggio; analogamente, innalzamento della bandiera bianca ed abbandono del mezzo implicano una resa in scontri tra veicoli armati, cfr. *ibid.*, 717-718.

<sup>113</sup> D. AMOROSO, G. TAMBURRINI, *The Ethical and Legal Case Against Autonomy in Weapons Systems*, cit., 20.

<sup>114</sup> L'argomento postula, difatti, che dalla rapidità ed estensione dell'avanzamento tecnologico nell'età moderna consegue che ogni risultato astrattamente possibile sarà inevitabilmente raggiunto, come riporta M. BODEN, *op. cit.*, 136-137. Tale tesi poggia sulla c.d. Legge di Moore, formulata nel 1965, la quale statuisce che la complessità dei microcircuiti di un computer raddoppia con una periodicità che spazia fra i 18 ed i 24 mesi e, di conseguenza, che tutte le loro caratteristiche e capacità aumentino in maniera esponenziale. Per quanto tempo ancora essa continuerà ad essere confermata rimane una questione aperta, su cui la letteratura appare divisa. Alcuni autori sostengono che essa resterà valida ancora per pochi anni, cfr. M. MITCHELL WALDROP, M., *The chips are down for Moore's Law*, in *Nature*, 530, 2016, 145; J. SHALF, *HPC Interconnects at the End of Moore's Law*, IEEE Optical Fiber Communications Conference and Exhibition (OFC), 2019. Peraltro, è stato argomentato che, se pure si avvicinasse la «fine» per la Legge di Moore, i progressi e gli investimenti nel campo dell'intelligenza artificiale porteranno ad un aumento della capacità computazionale tale da superare persino la Legge di Moore stessa, cfr. A. TURCHIN, *Assessing the future plausibility of future catastrophically dangerous AI*, in *Futures*, 107, 1, 2019, 45.

sia interesse (e finanziamenti) per la costruzione di sistemi in grado di conformarsi agli obblighi di diritto umanitario<sup>115</sup>, ma ciò esula dall'oggetto di questo lavoro. Tuttavia, anche in caso di produzione di dispositivi di tal fatta, sembrerebbero permanere alcune questioni, per così dire, ontologiche, cui tale argomento non potrebbe essere esteso e che pertanto rappresenterebbero un ostacolo invalicabile per i LAWS. Tali questioni saranno oggetto dei prossimi paragrafi.

#### 4.1. Responsabilità: ipotesi ed interrogativi

Principi di giustizia insiti alla nostra tradizione giuridica impongono sia fornita la possibilità di contestare le decisioni prese da un sistema autonomo e di porvi rimedio, ovvero di ottenere ristoro<sup>116</sup>. Ciò, in considerazione anche dei beni giuridici che possono essere intaccati in un conflitto, appare fondamentale in caso di violazioni del diritto umanitario. Non rappresenta una opzione accettabile l'impiego di un'arma senza che sia definita una catena di responsabilità<sup>117</sup>. Tuttavia, la natura dell'AI sembra poter scardinare il nostro tradizionale impianto di responsabilità. Esso 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, deve adoperarsi per evitare quello stesso danno. L'imprevedibilità e l'autonomia dei sistemi dotati di AI diluiscono tale controllo, rendendo inadeguato siffatto modello. Inoltre, si può registrare il problema c.d. delle *many hands*: la complessità di tali sistemi e del loro processo di progettazione e sviluppo rischiano di portare ad una situazione in cui numerosi soggetti possano essere considerati responsabili cumulativamente, ma nessuno individualmente<sup>118</sup>. Per tali ragioni, l'impiego di LAWS potrebbe comportare, in caso di violazioni non discendenti da difettosità, errori umani o utilizzo scorretto degli stessi, pertanto da elementi patologici, (inaccettabili) vuoti di responsabilità.

In dottrina, numerose ipotesi sono state avanzate, allo scopo di fornire una possibile soluzione alla questione. Fermo che il punto richiederebbe una elaborazione approfondita *ad hoc*, che qui non si può offrire, risulta necessario delineare i tratti fondamentali delle principali. Una prima risposta potrebbe rintracciarsi nella dottrina della responsabilità di comando, secondo la quale dovrebbe rispondere delle infrazioni commesse da un LAWS l'ufficiale che ne ha ordinato l'impiego. D'altronde, la responsabilità del comandante è prevista per violazioni commesse da soldati umani, soggetti per definizione autonomi: per analogia, potrebbe imporsi anche nel caso di trasgressioni ad opera di LAWS<sup>119</sup>. Al comandante si richiederebbe di capire come tali sistemi presumibilmente funzionino (essendo sufficiente una consapevolezza in merito alle potenzialità di tali armi ed ai loro attendibili *output*) e di prendere decisioni di conseguenza<sup>120</sup>. Sembra, a chi scrive, che la responsabilità di comando possa effettivamente rappresentare la strada più opportunamente percorribile: pare indiscutibile che un ufficiale debba avere cognizione delle forze a propria disposizione; inoltre, l'iscrizione a suo carico della responsabilità potrebbe comportare maggiore cautela nell'impiego di LAWS, determinando un

<sup>115</sup> D. AMOROSO, *Jus in bello and jus ad bellum arguments against autonomy in weapons systems*, cit.

<sup>116</sup> INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE SET UP BY THE EUROPEAN COMMISSION, *Ethics Guidelines for Trustworthy AI*, Bruxelles, 2019, 13.

<sup>117</sup> N.E. SHARKEY, *op. cit.*, 791.

<sup>118</sup> D. AMOROSO, G. TAMBURRINI, *The Ethical and Legal Case Against Autonomy in Weapons Systems*, cit., 15.

<sup>119</sup> C. HEYNS, *op. cit.*, para. 78.

<sup>120</sup> M. SASSÒLI, *op. cit.*, 324.

effetto di deterrenza dal loro uso in situazioni ad alto rischio. Tuttavia, il diritto umanitario, ai fini dell'imposizione della menzionata forma di responsabilità, richiede la conoscibilità dell'illecito e la mancata attuazione di misure di prevenzione e/o repressione, nonché di punizione del colpevole<sup>121</sup>. Non avendo – di norma – approfondite conoscenze di programmazione, un comandante potrebbe trovarsi in una situazione di oggettiva difficoltà nel tentativo di prevedere il comportamento di un'AI<sup>122</sup>. Si potrebbe ribattere che anche il comportamento umano può rivelarsi di complessa previsione, e, ciò nonostante, un comandante può dover rispondere dei crimini commessi dai sottoposti. Tuttavia, la differenza risiede nel requisito della punibilità del colpevole, il quale, in caso di violazioni commesse da LAWS, non potrebbe essere integrato, ferma l'impossibilità (o, meglio, la supposta inutilità) dell'imposizione di una sanzione su una macchina<sup>123</sup>.

