GIUSTIZIA INSIEME

Diritto e innovazione

Intelligenza artificiale e diritti. L'UE emana il suo regolamento

di Maria Teresa Covatta

16 gennaio 2024

A fine novembre si è celebrato il primo anniversario di ChatGPT, il software di intelligenza generativa più noto che, insieme ad altre forme di applicazioni dell'intelligenza artificiale dai nomi affascinanti (Bard, LLaMM, Claude e simili), ha riproposto il problema della crescita esponenziale e inevitabile dell'intelligenza artificiale, delle sue prospettive, e, secondo molti, delle sue minacce.

Mettendo da parte le riflessioni morali-filosofiche che le nuove tecnologie impongono, prima tra tutte il concreto pericolo che in un futuro vicino esse possano superare e travolgere la centralità dell'uomo, non vi è dubbio che da quando OpenAI ha lanciato ChatGPT nell'autunno del 2022 l'attenzione globale è stata costretta a concentrarsi sui rischi e sui benefici dell'Intelligenza artificiale in campo politico e sociologico. Soprattutto è diventato evidente che, con così tante incognite, è difficile identificare un percorso chiaro e diretto che consenta di sfruttare al meglio queste tecnologie e di modellarne gli usi ai fini di pace, sicurezza e garanzia per i diritti, sia nelle relazioni internazionali sia nei vari campi del diritto interno agli Stati.

È recentissima la notizia che il New York Times ha portato Elon Musk a giudizio davanti alla Corte di Manhattan per la violazione di diritti di copyright, con l'accusa di aver "dato in pasto" a ChatGPT centinaia e centinaia di articoli per istruirlo nell'arte del giornalismo, impossessandosi di risultati per i quali il NYT investe da sempre migliaia e migliaia di dollari in formazione.

Ed è di questi giorni anche la notizia che nel 2021 un robot operaio abbia quasi atterrato un ingegnere della Tesla con cui lavorava, evocando scenari da 2001 Odissea nello Spazio e ponendo ancora una volta il problema del rapporto uomo macchina che Stanley Kubrick aveva già immaginato nel lontano 1968.

Ma passando a rischi ben più gravi di quello economico lamentato dal NYT, e superando l'episodio del 2021 come un mero incidente di percorso (così lo ha definito lo stesso Elon Musk) non vi è dubbio che la crescita e la diversità delle applicazioni dell'intelligenza artificiale, ove lasciata senza controllo, non potrà che aggravare il già netto divario digitale e tecnologico tra chi ha e chi non ha all'interno del sistema internazionale, con nette ricadute anche in ambito nazionale.

Così alimentando i timori più correnti circa l'impatto dell'AI sul libero esercizio di diritti fondamentali quali, ad esempio, il diritto al lavoro, implicando una riduzione dell'occupazione disponibile e talora addirittura la scomparsa di talune tipologie di lavoro che l'AI sarebbe in grado di svolgere in modo automatico, più veloce e meno costoso per le aziende. O, ancora a titolo di esempio, sul diritto alla libertà di riunione e di manifestazione del pensiero che potrebbe essere limitata dall'uso dell'AI, in grado di profilare individui legati a determinati gruppi politici o di opinione. O infine, incidere sulla libertà di informazione e segnatamente sul diritto ad una corretta informazione, alimentando il terreno già fertile della disinformazione.

Poiché la creazione e lo sviluppo dell'AI è di fatto concentrata in pochi Paesi e gestita attraverso una manciata di società private, la maggior parte del mondo dovrà concentrarsi sul problema di creare un quadro di governance comune a tutti per evitare o almeno limitare le possibili azioni avverse di chi, in possesso di tali tecnologie, possa condizionare il resto del pianeta.

Le sfide del 2024 sono tante, a cominciare da quella della gestione dell'informazione, che a livello globale si troverà a fronteggiare la massiccia ondata di elezioni che si svolgeranno in tutto il mondo e il rischio di disinformazione che vi è connesso.

Un numero senza precedenti di oltre 50 Paesi, tra cui alcuni molto popolosi (Bangladesh, India, Iran, Pakistan, Russia, Stati Uniti, solo per citarne alcuni) e che coprono quasi 2 miliardi di persone, si recheranno alle urne nel 2024. Questo rende evidente come disinformazione o fake news potrebbero mettere a serio rischio il democratico svolgimento delle competizioni elettorali.

