

## Intervista al Prof. Guglielmo Tamburrini

Il settore militare ha visto l'impiego di diversi tipi di sistemi automatici, con un deciso aumento della loro produzione e impiego almeno a partire dalla Prima Guerra Mondiale. Ma i recenti sviluppi dell'Intelligenza Artificiale (IA) e della robotica, congiuntamente applicate allo sviluppo di sistemi d'arma, hanno sollevato problemi etici e legali nuovi e particolari, che riguardano la capacità di un sistema d'arma di agire autonomamente, selezionando e attaccando un obiettivo senza dover necessariamente passare per l'autorizzazione preventiva di un'essere umano. L'autonomia di un sistema d'arma nelle funzioni critiche di selezione e attacco di un obiettivo militare ha generato preoccupazioni nella pubblica opinione, ma non solo.

Il Parlamento europeo si è espresso contro lo sviluppo di questa forma di autonomia nei sistemi d'arma. Un gruppo di ONG ha lanciato nel 2013 una campagna per la messa al bando delle armi autonome (la "Stop Killer Robots Campaign"). Migliaia di ricercatori nei settori dell'IA e della robotica si sono apertamente schierati contro lo sviluppo e l'impiego delle armi autonome. Nella sede ONU di Ginevra si discute dal 2014 della compatibilità delle armi autonome con i principi delle leggi umanitarie in guerra nel quadro diplomatico della CCW (Convention on Conventional Weapons). Nell'ambito della CCW un gruppo di esperti internazionali ha ricevuto il mandato di approfondire il problema dell'autonomia con lo scopo di fornire raccomandazioni agli stati membri in merito alla loro ammissibilità in base alle leggi umanitarie in guerra.

Nell'aprile del 2017 il Department of Foreign and Security Policy della Heinrich Böll Stiftung (<https://www.boell.de/en>) – la fondazione tedesca Heinrich Böll con sede a Berlino – ha dato vita a una Task Force sulle armi autonome, formata da esperti di varie discipline tra cui antropologi, informatici e robotici, studiosi di diritto internazionale, filosofi e studiosi di politica. La fondazione Heinrich Böll, che si trova sotto l'egida del partito dei Verdi, è una delle fondazioni di ricerca e studi che la Germania permette ai partiti presenti nel Parlamento tedesco di avere e gestire con il sostegno del finanziamento pubblico.

Obiettivo degli studi condotti dalla Task Force è stato quello di redigere un Rapporto che analizzasse il potenziale dirompente (disruptive) delle applicazioni della IA e della robotica nel settore militare da una prospettiva etica e giuridica. Il Rapporto, intitolato *Autonomy in Weapon Systems. The Military Application of Artificial Intelligence as a Litmus Test for Germany's New Foreign and Security Policy* è stato redatto da Daniele Amoroso, Frank Sauer,

Noel Sharkey, Lucy Suchman e Guglielmo Tamburrini, è uscito nel 2018 in Creative Commons ed è scaricabile da <https://www.boell.de/en/2018/05/17/autonomy-weapon-systems>.

Il Rapporto, rivolto inizialmente alle istituzioni politiche e ai decisori della Germania, delinea e motiva diverse obiezioni a futuri sviluppi delle armi autonome, fa appello affinché il governo intervenga per fermarne la progettazione, lo sviluppo e l'uso, auspicando che le Nazioni Unite arrivino a bandire le armi autonome, come risultato dell'azione congiunta di vari governi e stakeholder.

Il Rapporto sviluppa riflessioni di carattere etico e giuridico sull'impiego delle armi autonome dal punto di vista dei principi dell'International Humanitarian Law (IHL), con particolare riferimento al principio di distinzione tra combattenti e popolazione civile. Sviluppa inoltre i temi della imprevedibilità del comportamento di sistemi autonomi e robotici e della responsabilità finale degli effetti dell'uso di queste armi. Mette in guardia rispetto ai rischi della proliferazione degli armamenti causata dall'applicazione dell'IA alle armi e fa appello ai principi della dignità umana, dell'umanità e della coscienza pubblica rispetto agli effetti di queste armi. Conclude con una serie di raccomandazioni al Governo e al Ministero della Difesa della Germania.

Abbiamo chiesto al prof. Tamburrini di descrivere l'origine della Task Force e le sue attività.