Ulteriore ipotesi suggerisce di volgere lo sguardo verso la «fonte» della condotta di un'AI, imponendo una responsabilità sul soggetto che ha programmato e sviluppato il software. Tuttavia, l'imputazione del comportamento di un LAWS al programmatore non impedirebbe l'utilizzo dell'imprevedibilità di tale sistema quale argomento liberatorio, quando esso è dotato di funzioni di apprendimento. Per di più, codici ed algoritmi sono tendenzialmente scritti da squadre di sviluppatori ed individuare i singoli colpevoli potrebbe rivelarsi complesso<sup>124</sup>. L'ascrizione 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 di produzione); implicherebbe, ulteriormente, in considerazione del grado di rischio e del valore dei beni giuridici a rischio, un disincentivo con ogni probabilità insuperabile alla produzione di LAWS<sup>125</sup>. Inoltre, la responsabilità del produttore rappresenta un istituto a carattere civilistico e pare ben poco opportuno imporre l'onere di agire in giudizio sulle vittime, appartenenti tipicamente a popolazioni impoverite dalla guerra<sup>126</sup>.

La difficoltà di individuare un soggetto responsabile e la questione delle *many hands* ha portato 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<sup>127</sup>, tale dottrina predica, per sommi capi, che, in presenza di violazioni perpetrate da più individui nella cornice di un programma criminoso comune, tutti gli individui possono essere reputati *principal perpetrator*<sup>128</sup>. Tale modello, tuttavia, nel richiedere un intento comune, non sembra potersi applicare al caso in oggetto, in cui alla maggior parte, se non a tutti, dei soggetti coinvolti si potrebbe al più ascrivere un dolo eventuale, insufficiente *de iure condito* per una imputazione dinanzi ad un tribunale internazionale<sup>129</sup>. Tale considerazione ha condotto a guardare anche alla responsabilità dello Stato quale possibile

<sup>121</sup> *I Protocollo Aggiuntivo*, art. 86(2), 87(3).

<sup>122</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 43.

<sup>123</sup> HEINRICH BÖLL FOUNDATION (a cura di), *op. cit.*, 30.

<sup>124</sup> G. SOLIS, *op. cit.*, 545.

<sup>125</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 43.

<sup>126</sup> J.M. BEARD, *op. cit.*, 660.

<sup>127</sup> ICTY, *Prosecutor v Tadić*, Appeals Chamber Judgment, del 15 luglio 1999, para. 220.

<sup>128</sup> D. AMOROSO, B. GIORDANO, *Who is to Blame for Autonomous Weapons Systems' Misdoings?*, in E. CARPANELLI, N. LAZZERINI (a cura di), *Use and Misuse of New Technologies in International and EU Law*, Cham, 2019, 220.

<sup>129</sup> Emblematicamente, in tal senso, lo Statuto della Corte Penale Internazionale, che richiede, ai fini della responsabilità criminale, «intent and knowledge», cfr. INTERNATIONAL CRIMINAL COURT (ICC), *Rome Statute of the International Criminal Court*, art. 30.

soluzione, essendo in tale regime assente – ovvero meno rigoroso<sup>130</sup> – l'elemento della *mens rea*. Tuttavia, una soluzione analoga appare al più parziale. Come è stato giustamente argomentato<sup>131</sup>, non sembra, a chi scrive, che la responsabilità statale possa integralmente surrogare quella individuale: istanze di prevenzione generale rimarrebbero, in questo modo, sostanzialmente inascoltate. Inoltre, la responsabilità dello Stato non potrebbe applicarsi ai casi di violazioni commesse da LAWS impiegati da gruppi armati non-statali – i quali, statisticamente, rappresentano i soggetti più coinvolti nei conflitti dell'età contemporanea.

Di fronte a tali complessità, si è tentato di raccogliere i suggerimenti derivanti dalla dottrina che si interroga del rapporto tra AI e diritto, secondo cui una soluzione generale potrebbe essere costituita dall'adozione di un regime di responsabilità oggettiva<sup>132</sup>. Anche con specifico riferimento ai LAWS, taluni hanno consigliato questo modello, in quanto il loro utilizzo potrebbe ben essere assimilato all'esercizio di attività pericolose e tale regime potrebbe favorire un loro impiego cauto e discriminato<sup>133</sup>. Tuttavia, sembra necessario – per quanto non certo semplice – individuare una via mediana tra rischio di impunità e ricerca di capri espiatori. L'adozione di un regime di responsabilità oggettiva potrebbe, difatti, comportare una deviazione significativa da principi di colpevolezza che rivestono un ruolo fondamentale nella nostra tradizione giuridica. Taluni hanno allora suggerito che la strada potrebbe essere quella dell'imputazione della responsabilità alle macchine stesse. Per quanto paradossale possa apparire, questo potrebbe trovare una base giuridica nell'ipotesi di uno *status* specifico per i sistemi di AI<sup>134</sup>. La personalità giuridica di diritto comune è una nozione artificiale: storicamente, è stata negata ad intere categorie di persone fisiche (schiavi) ed è stata attribuita ad enti di ben altra natura (si pensi, ad esempio, al fiume Whanganui, in Nuova Zelanda)<sup>135</sup>. Da un punto di vista di stretta tecnica giuridica, l'imposizione di una responsabilità sui LAWS appare pertanto praticabile, ma occorre interrogarsi in merito alla sua opportunità. La risposta, sotto questo aspetto, pare dover essere negativa: gli obblighi di diritto umanitario sono rivolti ad umani, non 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*<sup>136</sup>. È forte, inoltre, il rischio che tale regime fornirebbe agli uomini la possibilità di usare i LAWS quali «scudi» dalla propria responsabilità<sup>137</sup>. Ferma la mera parzialità di soluzioni alternative (quali l'istituzione di un regime di assicurazione obbligatoria), il quadro giuridico attuale non pare idoneo a rispondere alle questioni di responsabilità sollevate dai LAWS.

<sup>130</sup> D. AMOROSO, B. GIORDANO, *op. cit.*, 224-225.

<sup>131</sup> D. AMOROSO, G. TAMBURRINI, *The Ethical and Legal Case Against Autonomy in Weapons Systems*, cit., 17.

<sup>132</sup> EUROPEAN PARLIAMENT, *European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics*, 2017, para. 53-55.

<sup>133</sup> E. FUZAYLOVA, *op. cit.*, 1359-1362.

<sup>134</sup> EUROPEAN PARLIAMENT, *European Parliament resolution of 16 February 2017*, cit., para. 59, letter f.

<sup>135</sup> B. BROŽEK, M. JAKUBIEC, *On the legal responsibility of autonomous machines*, in *Artificial Intelligence and Law*, 25, 3, 2017, 295.

<sup>136</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 45. Il mantenimento della responsabilità sugli umani è stato affermato quale principio guida anche nel corso degli incontri tenutisi negli ultimi anni a Ginevra, cfr. GGE ON LAWS, *Draft Report of the 2019 session of the Group of Governmental Experts*, cit., Annex IV, letter b.

<sup>137</sup> J.J. BRYSON, M.E. DIAMANTIS, T.D. GRANT, *Of, for, and by the people: the legal lacuna of synthetic persons*, in *Artificial Intelligence and Law*, 25, 3, 2017, 285-287.