Ancor prima che l'intelligenza artificiale facesse irruzione nella nostra vita quotidiana con la visibilità che oggi la caratterizza, la crescita di internet e dei social media avevano reso il problema dell'informazione, misinformazione e disinformazione (MDM) una delle più grandi

sfide del nostro tempo, con un impatto deflagrante sulle politiche pubbliche e sulla società. In questo scenario, che già si rappresentava come una guerra in perdita, l'AI, ove non opportunamente regolamentata, ha indubbiamente la capacità di aggiungere ulteriore potenza alla pozione avvelenata delle deep fake, dell'infoware e dell'ipercomunicazione**[1]**.

Le premesse non sono buone. L'AI generativa ha un ruolo critico nello sviluppo di sofisticati sistemi di deepfake e ha già contribuito alla proliferazione di immagini e audio fuorvianti.

A mero titolo di esempio la pubblicazione negli Stati Uniti, a novembre di quest'anno, di un video che immagina, in caso di rielezione del presidente Biden, un futuro in cui sciami di migranti attraversano il Paese, mentre soldati armati pattugliano strade vuote in un clima da terza guerra mondiale. I Russi, del resto, avevano già fatto la loro parte pubblicando, a maggio di quest'anno, sul canale russo RT.com, un video che mostrava il Pentagono in fiamme colpito da missili.

Documenti che evidenziano come l'uso della disinformazione – e dell'AI che può renderla sempre più raffinata e credibile – possa falsare la libertà di voto dei cittadini e comunque influenzarne le coscienze.

Senza dimenticare che l'uso dei deep fake prende di mira in particolare le donne, aggiungendo un altro strumento agli abusi digitali e alle molestie che notoriamente le vittimizzano in tutto il mondo.

Forse uno degli ambiti che riceve maggiore attenzione riguarda l'uso militare delle nuove tecnologie nei conflitti attuali e futuri.

Lo stato attuale del mondo, dai conflitti in Ucraina e Gaza ai tanti focolai di tensione che esistono da tempo o che spuntano di continuo, ci offriranno ampie opportunità di testare cosa ci riserva il futuro.

È cresciuta la consapevolezza del pericolo che l'intelligenza artificiale può rappresentare nell'uso delle armi nucleari e di altre armi avanzate ma anche nell'individuazione degli obiettivi da colpire.

Purtroppo anche in quest'ultimo campo i timori si sono rivelati fondati, visto che le recenti esperienze nei conflitti in corso mostrano il pessimo uso fatto dell'AI che non ha operato nel senso di salvaguardare vite umane e garantire il diritto umanitario internazionale, quanto piuttosto è stata finalizzata ad ottenere una maggiore capacità di raggiungimento degli obiettivi bellici, con buona pace delle speranze di chi ipotizzava l'uso benefico dell'intelligenza artificiale nel *peacebuilding***[2]**.

Senza contare che la costruzione della pace passa anche attraverso la creazione di un sistema internazionale di limitazione delle disuguaglianze sociali che per contro potrebbero essere esacerbate dall'applicazione dell'AI nel settore finanziario.

I governi e le banche private stanno da tempo utilizzando questi nuovi strumenti per prendere decisioni sui prestiti concedibili e per valutare i rischi rispetto alle operazioni bancarie di ogni livello. E tuttavia un maggior coinvolgimento dell'intelligenza artificiale in termini decisionali nei servizi finanziari, con strumenti che individuano con criteri automatizzati Soggetti e Paesi da considerarsi più o meno rischiosi in termini di investimenti, potrebbe limitare l'accesso alle banche e ai crediti d'aiuto proprio per quei popoli che ne avrebbero più bisogno perché afflitti da guerre, carestie e crisi economiche.

Un'altra area di utilizzo dell'AI che sta richiamando l'attenzione a livello internazionale è la tecnologia di sorveglianza, che può condizionare anche l'esercizio delle attività di polizia e della giurisdizione.