“Il Dr. Frank Sauer della Bundeswehr Universität di Monaco di Baviera ha organizzato e coordinato la Task Force sull'impatto dirompente delle tecnologie dell'IA e della robotica nelle attività belliche, con particolare attenzione all'autonomia dei sistemi d'arma. Compito della Task Force è stato quello di redigere un rapporto con una serie di indicazioni e raccomandazioni da sottoporre all'attenzione delle istituzioni politiche della Germania sulla gestione di questi processi d'innovazione dal punto di vista etico e giuridico.

Un primo problema affrontato dalla Task Force riguarda la definizione di arma autonoma. L'autonomia operativa di un dispositivo tecnologico è infatti un concetto molto ampio. Alcuni potrebbero identificare una mina antiuomo come arma autonoma poiché questo dispositivo è in grado di gestire una situazione in cambiamento e di modificare il proprio comportamento in funzione dei dati sensoriali: infatti, solo se il sensore rileva un peso che eccede una determinata soglia la mina esploderà. In termini più generali, il problema definitorio da affrontare era quello di trovare un punto di equilibrio tra una definizione di autonomia molto liberale, che dal punto di vista politico-normativo sarebbe stata di difficile

applicazione perché troppo generica e inclusiva, e una definizione di autonomia così restrittiva da proiettare l'autonomia dei sistemi d'arma in un futuro indeterminato senza riuscire a cogliere le novità tecnologiche dirompenti dell'IA e della robotica che fin d'ora stanno cambiando l'organizzazione e la conduzione dei conflitti bellici.

Optando per una definizione eccessivamente restrittiva, per esempio, il Ministero della Difesa della Gran Bretagna ha dato una definizione di autonomia che proietta le armi autonome in un futuro lontano. Secondo la loro definizione, i sistemi d'arma autonomi devono in sostanza essere dotati della capacità di comprendere l'ambiente e i suoi mutamenti, di dialogare con gli umani e di interpretare le situazioni al livello degli esseri umani. Queste capacità costituiscono un punto di arrivo ideale per l'intelligenza artificiale, ma se intese come condizioni necessarie per l'autonomia, si proiettano ipso facto per definizione le armi autonome in un futuro talmente lontano da negare l'esistenza di tali armi nell'oggi o nel futuro prossimo. In conclusione, secondo questo approccio, non esiste il problema dell'autonomia delle armi che agita l'opinione pubblica, le discussioni diplomatiche in sede ONU e le discussioni politiche nei parlamenti di numerosi stati democratici.

Ma questo è assurdo. Oggi possiamo già dotare un robot della capacità di attaccare e uccidere un essere umano senza l'autorizzazione preventiva di un essere umano, anche se il robot in questione non ha un'intelligenza di tipo generale paragonabile a quella di un essere umano. Evidentemente la proposta britannica non costituisce un punto di partenza utile per comprendere le modifiche tecnologiche già in atto e per affrontarle da una prospettiva etica e giuridica appropriata.

Una posizione più equilibrata e utile sul problema definitorio è quella assunta dal Comitato Internazionale della Croce Rossa (ICRC) che ha adottato e fornito una definizione funzionale di autonomia dei sistemi d'arma, focalizzando la questione intorno alle funzioni eticamente e giuridicamente critiche per l'autonomia di un sistema d'arma. Per l'ICRC, un sistema d'arma è autonomo se ha la capacità di selezionare e attaccare un obiettivo senza nessun intervento da parte di un essere umano. Per fare ciò non è necessario, in molte circostanze, avere le capacità di comprensione e di intelligenza generale possedute dagli esseri umani, come si richiede nella definizione del Ministero della difesa britannico.

La definizione più inclusiva proposta dall'ICRC deve essere comunque adattata per fornire una base appropriata per la discussione sulla regolamentazione e sull'eventuale messa al bando internazionale di armi autonome. Infatti, non è ragionevole considerare come armi da

mettere al bando quei sistemi di protezione antimissile che sono attualmente in uso e che pure godono delle capacità funzionali proposte dall'ICRC. Tali sistemi agiscono autonomamente con un tempo di reazione molto più basso di quello che potrebbe avere un essere umano nell'individuare, puntare e fermare un'eventuale minaccia rappresentata da missili in arrivo. Sono delle armi autonome a tutti gli effetti, secondo quanto stabilito dal Comitato internazionale della Croce Rossa, ovvero capaci di identificare e attaccare un bersaglio senza essere sottoposti a un controllo umano preventivo. Questi sistemi hanno tuttavia solo uno scopo difensivo e soprattutto rispondono solo a minacce poste da altre armi senza esseri umani a bordo (unmanned).

