

Incontro di Studio - «Soluzioni giuridiche e pratiche all'uso dei sistemi aerei, marittimi e terrestri automatizzati in campo civile e militare», Macerata, 11 febbraio 2020

Le implicazioni giuridiche dell'autonomia nei sistemi d'arma

Prof. Daniele Amoroso
Università degli studi di Cagliari, ICRC



Alle origini del dibattito sui sistemi d'arma autonomi

“[R]obots not only can be better than soldiers in conducting warfare in certain circumstances, but they also can be more humane in the battlefield than humans”

Ronald Arkin



“Rather than making war more humane and ethical, autonomous armed robotic machines are simply a step too far in the dehumanization of warfare. We must continue to ensure that humans make the moral decisions and maintain direct control of lethal force”

Noel Sharkey





Convention on Certain Conventional Weapons (CCW)
Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)
11-15 April 2016, Geneva

1. Definitions

The ICRC has defined autonomous weapon systems as: "Any weapon system with autonomy in its critical functions. That is, a weapon system that can select (i.e. search for or detect, identify, track, select) and attack (i.e. use force against, neutralize, damage or destroy) targets without human intervention."



Department of Defense DIRECTIVE

NUMBER 3000.09
November 21, 2012

Incorporating Change 1, May 8, 2017

USDX(P)

SUBJECT: Autonomy in Weapon Systems

autonomous weapon system. A weapon system that, once activated, can select and engage targets without further intervention by a human operator.



**CAMPAIGN TO STOP
KILLER ROBOTS**

Urgent Action Needed to Ban Fully Autonomous Weapons

Non-governmental organizations convene to launch Campaign to Stop Killer Robots

(London, April 23, 2013) – Urgent action is needed to pre-emptively ban lethal robot weapons that would be able to select and attack targets without any human intervention, said a new campaign launched in London today. The Campaign to Stop Killer Robots is a coordinated international coalition of non-governmental organizations concerned with the implications of fully autonomous weapons, also called "killer robots."

COS'È UN SISTEMA D'ARMA AUTONOMO?

Un sistema d'arma che – servendosi di tecnologie più o meno avanzate di IA – è in grado, una volta attivato, di svolgere le funzioni "critiche" di selezione ed ingaggio degli obbiettivi senza intervento umano

Tipologie di sistemi d'arma autonomi esistenti...



Sistemi di difesa aerea
Nächstbereichschutzsystem
MANTIS – Germania



Sentinelle robotiche
Super aEgis II – Corea del
Sud



Munizioni *fire-and-forget*
Brimstone – Regno Unito



**Sistemi di
munizionamento *loitering***
Harpy NG – Israele



... e futuri



PERCHÉ LA SELEZIONE E L'INGAGGIO DEGLI OBBIETTIVI SONO CONSIDERATE FUNZIONI «CRITICHE»?