In questo campo l'intelligenza artificiale senz'altro costituisce un ausilio ai sistemi di sicurezza, avendo la capacità di rivedere e analizzare dati e documenti audio con capacità ben più veloci di quanto siano in grado di fare gli umani specialisti nel campo; con risultati forse – a saldo dei margini di errore – anche più precisi.

E tuttavia pone non pochi problemi per il suo possibile utilizzo nel monitoraggio su larga scala delle azioni dei cittadini, in particolare sul fronte della polizia e della giustizia predittive.

Nonostante tutta l'attenzione sembri concentrarsi su ChatGPT, esiste, in tutti gli angoli dell'internet, un incredibile numero di opzioni di intelligenza artificiale generativa prodotta da aziende, governi e enti di ricerca che attingono ad un enorme volume di dati e che mettono in pericolo la privacy della vita quotidiana, delle scelte religiose, sessuali e politiche dei cittadini del mondo: pericolo reso sempre più evidente dall'aumento dell'autoritarismo mondiale e dei continui attacchi ai processi democratici.

A fronte di questo scenario negativo e pieno di dubbi le speranze non mancano.

Prima di tutte quella di usare il tempo che abbiamo per far sì che l'intelligenza artificiale possa rivoluzionare davvero le nostre vite in positivo così come è stato per il progresso tecnologico che fin ora ha comportato un beneficio netto per l'umanità. E che il tempo possa essere usato per conformare risposte normative idonee per plasmare responsabilmente l'evoluzione di questa tecnologia.

L’utilizzo dell’intelligenza artificiale nei conflitti potrebbe essere capovolta rispetto all’uso bellico fin ora registrato per divenire uno strumento prezioso nei processi di pace.

I droni non armati potrebbero svolgere un ruolo importante nel monitoraggio delle linee di contatto e delle violazioni del cessate il fuoco e potrebbero fare un’operazione di verità sulle azioni di guerra, contribuendo, unitamente alle immagini satellitari, a identificare i crimini di guerra e forse evitare, per il timore di essere denunciati alla giustizia internazionale, devastazioni, violenze di massa e violazioni del diritto umanitario internazionale.

Tra gli altri possibili usi benevoli l’utilizzo dei dati storici che l’AI può ricercare e identificare con maggiore facilità e precisione di quanto potrebbero fare significative risorse umane e che potrebbero essere utilizzati in vari settori quali sanità, giustizia e legislazione tanto per citare le più importanti.

Ma affinché questi obiettivi possano essere realizzati, la prima esigenza da soddisfare è la regolamentazione dell’AI con criteri ben determinati, in un quadro di livello globale che tenga conto, in primo luogo, della salvaguardia dei diritti umani e che consenta di sviluppare risposte politiche adeguate a trarne il massimo beneficio e penalizzarne l’uso nocivo.

In quest’ottica sembra che si stiano orientando i governi e le organizzazioni internazionali che stanno discutendo del futuro dell’AI e di come dovrebbe essere gestito, con la proposta di possibili modelli da usare per l’obiettivo generalmente condiviso di evitare i suoi danni, voluti o non voluti che siano.

E, sempre sulla stessa strada, il Segretariato Generale dell’ONU ha lanciato un progetto di discussione che coinvolgerà tutti gli stati membri e culminerà nel “Summit of the Future” previsto per il settembre del 2024 con l’ambiziosa finalità di valutare tutte le implicazioni della *computing capacity* e degli algoritmi e di raggiungere l’accordo su un Global Digital Compact, valevole e vincolante per tutti i firmatari.

Anche l’Unione Europea sta continuando il suo sforzo, avviato già nel 2018, di creare una governance idonea a mitigare i rischi della tecnologia emergente e sempre più in fase avanzante, al fine di fare in modo che essa possa contribuire al bene pubblico comune.

Il 13 dicembre, dopo più di 2 anni di gestazione. l’Unione ha finalmente raggiunto un accordo politico per l’*Artificial Intelligence Act*, il regolamento sull’intelligenza artificiale, frutto di una negoziazione su base trilaterale tra Parlamento,Consiglio e Commissione, finalizzata a limitare l’uso dei software deputati alla ricognizione facciale, richiedere alle compagnie che creano e gestiscono i software di elaborare modelli di linguaggio informatico (LLMs) per divulgare e

rendere accessibili i dati utilizzati per creare i loro software e imporre alle compagnie che operano nel campo delle nuove tecnologie di effettuare la certificazione del rischio prima che i loro prodotti siano utilizzati.