Dopo aver affrontato la questione definitoria, la Task Force ha preso in considerazione e si è concentrata specificamente su sistemi d'arma in cui la robotica e l'intelligenza artificiale congiunte riescono ad abilitare grandi capacità di tipo percettivo, di riconoscimento di obiettivi militari e nello stesso tempo attribuiscono a queste armi la decisionalità di attaccare autonomamente anche degli esseri umani, senza essere sottoposte al controllo umano, in scenari bellici che normalmente sono molto complessi e di difficile interpretazione.

In guerra, i belligeranti si impegnano nella ricerca della sorpresa e nella modifica delle condizioni ambientali più prevedibili dal nemico. È molto difficile che la robotica e l'intelligenza artificiale attuali riescano a creare dei sistemi sufficientemente flessibili e capaci di riconoscere e valutare in situazioni difficilmente prevedibili di questo genere la proporzionalità dell'attacco e la distinzione civile-combattente, così come è invece richiesto dalle leggi umanitarie in guerra. Inoltre, c'è un esempio molto importante che è stato studiato nel Rapporto, cioè l'integrazione di sistemi robotici dotati di capacità cinetiche e militari con sistemi informatici. L'integrazione di questi vari sistemi può determinare dei comportamenti imprevedibili e incontrollabili, non solo rispetto alle condizioni dell'ambiente esterno ma anche rispetto al comportamento che emerge dalla loro interazione, con particolare riferimento all'interazione con altri sistemi software autonomi che vivono nel mondo non tangibile della rete. Nel caso di azioni imprevedibili di armi autonome che sarebbero giudicate come dei crimini di guerra nel caso in cui esse fossero compiute da un essere umano, chi potrà risponderne o esserne considerato responsabile sotto il profilo etico e legale? Il rapporto punta il dito sull'interruzione della catena di responsabilità militari e politiche che lo sviluppo delle armi autonome può comportare.

A questa considerazione bisogna affiancare il problema dei cyberattacchi e le operazioni di hacking sulle armi autonome. Si tratta di grandi pericoli, già denunciati da ONG e altre agenzie internazionali che monitorano le attività militari, lo sviluppo di nuovi sistemi d'arma e la

protezione dei diritti umani in guerra. In particolare, l'hackeraggio di sistemi autonomi non sottoposti al controllo umano può ritorcersi proprio contro coloro che li hanno dispiegati, portando a perdite di controllo e a casi di fuoco amico da parte di sistemi autonomi

Il rapporto però non riguarda solo l'essere umano, ma affronta anche i pericoli che le armi autonome potrebbero sollevare per la sicurezza globale e la stabilità. Non è solo un problema di scienza politica e di relazioni internazionali. Si tratta di un problema che è anche di pertinenza dell'etica e in particolare della cosiddetta etica delle conseguenze. Bisogna tutelare il più possibile la pace e la stabilità internazionale, così come richiede la carta fondativa dell'ONU, mettendo in atto tutte le misure possibili per contrastare ciò che pone una minaccia alla convivenza e allo sviluppo pacifici dei popoli. Le armi autonome stanno già scatenando una nuova corsa agli armamenti. Esse possono costituire un potente incentivo a iniziare un conflitto in cui non ci si debba più preoccupare di mettere in gioco delle vite umane poiché la guerra sarebbe combattuta da armi autonome robotiche.

Lo sviluppo di sottomarini autonomi pone un rischio significativo alla pace e alla stabilità, anche ai fini della deterrenza nucleare, come è stato messo in luce in un recente rapporto del British Pugwash. Sciami di sottomarini autonomi che pattugliano gli oceani potrebbero presto essere in grado di tracciare gli spostamenti di tutti i sottomarini dotati di missili nucleari, ai quali è oggi affidata la principale forma di risposta in caso di attacco nucleare da parte del nemico. Ma neutralizzare la possibilità di una risposta nucleare a un attacco nucleare significa vanificare il tipo di deterrenza che si basa sulla distruzione reciproca assicurata.