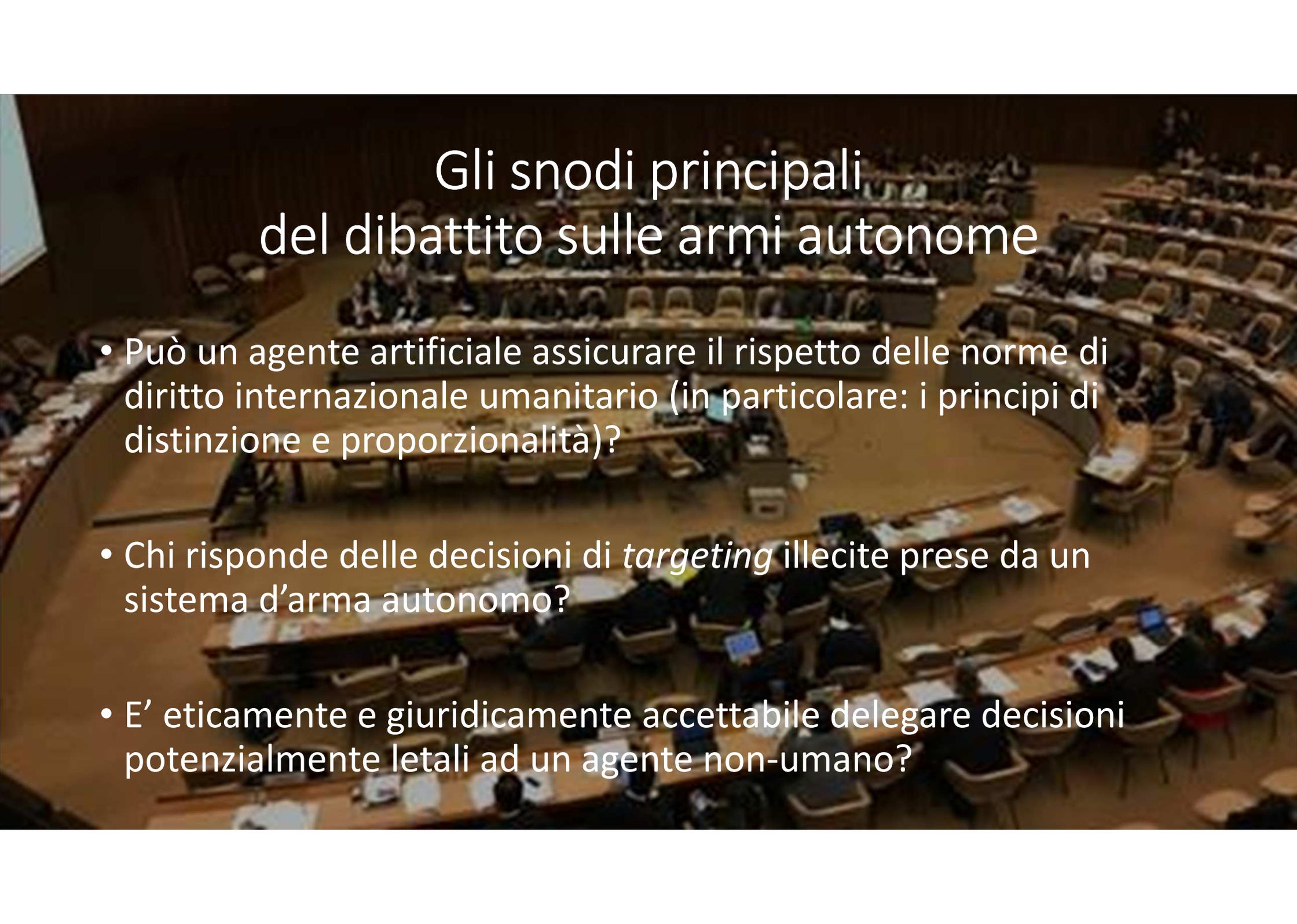
Sono disciplinate da principi-chiave del diritto internazionale umanitario



Il loro esercizio costituisce un fattore decisivo ai fini dell'attribuzione della responsabilità (individuale e statale)



Implicano scelte morali suscettibili di incidere su diritti fondamentali



Gli snodi principali del dibattito sulle armi autonome

- Può un agente artificiale assicurare il rispetto delle norme di diritto internazionale umanitario (in particolare: i principi di distinzione e proporzionalità)?
- Chi risponde delle decisioni di *targeting* illecite prese da un sistema d'arma autonomo?
- E' eticamente e giuridicamente accettabile delegare decisioni potenzialmente letali ad un agente non-umano?

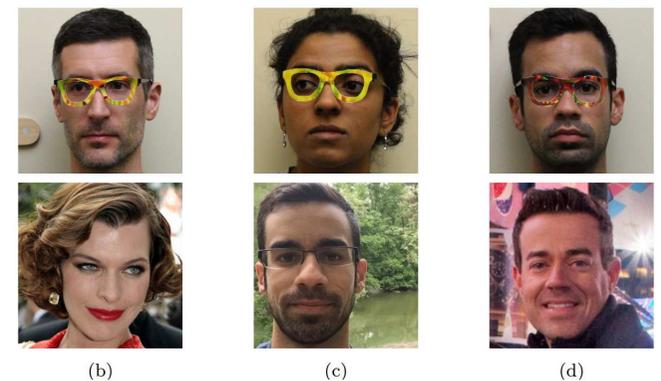
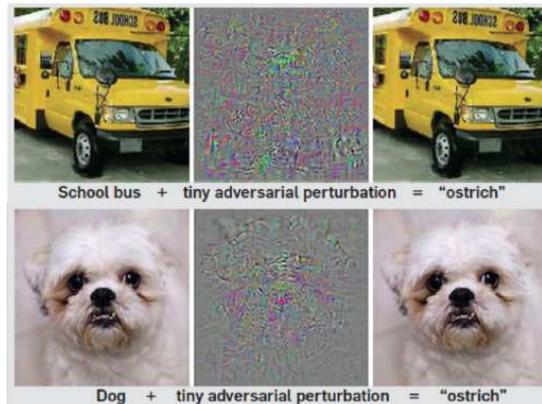
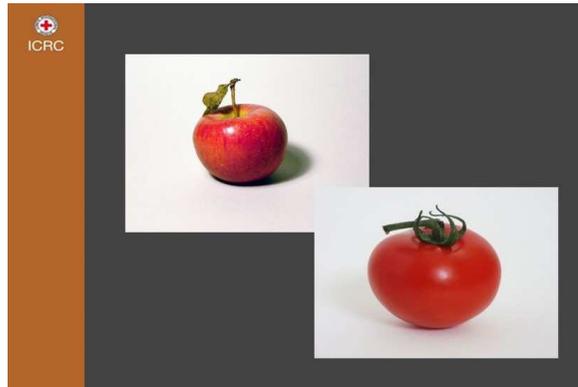