#### 4.2. Decisione algoritmica e dignità umana

La seconda questione «ontologica» è relativa al rapporto tra LAWS e dignità umana. Essa appare un valore piuttosto sfuggente, potendo il suo contenuto concretizzarsi in molteplici e, non di rado, persino opposte accezioni, il che tendenzialmente ne rende difficile una interpretazione univoca<sup>138</sup>. Sembra incontroverso, tuttavia, il *minimum* di significato da ricondurre a tale concetto, costituito dall'affermazione secondo cui ogni essere umano va trattato come individuo unico ed irripetibile, come fine, mai come mezzo. È stato in proposito affermato che «contrarie alla dignità umana sembrano essere tecnologie che manipolano l'utente - anche a fine di bene - o a cui sono delegate decisioni di grande importanza sociale o esistenziale senza che sia possibile comprenderne le dinamiche»<sup>139</sup>: ciò sembra particolarmente calzante per il caso dei LAWS, cui sarebbero rimesse *life-or-death decisions*<sup>140</sup>. Due sono gli argomenti principali portati da chi sostiene che la delegazione della scelta degli obiettivi a tali sistemi sarebbe contraria alla dignità umana. Da un lato, viene prospettato un dovere dell'agente di compiere una scelta non arbitraria, ossia informata e consapevole<sup>141</sup>: essa potrebbe derivare solo da un giudizio umano, in quanto solo esso può garantire il pieno apprezzamento del valore di una vita e del significato della sua perdita<sup>142</sup>. In secondo luogo, è stata affermata la sussistenza di un diritto, in capo al soggetto che costituisce il bersaglio di un attacco, di fare appello alla umanità del nemico per avere salva la vita<sup>143</sup> – cosa che gli sarebbe certamente impedita in caso di attacco lanciato da un LAWS.

Le argomentazioni sopra riportate sembrano avere connotazione spiccatamente morale. Una questione necessaria è se la contrarietà della dignità umana all'uccisione di un uomo da parte di una macchina possa discendere (anche) da una norma giuridica. La risposta può apparire affermativa con riferimento ad alcuni specifici ordinamenti – si pensi a quello tedesco, la cui Costituzione si apre con la statuizione per cui «die Würde des Menschen ist unantastbar»<sup>144</sup>. Nel diritto internazionale, sembra difettare una disposizione altrettanto chiara ed univoca sul punto. Potrebbe menzionarsi la statuizione della Corte Internazionale di Giustizia secondo cui «the right not *arbitrarily* to be deprived of one's life» si applica alle ostilità<sup>145</sup>; tuttavia, prosegue la Corte, la determinazione di cosa si possa considerare una arbitraria deprivazione della vita può essere effettuata solo sulla base del diritto applicabile e, nel diritto umanitario positivo, non sembra riscontrarsi una norma che avvalori

<sup>138</sup> C. CASONATO, *Introduzione al Biodiritto*, Torino, 2012, 81-84.

<sup>139</sup> FONDAZIONE LEONARDO-CIVILTÀ DELLE MACCHINE, *Statuto Etico e Giuridico dell'IA*, 2019, 31.

<sup>140</sup> Ciò sembra costituire una risposta a chi sostiene che l'affermazione della contrarietà dei LAWS alla dignità umana dovrebbe implicare eguale illiceità di qualsiasi arma (ad esempio ordigni esplosivi e missili da crociera) che impedisca un «rapporto interpersonale» tra combattente e vittime, non permettendo all'operatore di sapere precisamente chi sarà ucciso, cfr. M. BALISTRERI, *Robot Killer. La rivoluzione robotica nella guerra e le questioni morali*, in *Etica & Politica – Ethics & Politics*, XIX, 2, 426; M. SASSÒLI, *op. cit.*, 324. La differenza fondamentale sembra, tuttavia, risiedere nella identità del soggetto (o oggetto) cui è demandata la concreta decisione.

<sup>141</sup> P. ASARO, *On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making*, in *International Review of the Red Cross*, 94, 886, 2012, 689.

<sup>142</sup> HEINRICH BÖLL FOUNDATION (a cura di), *op. cit.*, 32.

<sup>143</sup> D. AMOROSO, G. TAMBURRINI, *The Ethical and Legal Case Against Autonomy in Weapons Systems*, cit., 19.

<sup>144</sup> *Grundgesetz für die Bundesrepublik Deutschland*, 23. Mai 1949, art. 1(1).

<sup>145</sup> ICJ, *Legality of the Threat or Use of Nuclear Weapons*, cit., para. 25.



l'argomento sopra riportato. Risulta impossibile non richiamare, ciò nondimeno, la c.d. *Martens Clause*. Introdotta dalle Convenzioni dell'Aia del 1899 e ribadita dal I Protocollo Aggiuntivo, la Clausola sottolinea che, in assenza di norme di diritto convenzionale, civili e combattenti rimangono sotto la protezione dei principi di diritto internazionale derivanti da consuetudine, «principi di umanità» e «dettami di pubblica coscienza»<sup>146</sup>. La *Martens Clause* è stata definita un anello tra considerazioni etiche e diritto umanitario<sup>147</sup>. Per quanto qui più interessa, alcuni autori hanno sostenuto che essa possa costituire la base giuridica per il bando di armamenti<sup>148</sup>. Sembrano tuttavia difettare argomenti, tanto nella prassi statale, quanto nell'elaborazione giurisprudenziale internazionale, che possano supportare tale tesi. In generale, parte consistente della letteratura nega rilievo precipuo e diretta applicabilità della *Martens Clause* nelle operazioni di selezione degli obiettivi e nella valutazione della liceità di nuove armi<sup>149</sup>. Anche allineandosi a tale orientamento, tuttavia, ciò che non sembra possibile negare è il fatto che il rispetto della dignità umana sia annoverato tra i «principi di umanità» citati nella Clausola. Pertanto, essa potrebbe costituire la «porta» attraverso cui le argomentazioni di cui sopra potrebbero entrare nell'impianto di diritto umanitario, fornendo una base giuridica alle stesse. Ciò sembra particolarmente importante in quanto, se pure in futuro le caratteristiche tecniche dei LAWS migliorassero a tal punto da rendere loro possibile la conformità ai principi fondamentali di diritto umanitario, l'argomento relativo alla dignità umana – il quale non potrebbe estendersi, peraltro, oltre gli utilizzi *letali* dei LAWS<sup>150</sup> – renderebbe comunque inaccettabile la collocazione dell'uomo *out-of-the-loop*. In ogni caso, va sottolineato, l'assunto concernente il rapporto tra dignità umana e *Martens Clause* sembra rappresentare, allo stato attuale, al più una prospettiva evolutiva del diritto umanitario; pur componente il diritto consuetudinario, la Clausola appare eccessivamente indeterminata ed aperta ad interpretazione per assicurare un fondamento positivo alle argomentazioni sopra accennate. Ciò nondimeno, esse appaiono esprimere un sentimento comune a non pochi commentatori. In tal senso, è stato segnalato, la funzione della *Martens Clause* potrebbe consistere nel fornire un «appiglio» positivo alle argomentazioni in favore della necessità di una regolazione *ad hoc*

<sup>146</sup> I Protocollo Aggiuntivo, art. 1(2).

<sup>147</sup> N. DAVIDSON, *A legal perspective: Autonomous Weapon Systems under International Humanitarian Law*, United Nations Office for Disarmament Affairs (UNODA) Occasional Papers, 30, 2017, 8.

