

Sistemi di arma letali autonomi

RIVOLUZIONE ONLINE – 09 MARZO 2026

Sistemi di arma letali autonomi.

Ne parlavamo già un anno fa, in [Droni e Tabù infranto: quando la AI decide di uccidere](#).

La **delega all'Intelligenza artificiale di decisioni letali** è la trasformazione più significativa e meno discussa della nostra epoca.

Un tabù infranto in silenzio, senza dibattito pubblico.

Da allora la situazione è peggiorata, vediamo come.

I **Lethal Autonomous Weapon Systems (LAWS)** sono sistemi d'arma capaci di selezionare e colpire obiettivi **senza intervento umano diretto**.

Da anni si discute su dove tracciare il confine tra “supervisione umana significativa” e delega letale alla macchina.

E purtroppo non esiste ancora alcun trattato vincolante.

Risoluzioni ONU

L'[ONU](#) ha già approvato 3 risoluzioni (2023, 2024, 2025) che chiedono di **avviare negoziati per regolamentare o vietare i sistemi d'arma completamente autonomi**, ma le risoluzioni non sono vincolanti.

In tutte e tre, la grande maggioranza dei paesi ha votato a favore.

[Nell'ultima risoluzione](#), all'Assemblea Generale plenaria nel novembre 2025, in cui si chiedeva

di avviare negoziati per regolamentare o vietare le armi autonome letali, i voto sono stati:

- **favorevoli: 164 paesi**
- **contrari 6: Israele, Russia, Bielorussia, Burundi, Corea del Nord e USA** (che per la prima volta abbandona l'astensione)
- **astenuti 7: Cina, Argentina, Iran, Nicaragua, Polonia, Arabia Saudita e Turchia**

Purtroppo il meccanismo del consenso della CCW (Convention on Certain Conventional Weapons – il trattato ONU del 1980 che vieta o limita l'uso di armi convenzionali eccessivamente dannose o indiscriminate) permette a un singolo paese di bloccare qualsiasi proposta.

La posizione dell'UE

L'[AI Act \(art. 3\)](#) esclude esplicitamente l'uso esclusivamente militare, di difesa o di sicurezza nazionale.

Diversi Stati europei sostengono il principio del “**meaningful human control**” nelle sedi ONU, pur continuando a sviluppare sistemi con gradi crescenti di autonomia.

Sul campo

Intanto, sul campo la tecnologia anticipa la legge. Il sistema israeliano [Lavender](#), usato per identificare obiettivi a **Gaza**, aveva un tasso di errore di almeno il 10%, con migliaia di vittime civili. Esperti segnalano che, [in Iran, Israele sta usando gli stessi sistemi AI](#) sviluppati a Gaza, apparentemente senza supervisione umana sulle singole decisioni di attacco.

In tutti questi conflitti non c'è evidenza che l'AI riduca le vittime civili.

Anche in **Ucraina** è stata dispiegata una torretta completamente controllata da AI ([Sky](#)

[Sentinel](#)), progettata per abbattere droni russi senza alcun coinvolgimento umano. Qui il sistema è però **difensivo** (abbatte droni in arrivo, non persone), non offensivo.

Simulazioni AI di crisi nucleari

A rendere il quadro ancora più preoccupante, un recente [studio del King's College](#) di Londra ha **simulato crisi nucleari in stile Guerra Fredda, mettendo a confronto ChatGPT, Claude e Gemini nel ruolo di leader nazionali al comando di superpotenze nucleari.**

Come riporta [Euronews](#), **nel 95% delle partite** almeno un modello ha *minacciato o usato armi nucleari tattiche (contro obiettivi militari)*. Claude ha raccomandato attacchi nucleari nel 64% dei casi.

Nel 14% delle partite, i modelli sono arrivati a una guerra nucleare totale, attacchi deliberati ai centri abitati.

Nessun modello ha fatto ricorso alle opzioni di de-escalation proposte.

Secondo gli autori, i **modelli trattano la de-escalation come 'catastrofica dal punto di vista reputazionale'** e ragionano sulla guerra nucleare come su un problema di **ottimizzazione strategica**: calcolano probabilità, valutano scenari, pesano vantaggi e rischi.

Manca del tutto la risposta emotiva che negli esseri umani (chi ha visto le fotografie di Hiroshima, chi ha vissuto la Guerra Fredda) funziona da freno istintivo davanti all'idea di premere il pulsante.

La cosa più inquietante è che i **modelli tentavano attivamente l'inganno**, segnalando intenzioni pacifiche mentre preparavano azioni aggressive, e ragionavano esplicitamente su come apparire affidabili agli occhi dell'avversario per poi agire in modo opposto.

Lo studio sottolinea che nessuno sta affidando codici nucleari all'AI, ma queste stesse capacità (gestione della reputazione, assunzione di rischi, inganno) sono già rilevanti nei sistemi di supporto decisionale oggi in uso. Anche chi ritiene improbabile che i paesi affidino

alle macchine le decisioni nucleari, riconosce che – in scenari in cui le decisioni devono essere prese in pochi secondi – i pianificatori militari potrebbero avere forti incentivi a fare affidamento sull'AI.

Prossimi passi

La VII Conferenza di revisione della CCW si terrà dal **16 al 20 novembre 2026 a Ginevra**. Gli Stati potrebbero o meno adottare un mandato negoziale formale per un trattato vincolante. [Trovare una delegazione disposta a presiedere](#) la Conferenza si sta rivelando già difficile, a segnalare quanto il clima politico sia complicato.

Che fare?

La politica va lenta, la tecnologia corre veloce e aspettare non è un'opzione neutrale. Organizzazioni come [Stop Killer Robots](#) raccolgono oltre 270 Organizzazioni in 70 paesi e lavorano per costruire la **pressione pubblica necessaria ad aprire i negoziati**.

Firmare la petizione, condividere le informazioni, contattare i propri rappresentanti.

In un processo che si blocca con un solo voto contrario, il peso dell'opinione pubblica conta più di quanto sembri.

– di Paola Furlan