Il principio di distinzione

Persone ed oggetti protetti dal diritto internazionale umanitario

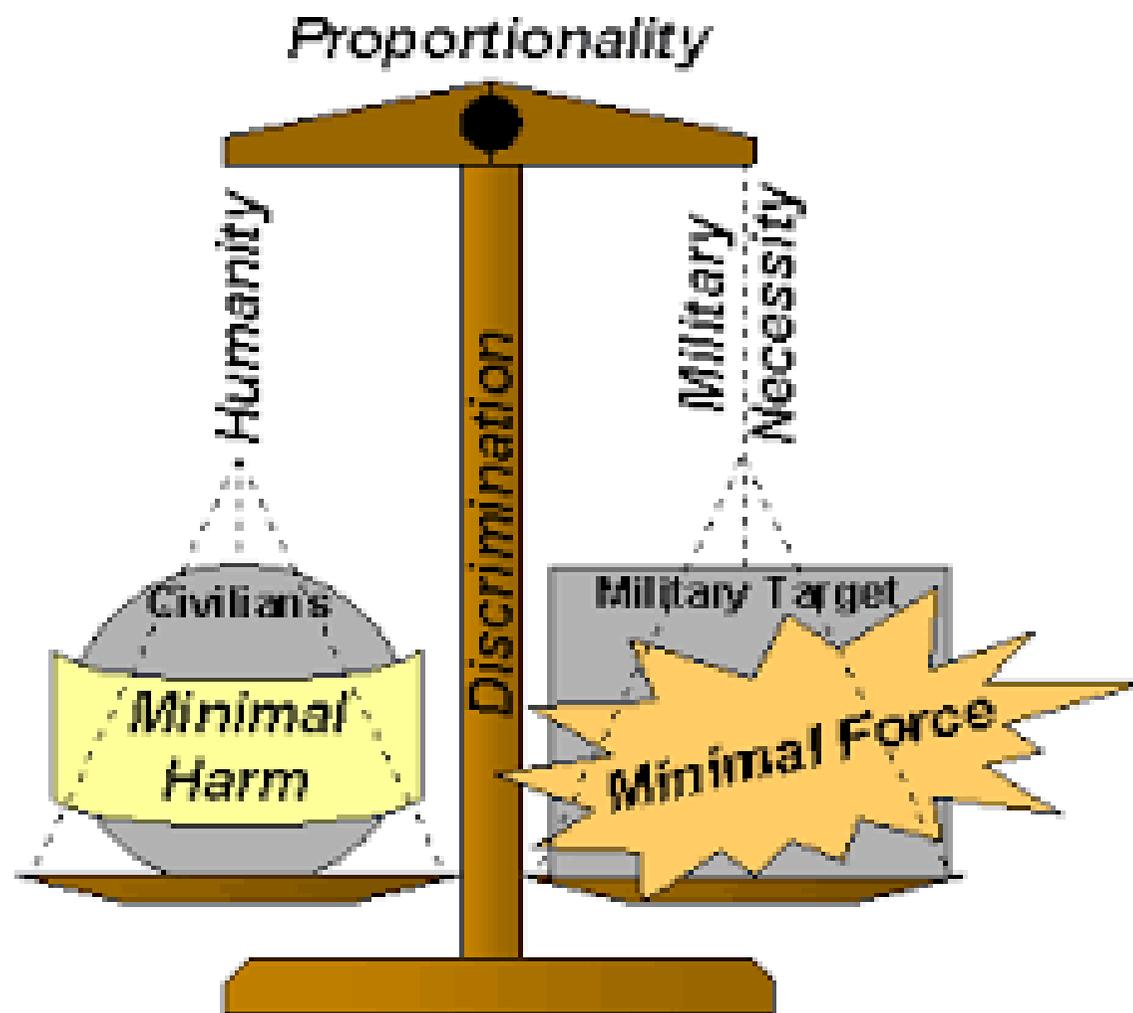
- Civili
...a meno che non partecipino direttamente alle ostilità
- Oggetti civili
...a meno che il loro uso o la loro destinazione li rendano obiettivi militari
- Forze di *peace-keeping*
- Personale militare assegnato in via esclusiva a funzioni religiose o sanitarie
- Nemici *hors de combat*