<sup>148</sup> S.J. RUSSELL, *Take a stand on AI weapons*, in *Nature*, 521, 2015, 415. È stato inoltre argomentato che l'impiego della *Clause* quale base giuridica per l'imposizione di un divieto abbia già un precedente nel bando (preventivo), risalente al 1998, dei laser accecanti, cfr. D. AKERSON, *op. cit.*, in D. SAXON (a cura di), *op. cit.*, 92-96. Peraltro, non sembra inconcepibile che il bando in questione sia stato motivato dal divieto di sofferenze non necessarie (possibilità che pure l'autore aveva escluso, affermando che i laser accecanti, rispetto ad altre armi, avrebbero potuto comportare una riduzione delle vittime), violato dall'inflizione di una disabilità permanente, ovvero da ragioni di opportunità o comunque di diverso segno.

<sup>149</sup> Tali argomentazioni sono rivolte in particolare contro il principio di umanità, di cui la *Martens Clause* costituisce, si ritiene, estrinsecazione: C.M. CHINKIN, M. KALDOR, *op. cit.*, 259; G. SOLIS, *op. cit.*, 307-309; Y. DINSTEIN, *The Principle of Proportionality*, in G.D. COOPER, K.M. LARSEN, G. NYSTUEN, (a cura di), *Searching for a 'Principle of Humanity' in International Humanitarian Law*, Cambridge, 2013, 73.

<sup>150</sup> Solo in questi casi, difatti, il bene vita sarebbe minacciato dall'impiego dei LAWS. Peraltro, è stato segnalato come la distinzione tra armi letali e non-letali sia, nei fatti, labile e come sarebbe più corretto riservare per queste ultime la denominazione di armi *sub-letali*, in quanto l'utilizzo di ogni tipo di armamento o munizione comporta, seppur con diverse proporzioni, il rischio di uccisione dell'obiettivo di un attacco, cfr. R. SPARROW, *op. cit.*, 715.



per i LAWS<sup>151</sup>. In questa ottica, rimarchevole appare che la formulazione della Clausola sia stata riportata nel *report* conclusivo della sessione di incontri in seno alla CCW dell'agosto 2019<sup>152</sup>.

## 5. Una possibile soluzione? La prospettiva del controllo umano significativo

Nel dibattito relativo alla compatibilità dei LAWS con il diritto umanitario sempre maggiore rilevanza sembra assumere la nozione, emersa nel corso degli incontri tra rappresentanti nazionali presso la UN CCW, di controllo umano significativo (*meaningful human control*, o MHC). Tale espressione indica, genericamente, il livello minimo di supervisione umana da esercitarsi sui LAWS<sup>153</sup>. Il concetto di MHC appare meritevole di interesse in primo luogo per la sua semplicità linguistica, che lo rende comprensibile ed accessibile anche a non tecnici<sup>154</sup>; inoltre, per il costituire una formula *costruttivamente* ambigua<sup>155</sup> e per il fatto che non solo il settore civile, ma anche quello militare sembra fondato sulla nozione di controllo – e ciò potrebbe rendere contestabile, nonostante i vantaggi operativi che essi sembrano offrire, una sua *completa* cessione ai LAWS<sup>156</sup>.

L'imposizione di un MHC potrebbe rappresentare la chiave per la soluzione di numerose questioni precedentemente illustrate. Per quanto attiene alla dignità umana, se un uomo rimanesse *significativamente* in controllo, le operazioni letali non potrebbero dirsi compiute da un soggetto incapace di comprendere il valore intrinseco della vita. L'incapacità, quantomeno allo stato attuale, dei LAWS di conformarsi ai principi fondamentali di diritto umanitario è in larga parte motivata dal loro difetto di fattori (percezione di contesto e rilevanza, buon senso) che connotano il giudizio umano: il mantenimento di un MHC comporterebbe l'implementazione di tali elementi. Ciò non significa, evidentemente, che gli obblighi da tali principi discendenti non potrebbero, in tal caso, essere comunque violati: tuttavia, un controllo umano concreto ed effettivo potrebbe assicurare la responsabilità del soggetto che si è reso colpevole di tali atti o avrebbe dovuto prevenirli<sup>157</sup>.

Ferme le potenzialità di tale nozione, una questione aperta è se sia possibile rintracciare una base giuridica per la stessa. Sul punto, ampia sembra la disparità di vedute, tanto che alcuni hanno suggerito una sua rilevanza al solo livello politico<sup>158</sup>, mentre altri si sono spinti ad affermare che il MHC costituisca *già* un requisito imposto dal diritto internazionale umanitario<sup>159</sup>. Tale asserzione, secondo

<sup>151</sup> D. AMOROSO, *Jus in bello and jus ad bellum arguments against autonomy in weapons systems*, cit.; D. MAURI, "Laws...of Humanity?". *Un resoconto (critico) del primo gruppo di esperti sulle armi autonome*, in *SIDIBlog*, <http://www.sidiblog.org/2018/01/12/laws-of-humanity-un-resoconto-critico-del-primo-gruppo-di-esperti-sulle-armi-autonome/> (ultima consultazione: 12/01/2018).

<sup>152</sup> GGE ON LAWS, *Draft Report of the 2019 session of the Group of Governmental Experts*, cit., para. 17, letter b.

<sup>153</sup> R. MOYES, H. ROFF, *Meaningful Human Control, Artificial Intelligence and Autonomous Weapons*, Briefing paper prepared for the Informal Meeting of Experts on LAWS, UN CCW, Ginevra, 2016, 1.

<sup>154</sup> K. NESLAGE, *Does "Meaningful Human Control" Have Potential for the Regulation of Autonomous Weapon Systems?*, in *University of Miami National Security & Armed Conflict Law Review*, VI, 1, 176.

<sup>155</sup> D. AMOROSO, G. TAMBURRINI, *What makes human control over weapons systems "meaningful"?*, cit., 5.

<sup>156</sup> I. BODE, H. HUELSS, *op. cit.*, 411.

<sup>157</sup> T. CHENGETA, *op. cit.*, 863-864.

<sup>158</sup> T. MARAUHN, *Meaningful Human Control – and the Politics of International Law*, in R. FRAU, W. HEINTSCHEL VON HEINEGG, T. SINGER (a cura di), *op. cit.*, 210.

<sup>159</sup> J.M. BEARD, *op. cit.*, 665-666; N. DAVIDSON, *op. cit.*, 11; ICRC, *Views of the ICRC on autonomous weapon systems*, cit., 5-6.



questa parte di letteratura, troverebbe diretta implicazione negli obblighi stesse di diritto umanitario, che imporrebbero indirettamente un giudizio di umana ragionevolezza – e proprio per questa ragione rappresenterebbero sfide tanto complesse per i LAWS<sup>160</sup>. Tuttavia, le regole relative a distinzione, proporzionalità, precauzione e necessità militare, per loro natura generali ed astratte, impongono uno standard di diligenza umano, ma non prescrivono in via generale un controllo umano<sup>161</sup>. Tale elemento appare necessario allo stato attuale, ma non sembra preclusa la possibilità di innovazioni tecniche che rendano i LAWS in astratto conformi ai principi fondamentali, sia pure in assenza di un MHC<sup>162</sup>. Si potrebbe dire, in sostanza, che quelli discendenti dal diritto umanitario siano obblighi di risultato, non di mezzi. È stato affermato che l'esempio fornito dalla citata (in nota) Convenzione di Ottawa potrebbe costituire un argomento in favore della tesi che afferma la sussistenza di un requisito di MHC nel diritto umanitario, avendo tale Trattato sancito il bando per le mine antiuomo *quando non controllate a distanza*: questo, si è sostenuto, suggerirebbe la necessità di una supervisione umana sulle armi<sup>163</sup>. Tuttavia, ciò non viene esplicitato: tale Convenzione proibisce semplicemente l'impiego di una categoria specifica di armamenti, non di qualsiasi arma non soggetta a controllo umano.