Essere contrari alle armi autonome vuol dire richiedere che ci sia sempre un essere umano a governare le macchine, anche nella decisione di sopprimere un'altra vita umana. La richiesta di imporre un controllo umano significativo su ogni sistema d'arma trova un suo importante fondamento nella protezione della dignità dell'uomo, sia in quanto potenziale paziente di una decisione di vita e di morte, sia in quanto attore responsabile di una tale decisione. La costituzione tedesca pone, con il suo primo articolo, la protezione della dignità dell'essere umano sopra ogni altra cosa. Il filosofo Immanuel Kant considerava l'uomo come portatore nell'universo fisico della responsabilità morale, che a sua volta è fonte della speciale dignità che caratterizza l'essere umano in quanto tale. Il collegamento con l'art. 1 della Costituzione tedesca è stato un punto di ispirazione molto importante per questo Rapporto e per le conseguenti raccomandazioni al governo tedesco.

Come è stato accolto il Rapporto?

Il rapporto è liberamente disponibile sul sito della Fondazione Heinrich Böll (<https://www.boell.de/en/2018/05/17/autonomy-weapon-systems> ). Il rapporto è stato presentato in un evento pubblico a Berlino. L'ICRAC (Comitato internazionale per il controllo delle armi robotiche) lo sta attivamente diffondendo. Vi sono iniziative analoghe che sono state avviate a livello nazionale per sensibilizzare i decisori politici anche nella scia dell'iniziativa della Fondazione Böll. In Italia, la Rete italiana per il disarmo, in collaborazione anche con l'USPID (Unione degli scienziati per il disarmo), ha recentemente organizzato (il 5 luglio 2018) un incontro-dibattito sulle armi autonome e sui problemi etici e giuridici che esse sollevano con la partecipazione di deputati di vari gruppi parlamentari intitolato "Killer robots: futuro obbligato della guerra? Le armi completamente autonome e il loro impatto su diritti e conflitti." Nel complesso, la protesta e le azioni contro lo sviluppo e l'impiego delle armi autonome stanno crescendo. E' da segnalare infine l'opposizione crescente agli usi dell'IA e della robotica per lo sviluppo di strumenti di sorveglianza di massa eventualmente collegati anche allo sviluppo di armi autonome. Nell'aprile del 2018, alcuni ingegneri di Google si sono dimessi in segno di protesta contro la collaborazione fra il Ministero della difesa degli Stati Uniti e Google per utilizzare gli algoritmi di deep learning di Google ai fini del riconoscimento di oggetti di interesse militare da parte dei droni da ricognizione (Project Maven). I dipendenti di Google hanno chiesto e ottenuto che Google non rinnovasse questo accordo di collaborazione con il Ministero della difesa statunitense. Inoltre, con un documento pubblicato sul sito dell'ICRAC (<https://www.icrac.net/open-letter-in-support-of-google-employees-and-tech-workers/> ), più di mille ricercatori nei vari settori delle tecnologie dell'informazione hanno appoggiato la lotta dei dipendenti di Google, sottolineando la relazione di fiducia e le corrispondenti responsabilità che questo colosso dell'informazione ha verso la comunità internazionale dei suoi utenti e verso la protezione dei dati personali di ciascuno di essi.

Il Prof. **Guglielmo Tamburrini** è Professore di logica e filosofia della scienza presso l'Università degli Studi di Napoli "Federico II"

<https://www.docenti.unina.it/guglielmo.tamburrini>

Guglielmo Tamburrini (PhD 1987, Columbia University, New York) è professore ordinario di filosofia della scienza e della tecnologia presso il DIETI, Dipartimento di Ingegneria Elettrica e Tecnologie dell' Informazione dell' Università di Napoli Federico II. I suoi interessi di

ricerca si concentrano sulle problematiche metodologiche delle scienze cognitive e delle neuroscienze computazionali, nonché sulle implicazioni etiche e sociali della robotica e dell' IA. È stato coordinatore del primo progetto europeo sull' etica dei sistemi robotici e dell' IA dal 2005 al 2008 (ETHICBOTS). Nel 2009-10 è stato Visiting Scholar presso lo ZIF (Zentrum für Interdisziplinäre Forschung, Universität Bielefeld, Germania). Nel 2014 gli è stato conferito il Premio Internazionale Giulio Preti del Consiglio Regionale della Toscana per il suo lavoro didattico e di ricerca sulle implicazioni etiche e sociali dell' intelligenza artificiale e delle tecnologie robotiche.