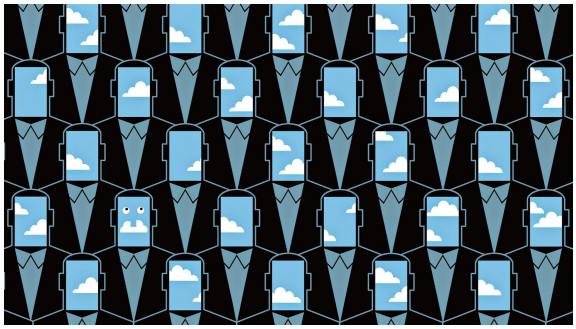
la Repubblica.it
Il mondo in diretta 24 ore su 24

L'IA e il confine tra guerra e barbarie

LINK: https://www.repubblica.it/robinson/2026/09/10/news/l_ia_e_il_confinetra_guerra_e_barbarie-425576110/



L'IA e il confine tra guerra e barbarie di Mariarosaria Taddeo L'opinione pubblica deve intervenire sui rischi delle armi autonome e dell'IA. Ecco la lectio della studiosa di Oxford Ascolta l'articolo L'intelligenza artificiale è entrata nella difesa. Quest'ingresso, rapido e privo di regole, rischia di spezzare la linea sottile che separa la guerra dalla barbarie, perché l'IA ottimizza decisioni e operazioni ma al costo del controllo e della responsabilità umana. È una svolta preoccupante, ancor più perché il dibattito è spesso inquadrato in termini fantascientifici che distraggono dalle sfide concrete che dovremo affrontare nei prossimi giorni. Il repertorio della fantascienza è sempre lo stesso: da Frankenstein a HAL per arrivare a Terminator. Nel caso dell'IA nella difesa, Terminator ha il dito sul bottone nucleare. Da mesi, appelli e prime pagine usano a queste immagini per descrivere i rischi dell'IA nella difesa. È il meccanismo della distrazione: si prende un problema reale, se ne immagina il caso più estremo, anche se non realistico, purché attiri l'attenzione mentre i rischi più concreti si materializzano. Se il meglio è nemico del buono, il peggio è amico del male. Partiamo dal peggio, la deterrenza nucleare guidata dall'IA . La deterrenza è una minaccia condizionata. Uno Stato promette la rappresaglia a chi lo attacca, e la minaccia funziona finché è credibile e controllata. Una risposta partita per errore, o sproporzionata, produce l'escalation, cioè il contrario della deterrenza. Per questo doppie chiavi e catene di autorizzazione fanno partire la risposta se e solo se l'ordine è legittimo, nella misura dovuta. È vero che una rappresaglia automatica rende la minaccia più credibile, perché assicura una risposta anche dopo che il primo attacco è stato lanciato. Per questo l'Unione Sovietica costruì un sistema semiautomatico di risposta, il Perimetr : ma la decisione finale restò agli ufficiali, anche in questo caso. Automatizzare del tutto aggiunge rischio senza aggiungere deterrenza, che gli arsenali da secondo colpo già forniscono. Nel 2024 Washington e Pechino hanno riaffermato che l'uso nucleare deve restare in mani umane, e nel 2025 l'Assemblea generale ha votato la prima risoluzione sull'IA nel comando nucleare, esigendo il controllo umano. Passiamo al male. In Ucraina un drone russo autonomo ha ucciso tre civili a Zaporizhzhia, primo caso documentato. A Gaza, il sistema Gospel identifica obiettivi fisici, 100 al giorno; mentre Lavender, secondo le inchieste del magazine investigativo +972 , ha identificato circa 37 mila persone come target militari in un processo decisionale che lascia agli ufficiali una manciata di secondi per validare le indicazioni del

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

sistema. In Iran , grazie a Maven, gli Stati Uniti hanno identificato e attaccato quasi 900 obiettivi nelle prime 12 ore dell'attacco. Per arrivare fin qui non è servito Terminator. È bastata l'assenza di regole. La risposta di rito dice che il diritto umanitario si applica anche all'IA. È vero, come è vero che la biologia si applica alla medicina. Ma senza protocolli, certificazioni e applicazioni specifiche la biologia non offre cure. L'applicazione del diritto richiede interpretazioni, la definizione di soglie di rischio e di casi limite. Quanto tempo è necessario lasciare agli ufficiali per validare le decisioni dell'IA? Quali livelli di predicibilità dei sistemi IA sono necessari? Tutto questo manca. Le regole andavano scritte a Ginevra, dove la Convenzione su certe armi convenzionali mise il tema in agenda nel 2013 e lo affidò a un Gruppo di esperti governativi. Quattordici anni hanno fruttato undici principi che ribadiscono l'ovvio, una definizione provvisoria di armi autonome, e un testo negoziale la cui premessa è "niente è concordato finché tutto non è concordato". Quattordici anni per concordare che niente è concordato. Altrove farebbe sorridere. Ma a Ginevra, il Gruppo delibera per consenso e un pugno di potenze lo usa come veto permanente. Il 31 agosto si apre l'ultima sessione del Gruppo, che consegnerà il rapporto finale. In autunno la Prima Commissione attende la quarta risoluzione sulle armi autonome , e a novembre la Conferenza di riesame sceglierà tra negoziare un trattato, prorogare la discussione e lasciar cadere il processo. Questo dibattito non può essere interamente demandato alla politica o alla diplomazia . Decidere che cosa è lecito in guerra significa fissare il confine che separa lo stato di diritto dal terrorismo, e la guerra dalla pura barbarie. Per questo nei prossimi mesi l'opinione pubblica deve richiamare a sé un ruolo centrale nel dibattito su come queste tecnologie si usano e si limitano nella difesa. Le distrazioni su questo tema sono dannose e colpevoli, ma lo è anche lasciarsi distrarre. L'EVENTO: FESTIVALFILOSOFIA Il tema della 26ª edizione, "Caos" , è inteso come disordine ma anche come forza generatrice di cambiamento. In programma oltre 50 lezioni magistrali e quasi 200 eventi. A ingresso gratuito Appuntamenti - Mariarosaria Taddeo insegna Digital Ethics and Defence Technologies all'Oxford Internet Institute si occupa di filosofia della tecnologia e etica applicata. Il suo ultimo libro è Etica dell'intelligenza artificiale nella difesa (Raffaello Cortina). È ospite a Carpi venerdì 18 settembre (ore 15) di Festivalfilosofia con la lectio 'Chi decide? Un'etica per l'intelligenza artificiale applicata alla sicurezza'"

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