Il fatto che il MHC non appaia un requisito imposto dal diritto umanitario non esclude la valenza che esso può rivestire nel dibattito in corso. Ciò sembra testimoniato dal ruolo centrale che a tale concetto è stato attribuito dai partecipanti ai citati incontri presso la UN CCW<sup>164</sup> – o quantomeno dalla «vasta maggioranza» degli Stati che hanno, allo stato attuale, espresso la propria posizione sul tema<sup>165</sup>. Più di 30 rappresentanze nazionali – oltre a numerose ONG – si sono dichiarate favorevoli all'introduzione di un vincolo di controllo umano *significativo* per i LAWS; altrettante (non sempre le medesime) hanno predicato il bando di tali armamenti; inoltre, anche buona parte delle delegazioni che non si sono allineate ai poli del dibattito hanno ciò nondimeno espresso la necessità di mantenere l'uso della forza nella sfera di controllo umana e di non delegare *life-or-death decisions* alle macchine<sup>166</sup>. A tali posizioni si contrappongono quelle di un gruppo di Stati – tendenzialmente quelli che più sembrano aver investito nello sviluppo di LAWS<sup>167</sup> – secondo i quali l'impianto di diritto umanitario attualmente esistente non necessiterebbe di integrazioni per disciplinarli efficacemente e l'imposizione di restrizioni ai LAWS costituirebbe una misura «prematura», miope di fronte ai poten-

<sup>160</sup> D. AMOROSO, G. TAMBURRINI, *What makes human control over weapons systems "meaningful"?*, cit., 7.

<sup>161</sup> K. NESLAGE, *op. cit.*, 170-171.

<sup>162</sup> D. AMOROSO, *Jus in bello and jus ad bellum arguments against autonomy in weapons systems*, cit.

<sup>163</sup> HUMAN RIGHTS WATCH, *Killer Robots and the concept of Meaningful Human Control*, 2016, 10-11.

<sup>164</sup> Il concetto di controllo umano è stato in tale sede ritenuto «un elemento imprescindibile per le discussioni future e, su un piano più tecnico, per lo sviluppo di» armi dotate di funzioni autonome, secondo quanto riporta D. MAURI, *op. cit.*

<sup>165</sup> B. STAUFFER, *Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control*, Human Rights Watch, <https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and> (ultima consultazione: 10/08/2020).

<sup>166</sup> *Ibid.* Nel report citato è possibile ripercorrere, in forma di sommario, le posizioni espresse dalle rappresentanze statali che si sono pronunciate sul tema fino all'incontro presso la UN CCW di agosto 2020.

<sup>167</sup> R. ACHESON, *Human control for human rights*, in *Reaching Critical Will*, 6, 8, CCW Report, del 29 agosto 2018, 2.





ziali benefici che questi potrebbero apportare<sup>168</sup>. Una posizione critica è stata assunta anche dalla Polonia, che, pur riconoscendo l'esigenza del mantenimento di un controllo umano sull'uso della forza, ha specificato che ciò va contenuto «to the extent possible in a non-disruptive way for the mission conducted by the system» e che «excessively intrusive human control can disturb mission effectiveness or its completion»<sup>169</sup>; mentre la Finlandia ha segnalato come, ferma la necessità di garantire ad un operatore umano la possibilità di annullare una operazione, questa potrebbe essere limitata e potrebbe sempre raggiungersi un «punto di non ritorno»<sup>170</sup>.

Tali annotazioni mal sembrano coniugarsi con i dettami di diritto umanitario, quantomeno finché, come in precedenza sostenuto, i LAWS, in mancanza di supervisione umana, non potranno fornire garanzie in termini di responsabilità e rispetto dei principi di distinzione, proporzionalità, precauzione e necessità militare. In ogni caso, le obiezioni cennate appaiono rendere decisamente implausibile rintracciare una base giuridica per il requisito di MHC nell'avvenuta o contingente emersione di una norma di diritto consuetudinario *ad hoc* – e, se pure ciò fosse ipotizzabile, una norma analoga avente quali *persistent objectors* potenze militari del calibro di Stati Uniti, Regno Unito e Russia difficilmente potrebbe dirsi risolutiva. Più certa ed efficace – nonché, ad opinione di chi scrive, auspicabile – sembra allora la soluzione rappresentata dall'affermazione di una norma di tal fatta per via pattizia. Questa pare la via tratteggiata anche dai 320 partecipanti (rappresentanti 63 Stati parte della CCW) al Forum di Berlino di aprile 2020, tenutosi in via telematica a causa dell'emergenza sanitaria incorsa. I lavori in seno al Forum rimarcano come il diritto umanitario non fornisca tutte le risposte necessarie alle questioni aperte dai LAWS e ciò enfatizzi la necessità di «internationally-agreed limits»<sup>171</sup>. Vie dirette, in tal senso, sono state individuate dai partecipanti al Forum di Berlino nei principi guida indicati dalle Conclusioni della sessione di lavori presso la UN CCW di agosto 2019, in particolare quelli riguardanti il ruolo umano nell'uso della forza: secondo essi, «human responsibility for decisions on the use of lethal force must be retained»<sup>172</sup>; l'operabilità di LAWS è possibile solo «within a responsible chain of human command and control»<sup>173</sup>; e l'interazione uomo-macchina deve garantire «that the potential use of weapons systems [...] is in compliance with applicable international law, in particular IHL»<sup>174</sup>. Quest'ultima statuizione, nello specifico, è stata indicata dai partecipanti al Forum di Berlino come potenziale «pietra angolare» per il prosieguo dei lavori in sede internazionale sul tema<sup>175</sup>.

<sup>168</sup> D. MAURI, *op. cit.*

<sup>169</sup> *Operationalization of the 11 Guiding Principles of Lethal Autonomous Weapons Systems, Poland's remarks, 2020.*

<sup>170</sup> *Considerations on the appropriate level of human involvement in LAWS, LAWS Group of Governmental Experts of the High Contracting Parties to the Convention on Certain Conventional Weapons, Food-for-thought paper by Finland, June 2020.*

<sup>171</sup> GGE ON LAWS, *Berlin Forum for Supporting the 2020 Group of Governmental Experts on Lethal Autonomous Weapons Systems*, Submitted by Germany, 25 June 2020, para. 17.

<sup>172</sup> GGE ON LAWS, *Draft Report of the 2019 session of the Group of Governmental Experts, cit.*, Annex IV, letter b.

<sup>173</sup> *Ibid.*, Annex IV, letter d.

<sup>174</sup> *Ibid.*, Annex IV, letter c.

