Avvenire

diffusione:94524
tiratura:125758

L'INTERVISTA

«Nessuno ha il diritto di scommettere la vita degli altri»

«L'intelligenza artificiale nella difesa è ormai adottata da tutte le grandi potenze e perfino dalle forze di Stati minori. Tutti usano una tecnologia per cui non esistono ancora leggi e norme e che, a ben guardare, viola quelle esistenti. Ne segue che chi la adotta è pronto a ignorare il diritto umanitario internazionale, compresi gli Stati che se ne proclamano paladini. È cominciata un'erosione sistemica di queste leggi e la cosa dovrebbe preoccuparci», a mettere insieme i pezzi è la filosofa Mariarosaria Taddeo, docente di Digital Ethics and Defence Technologies all'Università di Oxford e autrice di "Codice di guerra. Etica dell'intelligenza artificiale nella difesa" (Raffaello Cortina Editore).

In che modo l'IA sta ridefinendo la guerra, le armi e il concetto di plausibile?

La troviamo su tre fronti: il supporto alle funzioni strategiche, con l'intelligence aumentata e la logistica predittiva; i conflitti non cinetici, a partire dal cyberspace; i conflitti cinetici, dove l'integrazione nei sistemi d'arma autonomi pone la domanda più radicale, se sia lecito delegare la forza letale a entità non umane. Sono fattori dirompenti, che cambiano il nostro stesso modo di concepire la guerra e le sue implicazioni etiche e legali. Quanto al plausibile, è il punto che mi interessa di più. In pochi anni la generazione algoritmica di bersagli è passata da scenario ipotetico a routine operativa, e la soglia di ciò che appare normale, perfino inevitabile, si è spostata quasi senza che ce ne accorgessimo. L'hype pesa quanto i fatti. Le analisi indipendenti mostrano che le capacità dimostrate sul campo sono spesso inferiori alle attese, eppure l'adozione continua, in certi casi frettolosa, spinta da promesse di vantaggi decisivi che spesso non si realizzano.

Questo ha diverse derive etiche.

Quelle che mi preoccupano di più sono quattro. La prima è l'imprevedibilità. I sistemi di apprendimento automatico possono produrre esiti che chi li progetta e li impiega non aveva previsto e non poteva prevedere, minando le fondamenta del diritto internazionale e limitando di fatto il controllo umano sulla tecnologia. La seconda è la coppia velocità e scala: la mediazione tecnologica anestetizza il giudizio morale. La terza è l'erosione della responsabilità, uno degli argini (sottilissimi) che impediscono alla guerra di degenerare in atrocità. Se la responsabilità morale per le azioni di sistemi AI non può essere attribuita in partenza, l'adozione di questa tecnologia nella difesa incontra un limite serio. Io credo che la responsabilità morale per le azioni dell'IA possa essere ascritta solo parzialmente. La quarta è l'escalation cibernetica. In un vuoto regolatorio, il conflitto cibernetico abilitato dall'IA può degenerare fino a minacciare la stabilità internazionale.

La filosofa Mariarosaria Taddeo sulle regole per l'uso dell'IA in guerra: «Chi aspetta il trattato perfetto sta regalando tempo a chi non vuole alcun trattato»

In che senso?

Chi progetta o impiega questi sistemi fa una scommessa morale, moral gambit. Si accetta in anticipo di rispondere personalmente anche degli esiti che non si volevano. Come chi firma una cambiale in bianco. In difesa, per sistemi AI usati per funzioni di supporto e non-cinetiche questa scommessa è accettabile. Lo è anche per i sistemi d'arma autonomi non letali. Per un'arma autonoma letale no, perché la posta non è tua. Nessuno ha il diritto di scommettere la vita degli altri.

Quali argini ci sono all'IA "bellica"?

Il diritto internazionale umanitario si applica anche all'IA, sulla carta. Il problema è che è stato scritto per armi che non imparano. Un esempio: l'articolo 36 del primo Protocollo addizionale alle Convenzioni di Ginevra impone di testare le armi prima del loro uso in guerra. Con l'IA l'applicazione dell'articolo vacilla. Questi sistemi imparano e si trasformano imparando. Quelli ad apprendimento continuo sono sistemi diversi alla fine di ogni missione. Applicato alla lettera, l'articolo imporrebbe di ritestarli prima di ogni singolo impiego. Non è fattibile. Progressi ce ne sono stati, ma impercettibili. Penso, ad esempio, alla Ccw (Convenzione su certe armi convenzionali ndr) delle Nazioni Unite, dove il Gruppo di esperti governativi lavora a un testo che copre aspetti come proibizioni e responsabilità, ma viene bloccata dalla regola del consenso unanime. Uno strumento vincolante specifico ancora non esiste. Il risultato è sotto gli occhi di tutti. La scadenza del 2026 fissata da Guterres per un trattato vincolante è passata a luglio.

Come si supera l'impasse attuale?

Chi aspetta il trattato perfetto sta regalando tempo a chi non vuole alcun trattato. Vedo tre direttrici. La prima, "operazionalizzare" l'etica dentro le istituzioni della difesa, subito, applicando in progettazione, sviluppo e impiego i principi di responsabilità, spiegabilità, tracciabilità, affidabilità e governabilità. La seconda, è che Nato e Unione Europea preparino il terreno a un'iniziativa guidata dalle Nazioni Unite. La terza è il pragmatismo definitorio. Meglio convergere su caratterizzazioni funzionali che inseguire definizioni onnicomprensive. L'approccio a due livelli – proibire i sistemi imprevedibili e quelli che bersagliano esseri umani e regolare tutti gli altri – è il punto di convergenza più realistico. Ed è coerente con la conclusione a cui sono giunta da filosofa. La scommessa morale sulle vite altrui non è mai lecita. Tutto il resto si può negoziare, questo no.

Elisa Campisi

© RIPRODUZIONE RISERVATA

La corsa delle armi "intelligenti"

La proprietà intellettuale è riconducibile alla fonte specificata in testa alla pagina. Il ritaglio stampa è da intendersi per uso privato