Pur non essendo stato ancora, al momento, pubblicato il testo definitivo, gli aspetti salienti emersi dalla documentazione già pubblicata, in particolare il *draft* del testo elaborato dal Parlamento, molto avanzato sul profilo della tutela dei diritti, e dalla conferenza stampa, riguardano in particolare il nodo del riconoscimento facciale da remoto (RBI, *Remote Biometric Identification*), le proibizioni e le relative sanzioni in caso di violazioni.

Proibiti i sistemi di categorizzazione biometrica che utilizzano caratteristiche sensibili quali convinzioni politiche, religiose filosofiche, orientamento sessuale e razza; nonché la raccolta non mirata di immagini di volti da Internet o da filmati da telecamere a circuito chiuso per create database di riconoscimento facciale; e ancora proibito il riconoscimento e classificazione di emozioni sul posto di lavoro o nelle istituzioni scolastiche; o il cosiddetto social scoring, a classificazione degli individui basata sul comportamento sociale: nonché qualunque tipo di sistema di intelligenza artificiale che sia idoneo, in qualunque forma concepito, a manipolare il comportamento umano e coartare la libera volontà delle persone.

Bandita, infine, la cosiddetta polizia predittiva quanto meno nella forma in cui si valuta il rischio che un certo individuo possa commettere reati futuri sulla base di tratti personali.

Verosimilmente a causa delle pressioni dei governi sul punto, manca una disposizione di messa al bando totale sui sistemi di identificazione da remoto in spazi aperti al pubblico, con la previsione di eccezioni previamente autorizzate dall'autorità giudiziaria e solo per liste di reati rigorosamente definite dalla legislazione. La RBI da remoto *ex post* potrà essere utilizzata solo per la ricerca mirata di persone condannate o sospettate di aver commesso tali reati; mentre quella in tempo reale, invece, potrà essere utilizzata solo per le ricerche mirate di vittime di reato o per la prevenzione di minacce terroristiche che siano valutate come specifiche e attuali.

Da questo sintetico resoconto emerge che si tratta di una cornice normativa importante sull'uso globale dell'intelligenza artificiale che mira ad avere un impatto simile a quello della regolamentazione europea dell'*e-commerce*, sulla lotta alla disinformazione e ai linguaggi di odio sui social media e sulla normativa sulla tutela della privacy e sull'uso dei dati privati dei cittadini.

E non vi è dubbio che già il fatto di aver raggiunto un accordo politico vincolante sulla regolamentazione, nonostante la quantità di soggetti e interessi che remavano contro, sia già di

per sé un successo.

Ma il successo di maggior peso consisterà nella reale presa di coscienza a livello internazionale – e delle azioni conseguenti – che, dalle finanze al sistema degli armamenti, dai social media alla sanità, dalla giustizia e al ruolo nei conflitti, l'AI è destinata ad espandersi, che un uso non controllato delle sue applicazioni può causare pericoli forse allo stato neppure immaginabili e che la globalità del fenomeno impone che sia la comunità internazionale a dover valutare i possibili modelli per l'individuazione e il conseguimento di tutti gli obiettivi che l'intelligenza potrebbe consentire di raggiungere ove gestita adeguatamente.

Questo risultato consentirebbe anche di riconoscere e apprezzare l'indubbio fascino dell'intelligenza artificiale.

C'è da chiedersi se saremo in grado di cogliere questa opportunità.

---

[1] ISPI – Istituto Studi Politiche Internazionali Dossier 2023.

[2] USIP - *United States Institut of Peace* – Ather Ashby: "Un ruolo per l'intelligenza artificiale nella costruzione della pace" – 6.12.2023.

Per approfondire queste tematiche su Giustizia Insieme:

Machina delinquere non potest di Costanza Corridori, **L'Intelligenza Artificiale nei processi decisionali: il pericolo per la giustizia** di Rosita D'Angiolella, **Nomofilachia, giustizia predittiva e intelligenza artificiale** di Giovanni Ariolli.