<sup>175</sup> GGE ON LAWS, *Berlin Forum, cit.*, para. 22





Il disposto dei principi guida riportati sopra e delle posizioni espresse dalle delegazioni nazionali in tema di LAWS<sup>176</sup> sembra inequivoco nello stabilire che una forma di controllo umano su armamenti dotati di autonomia debba necessariamente mantenersi. Ciò che appare, per il momento, aver impedito a tale requisito di cristallizzarsi in un testo vincolante è il disaccordo in merito a *quale* forma esso debba rivestire – e, pertanto, al suo grado di incisività. In ragione di quanto si è argomentato in precedenza, tanto con riferimento alle questioni che si sono definite «tecniche», quanto a quelle etichettate come «ontologiche», il controllo umano, al fine di garantire il rispetto delle norme di diritto umanitario, sembra dover essere particolarmente qualificato (tornando alla terminologia che sembra aver maggior fortuna nel dialogo internazionale, *meaningful*). Peraltro, questa formulazione rischia di ridurre tale elemento ad un concetto tanto fluido, tanto aperto ad operazioni ermeneutiche di stampo differente, da fornire ben pochi punti fermi all'interprete e persino da ostacolare la stipulazione di accordi che lo abbiano ad oggetto. Si manifesta, pertanto, l'esigenza di «riempire di contenuto» la formula di MHC. Si tratta, indubbiamente, di un obiettivo decisamente ambizioso, eccedente gli scopi di questo lavoro. In ogni caso, al fine di definire quantomeno un *minimum* del concetto di MHC necessario per la conformità dei LAWS al diritto umanitario, è necessario in primo luogo interrogarsi quanto alle fasi di vita dell'arma in cui tale elemento deve intervenire. Un controllo umano sembra patentemente presente nelle fasi di sviluppo, programmazione e dispiegamento di un sistema dotato di AI – salve prospettive futuristiche coinvolgenti dispositivi progettati, costruiti e schierati sul campo da altre macchine. La questione è se il controllo umano in queste fasi possa rivelarsi sufficiente. Questione non del tutto retorica, in quanto è stato affermato che, ai fini dell'integrazione del requisito, potrebbe rivelarsi bastevole l'intervento umano nella fase di programmazione, ove questa delimiti rigidamente l'operabilità dei LAWS all'interno di condizioni e parametri rigorosi che non possano essere alterati dal sistema stesso<sup>177</sup>. Tuttavia, non sembra potersi accogliere tale interpretazione: l'imposizione di un MHC sulle fasi antecedenti alle ostilità, pianificazione dell'attacco inclusa, non appare sufficiente a causa dell'alto grado di imprevedibilità del campo di battaglia, che molto difficilmente, e solo in determinati casi, consentirebbe ai LAWS di mantenersi rigidamente all'interno di parametri predeterminati<sup>178</sup>. Non va inoltre dimenticato come molti obblighi di diritto umanitario si concretino nella fase di selezione degli obiettivi e la possibilità di adempiervi discenda da comprensione situazionale e capacità di adattamento a rapidi mutamenti<sup>179</sup>. Si aggiunga poi che numerosi obblighi di diritto umanitario appaiono ben difficilmente riducibili a codici di programmazione, se non al prezzo di una loro eccessiva semplificazione<sup>180</sup>. Per tali ragioni, un controllo umano non potrebbe dirsi adeguato a garantire il rispetto delle norme di diritto internazionale – e *meaningful* – se non fosse esteso alle operazioni di *targeting*.

<sup>176</sup> Per le quali si rinvia nuovamente a B. STAUFFER, *op. cit.*

<sup>177</sup> Tale, ad esempio, la posizione espressa dai Paesi Bassi, cfr. KINGDOM OF THE NETHERLANDS, *Statement of the Netherlands on Human Machine Interaction*, 2018 GGE on LAWS, 11 aprile 2018, ribadita in KINGDOM OF THE NETHERLANDS, *Statement of the Netherlands delivered at Group of Governmental Experts on LAWS by Ms. Sandra de Jongh*, Agenda Item 5(b), Ginevra, 26 aprile 2019.

<sup>178</sup> T. CHENGETA, *op. cit.*, 872.

<sup>179</sup> K. NESLAGE, *op. cit.*, 165.

<sup>180</sup> HEINRICH BÖLL FOUNDATION (a cura di), *op. cit.*, 26.





Quanto alle forme e modalità di questo elemento, in generale il concetto di controllo non appare poter prescindere dalla possibilità di un intervento correttivo in caso di violazioni o imprevisti. Un monitoraggio che non consenta un coinvolgimento diretto sarebbe, inevitabilmente, inefficace, se non radicalmente inutile. Per evitare ciò, l'operatore umano che vigila su un LAWS dovrebbe sempre avere un potere di veto sulle scelte della macchina, la quale dovrebbe sottomettersi alle direttive dello stesso. Questo potere dovrebbe essere garantito non solo astrattamente, ma soprattutto in concreto. Ciò implica innanzitutto che il controllore dovrebbe disporre del tempo materiale necessario per ricevere le informazioni in tempo reale dal sistema, comprenderne le intenzioni e, se del caso, compiere una operazione di *override*: questo pure al «prezzo» del parziale venir meno dei vantaggi operativi legati alla maggiore velocità di reazione di un sistema autonomo, che non possono prevalere sugli altri beni giuridici interessati dalle operazioni militari<sup>181</sup>. Essenziale sarebbe, in tale ottica, la garanzia della piena affidabilità del collegamento fra uomo e macchina: in caso di interruzione temporanea dello stesso, si potrebbe ipotizzare di forzare il sistema ad uno *stand-by*, o quantomeno a non assumere nuovi obiettivi. L'effettività del controllo nel caso concreto discende anche dalla consapevolezza, nell'operatore e in chi prenda la decisione di impiegare un LAWS, del grado di avanzamento tecnologico e dei limiti del sistema in oggetto, nonché degli *attendibili* esiti derivanti dal suo utilizzo: ferma l'imprevedibilità di un dispositivo dotato di capacità di apprendimento, essa dovrebbe essere ridotta il più possibile per agevolare il compito dell'operatore, che dovrà avere coscienza del contesto di impiego e poter ricevere *feedback* e risposte intellegibili dall'AI<sup>182</sup>. A tale scopo, saranno indubbiamente necessari investimenti tesi alla configurazione di sistemi trasparenti ed attendibili (e quindi nei campi di *explicable* e *trustworthy AI*). Inoltre, un controllo effettivo non può prescindere dalla considerazione dell'effetto che le deliberazioni di sistemi dotati di AI potranno avere sui nostri processi decisionali<sup>183</sup>. Emblematico, in tal senso, appare il fenomeno del c.d. *automation bias*: esso è stato definito quale la tendenza umana a fidarsi di (ed affidarsi a) un sistema autonomo nonostante evidenze mostrino che esso non sia attendibile ovvero sia, nel caso di specie, fallace<sup>184</sup>. Così, la decisione dell'operatore, che pure avrebbe in astratto la possibilità di esercitare un controllo e partecipare attivamente alle operazioni, potrebbe dunque essere «catturata» da quella della macchina, ritenuta, a ragione o meno, più neutra ed attendibile. Ciò anche per motivazioni di pragmatica convenienza: in primo luogo, dimostrare che una decisione sia da attribuirsi esclusivamente ad un sistema di AI, e che l'uomo addetto al controllo si sia limitato ad avallarla acriticamente, risulta poco meno di una *probatio diabolica*; l'adesione a tale determinazione può contribuire ad affievolire l'onere di motivazione del soggetto (onere il quale, invece, dovrebbe con ogni probabilità essere rafforzato nel caso opposto) e permette allo stesso di «qualificare» la propria decisione<sup>185</sup>. Sembra evidente che il fenomeno dell'*automation bias* non potrà essere sottovalutato da una prospettiva che promuova istanze di MHC.