Il riconoscimento automatico delle immagini e i suoi limiti: il cd. gap semantico

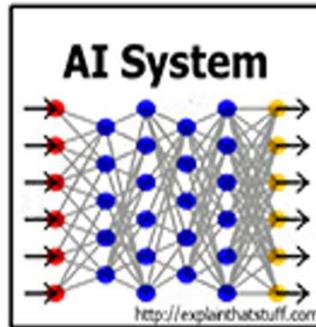


Il principio di proporzionalità

E' proibito lanciare attacchi dai quali ci si può attendere che provochino incidentalmente morti e feriti fra la popolazione civile, danni ai beni di carattere civile, o una combinazione di perdite umane e di danni, che risulterebbero eccessivi rispetto al vantaggio militare concreto e diretto previsto



Il problema del cd. «vuoto di responsabilità»



- We are entering a new age of AI applications
- Machine learning is the core technology
- Machine learning models are opaque, non-intuitive, and difficult for people to understand



- Why did you do that?
- Why not something else?
- When do you succeed?
- When do you fail?
- When can I trust you?
- How do I correct an error?

- I sistemi d'arma autonomi saranno sempre meno prevedibili:
 - *Machine learning*
 - Caratteristiche degli ambienti operativi
- Sarà sempre più difficile incardinare un rapporto di responsabilità normativamente accettabile per difetto della *mens rea*

L'AUTONOMIA (LETALE) DEI SISTEMI D'ARMA COME *MALUM IN SE*

-
- La delega ad un agente artificiale del potere di assumere decisioni letali è inaccettabile sul piano etico...
 - ... e giuridico in virtù del principio della dignità umana
 - Valutazione che prescinde dal rispetto delle norme del diritto internazionale umanitario

Due varianti argomentative

- «Diritto» individuale a non essere sottoposto ad una decisione letale «automatizzata»
- «Obbligo» degli operatori umani, in quanto agenti morali, di sostenere il peso dell'assunzione di decisioni letali

UNA POSSIBILE VIA D'USCITA: LA NOZIONE DI CONTROLLO UMANO SIGNIFICATIVO

- Consenso crescente a livello internazionale intorno all'idea secondo cui i sistemi d'arma dovrebbero essere soggetti ad un controllo umano significativo (CUS)
- Triplice ruolo del controllo umano «significativo»
 - *Meccanismo di salvaguardia* volto ad impedire che il cattivo funzionamento della macchina provochi danni altrimenti evitabili
 - *Catalizzatore di responsabilità* che consenta, in caso di eventi dannosi, di individuare il soggetto responsabile e le ragioni della riprovevolezza della sua condotta
 - «Attivatore» di *agenzia morale*, il quale assicuri che sia un essere umano, e non una macchina, a prendere decisioni riguardanti la vita e l'integrità fisica delle persone coinvolte in un conflitto armato

ELEMENTI DEL CONTROLLO UMANO SIGNIFICATIVO

- Obblighi primari
 - Privilegi di controllo esclusivo per gli operatori umani
 - Approccio differenziato
 - ... una limitata autonomia può essere consentita in presenza di certi limiti operativi (ad es. per i sistemi difensivi)
- Obblighi secondari
 - Necessari ad assicurare la *qualità* del coinvolgimento umano
 - Training
 - Design: trasparenza



GRAZIE PER L'ATTENZIONE!