<sup>181</sup> È stato, in merito, argomentato che «there should be no hurry for humans to kill each other», cfr. T. CHENGETA, *op. cit.*, 885.

<sup>182</sup> IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS, *op. cit.*, 125-129.

<sup>183</sup> M.A.C. EKELHOF, *Lifting the fog of targeting: "Autonomous Weapons" and Human Control through the Lens of Military Targeting*, in *Naval War College Review*, 71, 3, 2018, 85.

<sup>184</sup> HUMAN RIGHTS WATCH, *Losing our Humanity*, cit., 13.

<sup>185</sup> A. SIMONCINI, *L'algoritmo incostituzionale*, cit., 81.





Nel tentativo di «delineare il campo» per il controllo umano, una precisazione appare essenziale: l'esigenza di un MHC non potrebbe mai prevalere, in un bilanciamento, sui diritti umani all'integrità fisica ed alla vita. La proliferazione di tecnologie connotate da crescente autonomia e le tendenze delle guerre asimmetriche dell'età contemporanea potrebbero rendere il campo di battaglia un contesto troppo rapido e caotico per il processo decisionale umano: l'imposizione di un MHC *in ogni caso* sui LAWS potrebbe, in alcuni casi, essere di detrimento all'incolumità degli uomini stessi<sup>186</sup>. Riconoscendo ciò, alcuni autori hanno proposto una valutazione nel caso concreto del grado di autonomia che potrebbe essere consentito a specifici modelli di LAWS<sup>187</sup> – il quale appare maggiore per i sistemi difensivi. Essi sono presenti (ed operativi) sulla scena internazionale già da decenni e sembrano godere di un elevato grado di accettazione, come testimoniato anche in seno alla UN CCW<sup>188</sup>. Ciò anche in ragione del fatto che, quanto meno in via tendenziale, sistemi (meramente) difensivi, ove sufficientemente affidabili, appaiono più verosimilmente in grado di conformarsi ai principi fondamentali di diritto umanitario<sup>189</sup>: i loro obiettivi rappresentano indubbiamente obiettivi militari, come richiesto dal principio di distinzione; pur ferme le difficoltà di operare una valutazione di proporzionalità, la (legittima) difesa da un pericolo imminente può presumersi rappresentare un considerevole vantaggio militare; analogamente, una siffatta azione difensiva non potrebbe mai reputarsi *non necessaria*, alla luce del principio di necessità militare; infine, per quanto attiene al principio di precauzione, la ristrettezza dei tempi a disposizione e l'urgenza della situazione rende tipicamente impraticabile la messa in atto di misure preventive<sup>190</sup>. Una valutazione nel caso concreto sembra, peraltro, a propria volta maggiormente confacente al diritto umanitario, stante che l'art. 36 del I Protocollo Aggiuntivo prescrive l'obbligo di una *legal review* su ogni singola nuova arma, mentre i LAWS costituiscono una categoria generale<sup>191</sup>. L'elaborazione sopra riportata appare pertanto pienamente meritevole all'interno di una prospettiva sistematica; tuttavia, sembrano necessarie alcune considerazioni di carattere pragmatico. Un giudizio nel caso concreto rischia di cedere il fianco ad una incertezza applicativa che potrebbe trasformarsi in arbitrio. In ragione delle istanze di segretezza che caratterizzano il processo di sviluppo e produzione di nuove armi<sup>192</sup>, il giudizio sarebbe demandato, con ogni probabilità, a soggetti nazionali, presumibilmente già incaricati del compito di svolgere la valutazione di compatibilità con le norme di diritto internazionale applicabili ex art. 36. Ciò non potrebbe che comportare rischi connessi alla nascita di standard di valutazione divergenti, nonché di statuizioni «interessate» e concilianti. Inoltre, come alcuni autori hanno messo in luce, pochi Stati sembrano di-

<sup>186</sup> W.H. BOOTHBY, *Weapons and the Law of Armed Conflict*, cit., 250.

<sup>187</sup> D. AMOROSO, G. TAMBURRINI, *What makes human control over weapons systems "meaningful"?*, cit., 12-13. Nello specifico, gli autori costruiscono una classificazione su più livelli decrescenti, basata sul grado di controllo umano da imporsi su ogni sistema: gli ultimi tre livelli di tale scala corrispondono, sostanzialmente, ai gradi di autonomia indicati in questo elaborato con le locuzioni «in-the-loop», «on-the-loop» ed «out-of-the-loop».

<sup>188</sup> Ne è testimonianza la posizione, ad esempio, della Spagna, secondo la quale solo i sistemi difensivi rendono ammissibile un controllo attinente al modello «man-on-the-loop», cfr. *Commentaries on national implementation of the guiding principles on LAWS*, Spain, 2020.

<sup>189</sup> D. AKERSON, *op. cit.*, in D. SAXON (a cura di), *op. cit.*, 74.

<sup>190</sup> *Ibid.*

<sup>191</sup> C. GRUT, *op. cit.*, 21.

<sup>192</sup> Y. DINSTEIN, *The conduct of hostilities under the law of international armed conflict*, cit., 100-101; W.H. BOOTHBY, *Weapons and the Law of Armed Conflict*, cit., 345.



sporre effettivamente di meccanismi di *review* adeguati<sup>193</sup>. Per tali ragioni, sembra doversi ribadire l'esigenza, ad opinione di chi scrive, di un obbligo *vincolante* di MHC, stabilito per via pattizia: un MHC rispondente al *minimum* che si è tentato di tratteggiare, pertanto *esteso alla fase di targeting, effettivo, concreto e comprensivo di un potere di veto per LAWS offensivi*. Questo non significa che un controllo umano significativo non dovrebbe essere imposto su sistemi difensivi – affatto – bensì che questi, in casi straordinari di necessità ed urgenza, dovrebbero essere in grado di agire anche quando l'operatore umano non sia presente o non disponga della materiale possibilità di intervenire. Tale soluzione appare di analogo verso a quella emergente dalla lettera aperta pubblicata in occasione dell'apertura dell'*International Joint Conference on Artificial Intelligence (IJCAI)* del 2015, firmata da un lungo elenco di ricercatori ed esperti nei campi di intelligenza artificiale e robotica, la quale richiede un bando per armi autonome *offensive* sprovviste di controllo umano significativo<sup>194</sup>. Né il diritto internazionale è sordo alle istanze che possono essere sollevate da esigenze di autodifesa: così, la decisione – o, si potrebbe affermare, la non-decisione – della Corte Internazionale di Giustizia in merito alla legittimità delle armi nucleari potrebbe costituire, almeno parzialmente, una base giuridica per la soluzione di cui sopra, nella parte in cui i giudici dell'Aia hanno affermato di non poter stabilire definitivamente se la minaccia o l'utilizzo di armi nucleari siano legittimi in «an extreme circumstance of self-defence»<sup>195</sup>.

## 6. Conclusioni

Nell'impianto contingente di diritto umanitario non sembrano potersi riscontrare argomenti *decisivi* che segnalino la radicale incompatibilità dei LAWS – specie delle loro potenziali conformazioni future – con lo stesso. Ciò nondimeno, le questioni relative a responsabilità e dignità umana richiederanno quantomeno degli adattamenti nel diritto applicabile, come sottolineato anche dal Forum di Berlino. Come è stato giustamente osservato, tali modifiche potrebbero incontrare riluttanze persino maggiori da parte degli Stati: ad esempio, questo potrebbe avvenire in caso di mutamenti che tendano verso un allentamento dei requisiti per la responsabilità degli stessi o dei loro soldati<sup>196</sup>. Tale considerazione sembra rappresentare una ulteriore spinta nella direzione della necessità dell'imposizione di un MHC, di cui si sono tentati di delineare gli elementi essenziali. Si tratta, naturalmente, di una soluzione innovativa: non esistono, nel diritto umanitario, norme che impongano il mantenimento di un controllo umano sugli armamenti. Ciò, tuttavia, è motivato dal fatto che la questione stessa è completamente innovativa: mai, nella storia umana, si è presentata la prospettiva di un'arma che possa autonomamente scegliere i propri obiettivi. Il problema, inveterato ed endemico, è quello della differenza di velocità tra sviluppo tecnologico e adattamento del diritto. Nuove questioni invocano nuove soluzioni e il MHC appare la più appropriata alle domande aperte dai LAWS. Oltretutto, esso si pone nel solco del filone ermeneutico che pone la propria attenzione, nell'ottica del rapporto fra AI e diritto

<sup>193</sup> HEINRICH BÖLL FOUNDATION (a cura di), *op. cit.*, 25; C. GRUT, *op. cit.*, 9.

<sup>194</sup> FUTURE OF LIFE INSTITUTE, *Autonomous Weapons: an Open Letter from AI & Robotics Researchers*, Buenos Aires, del 28 luglio 2015.

<sup>195</sup> ICJ, *Legality of the Threat or Use of Nuclear Weapons*, cit., para. 105.

<sup>196</sup> D. AMOROSO, *Jus in bello and jus ad bellum arguments against autonomy in weapons systems*, cit.



to, sull'emersione di un diritto individuale ad una decisione non completamente automatizzata<sup>197</sup>. Nel corso delle operazioni militari, le decisioni da prendersi riguardano il bene giuridico vita: pertanto, l'elaborazione su questo tema appare particolarmente centrale per il diritto umanitario. Essa potrebbe fornire spunti rilevanti al fine dell'imposizione di un MHC sui LAWS – e, al contempo, da questa eventualità trarre ulteriore linfa.

L'AI sembra poter ricevere molteplici utilizzi nel contesto del conflitto armato e le questioni sollevate dai LAWS non devono portare ad ignorare i potenziali effetti benefici della stessa. L'AI potrebbe innanzitutto fornire un apporto fondamentale al settore della difesa, non solo, come accennato, su un piano strettamente fisico, ma anche a livello dematerializzato (in particolare, nella protezione dalle minacce evocate dalla frontiera del *cyber warfare*)<sup>198</sup>. Essa potrebbe offrire un supporto estremamente valido al settore militare nel raccogliere, immagazzinare ed analizzare dati di *intelligence*, nonché nella conseguente attività di pianificazione strategica<sup>199</sup>. Ulteriori impieghi suggeriti spaziano dallo svolgimento di operazioni di sorveglianza alla consegna di viveri e scorte, dal disinnescamento di esplosivi all'estrazione di combattenti feriti<sup>200</sup>. Inoltre, fermo quanto si è cercato di argomentare, l'AI potrebbe rappresentare un supporto al processo decisionale umano anche nelle operazioni di selezione degli obiettivi<sup>201</sup>: in tal caso, dovranno essere poste in essere misure adeguate per evitare il verificarsi di fenomeni di *automation bias* o analoghi. Allo scopo, sembra essenziale un adeguato lavoro di formazione del personale militare che si ritroverà a collaborare con LAWS o altre applicazioni dell'AI ed a controllarne l'operato. Comandanti e combattenti dovranno essere coscienti di interagire con macchine e dovranno essere consapevoli tanto delle potenzialità, quanto dei possibili difetti dei sistemi a loro disposizione, nonché delle questioni giuridiche che essi sollevano. Diretta implicazione di ciò è, inevitabilmente, il venir meno (o quantomeno l'affievolirsi) di alcune delle promesse dei LAWS, in particolare quelle relative a riduzione di carico di lavoro umano e costi. Affinché il modello auspicato possa funzionare, sarà inoltre necessario investire sul raggiungimento di una AI affidabile ed *explainable*, così come sul rafforzamento delle procedure di *legal review* ex art. 36, al fine di assicurare l'effettiva significatività del controllo umano. Ciò sembra suggerito anche da argomenti di carattere geopolitico: all'interno di un contesto globale sempre più polarizzato, nel quale la guerra si caratterizza per una asimmetria crescente fra schieramenti più e meno avanzati dal punto di vista tecnologi-

<sup>197</sup> Esso può essere indicato quale il diritto a non essere sottoposti a decisioni basate unicamente su un trattamento automatizzato. In assenza di un controllo umano effettivo, tale diritto sarebbe rispettato soltanto su un piano puramente formale, in quanto la determinazione finale risulterebbe solo formalmente risultante dall'interazione fra uomo e macchina. La principale formulazione del diritto ad una decisione non completamente automatizzata (o anche diritto alla non-esclusività della decisione algoritmica) si rinviene in *Regulation of the European Union 2016/679, Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, c.d. General Data Protection Regulation*, del 27 aprile 2016, art. 22(1). Tale statuizione è limitata, peraltro, da alcune eccezioni, tra cui spicca quella relativa al «consenso esplicito dell'interessato». Per un'elaborazione sul tema, si rinvia a C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., 127-129; A. SIMONCINI, *L'algoritmo incostituzionale*, cit., 79-84.

<sup>198</sup> THE HAGUE CENTRE FOR STRATEGIC STUDIES (a cura di), *Artificial Intelligence and the Future of Defense. Strategic Implications for small- and medium-sized force providers*, L'Aia, 2017, 67.

<sup>199</sup> *ibid*, 91-95.

<sup>200</sup> D. AKERSON, *op. cit.*, in D. SAXON (a cura di), *op. cit.*, 68.

<sup>201</sup> M.A.C. EKELHOF, *op. cit.*, 86.



co, le sfide al costituzionalismo presentate dai nuovi poteri economici che emergono nel mondo dei *big data* includono certamente anche quella di un controllo sul potere militare che può derivare da una supremazia nel campo dell'AI. L'imposizione di un controllo umano non rappresenterebbe, si ritiene, un ostacolo alla ricerca scientifica nel settore. Come la maggior parte di chimici e biologi non hanno interesse nello sviluppo di armi chimiche o biologiche, e ne hanno largamente supportato il bando, analogamente la maggior parte dei ricercatori nel campo dell'AI non sembra mostrare interesse nella produzione di armi dotate di crescente autonomia – anzi, sembra paventare, in questo caso, il rischio di una indesiderabile reazione negativa dell'opinione pubblica nei confronti dell'AI stessa<sup>202</sup>.

*IS  
&  
Law*

---

<sup>202</sup> FUTURE OF LIFE INSTITUTE, *op. cit.*

