

Intelligenza Artificiale e società

Versione finale

L'Intelligenza Artificiale (IA) sta rimodellando il nostro mondo, offrendo numerosi vantaggi ma sollevando anche molte preoccupazioni che devono essere affrontate attraverso una *governance* inclusiva e coordinata a livello globale. Nell'ottobre 2023, i leader del G7 hanno pubblicato i Principi Guida Internazionali dell'Accordo di Hiroshima e il relativo Codice di condotta indirizzati alle organizzazioni pubbliche e private che sviluppano e utilizzano sistemi avanzati di intelligenza artificiale. Con la presente dichiarazione congiunta, esprimiamo la posizione delle Accademie delle scienze dei paesi del G7 in merito alle preoccupazioni tecniche e sociali generate dall'IA con un elenco di raccomandazioni indirizzate ai responsabili politici.

1. I potenti sistemi di IA, che presentano rischi potenziali per la collettività se utilizzati in modo improprio, devono essere adeguatamente protetti dagli attacchi informatici e fisici. Per le applicazioni sensibili di intelligenza artificiale, le normative sulla protezione dei dati (quali, ad esempio, il Regolamento generale sulla protezione dei dati dell'UE) commisurate al livello di rischio e l'uso di tecnologie per la tutela della privacy, sono strumenti essenziali per garantire la protezione dei dati personali. Nelle interazioni con i sistemi di IA, gli utenti hanno bisogno di indicazioni chiare sulla protezione dei loro dati e su come e per quanto tempo i loro dati saranno utilizzati, riutilizzati e archiviati.
2. Con l'aumentare delle capacità dei sistemi di IA e dei rischi sempre maggiori che presentano per il controllo umano, sarà necessario avere garanzie di sicurezza formali prima di poter procedere alla loro implementazione. È essenziale aderire a standard rigorosi per la verifica dei requisiti, la convalida e il collaudo dei sistemi e il monitoraggio successivo una volta implementati, affidato preferibilmente ad organismi indipendenti responsabili anche del controllo di eventuali *bias* nascosti. È necessario promuovere leggi che impongano l'obbligo di fornire una documentazione adeguata affinché gli utenti possano comprendere le capacità, le limitazioni e gli ambiti appropriati delle applicazioni. La tracciabilità dei dati di base su cui sono costruiti i modelli di IA è fondamentale. Poiché l'IA influisce su decisioni delicate in settori come la medicina, le scienze, il diritto e la finanza, la trasparenza e la responsabilità sono di particolare importanza. È necessario definire chiare linee di responsabilità durante lo sviluppo e l'implementazione dei sistemi di intelligenza artificiale. Funzionalità come la spiegabilità, gli algoritmi impiegati e le modalità da seguire per presentare un reclamo consentono agli utenti di contestare le decisioni dell'IA.

3. L'affidabilità dei sistemi di IA è di vitale importanza. In considerazione del fatto che gli esseri umani trascorrono sempre più tempo ad interagire con i sistemi di IA e poiché i contenuti generati dall'IA costituiscono una porzione sempre maggiore del nostro ecosistema informativo, i rischi per la democrazia, per la società e per l'intelligenza critica umana aumentano. I governi devono promuovere standard legalmente riconosciuti per regolamentare i contenuti generati dall'IA. Le strutture istituzionali che promuovono l'accuratezza e l'autenticità diventeranno sempre più importanti. Inoltre, l'alfabetizzazione informatica continua delle popolazioni onde consentirle di identificare e interagire con i contenuti generati dall'IA sarà un elemento chiave. Gli sviluppatori di tecnologia e le piattaforme multimediali online devono considerare il valore dei sistemi di provenienza dei contenuti digitali e prevedere l'inclusione di esperti di dominio nelle valutazioni socio-tecniche (ad esempio, il red teaming).
4. L'IA generativa offre un grande potenziale per la creazione di contenuti e di sviluppo di software, ma solleva interrogativi sulla tutela dei diritti di proprietà intellettuale quando l'IA crea contenuti apparentemente nuovi da opere esistenti. Le industrie creative e la ricerca scientifica sono tra i settori principali su cui l'IA generativa sta avendo un impatto significativo. La normativa esistente sul diritto d'autore, l'uso equo e la protezione della proprietà intellettuale va aggiornata e adattata all'era dell'IA. Per affrontare questi aspetti è necessario disporre di norme chiare, dell'autoregolamentazione del settore e di una supervisione indipendente.
5. I sistemi di intelligenza artificiale su larga scala richiedono risorse computazionali costose, creando dunque una barriera all'ingresso. Ampliare l'accesso globale al calcolo ad alte prestazioni e creare centri di ricerca intergovernativi è vitale per una fruizione equa dell'IA in futuro. L'espansione delle risorse computazionali dell'IA solleva preoccupazioni ambientali a causa del crescente consumo di energia. Perché l'IA sia conforme ai principi di etica ambientale occorre promuovere la progettazione sostenibile dei data center, scegliere siti strategici dal punto di vista dell'efficienza energetica e del raffreddamento, e sviluppare nuove tecniche di IA efficienti dal punto di vista energetico.
6. L'open source di insiemi di dati selezionati attentamente e senza problemi di copyright è generalmente un bene pubblico da incoraggiare, ma i grandi modelli di IA open source presentano un complesso tradeoff tra vantaggi (ad esempio, per i ricercatori del settore pubblico e i piccoli sviluppatori a valle) e rischi derivanti da usi impropri che sarebbero a carico del pubblico. La decisione se consentire o limitare la diffusione open source di potenti sistemi di intelligenza artificiale deve essere soggetta a controllo democratico, e le norme di sicurezza che si applicano ai sistemi proprietari devono applicarsi anche a tutte i casi di sistemi open source.
7. La comparsa di sistemi d'arma autonomi solleva pressanti preoccupazioni etiche e militari. La loro autonomia potrebbe facilitare i conflitti, stimolare una corsa agli armamenti e rischiare abusi da parte di regimi canaglia o di terroristi. Alcuni sostengono il divieto per tali armi, mentre altri invocano norme rigorose in linea con il diritto umanitario internazionale, sotto la sorveglianza dell'uomo. E' fondamentale che vi sia un intervento rapido da parte della società per affrontare questa importante sfida. La trasparenza pubblica e il dibattito sono fondamentali per affrontare queste sfide, con l'obiettivo di sostenere i valori umani e le norme internazionali, nonché la sicurezza fisica a livello personale, locale, nazionale e internazionale.
8. Allineare l'IA con i valori umani e con l'etica è necessario poiché l'influenza dei sistemi di IA è destinata ad aumentare. La collaborazione intersettoriale è fondamentale per implementare un'IA etica, integrando aspetti sia tecnici che socio-culturali e tenendo

conto anche delle diversità culturali. Affrontare i rischi a breve termine, che potrebbero avere un impatto negativo sui gruppi sociali vulnerabili, è fondamentale per mantenere la fiducia e incoraggiare l'adozione di tecnologie vantaggiose. I costi economici e sociali dell'allineamento devono essere trasparenti e inclusi nel processo decisionale relativo a quando e come utilizzare i sistemi di intelligenza artificiale.

9. Un'amministrazione efficace dell'IA è fondamentale per una distribuzione uniforme dei suoi benefici. Sistemi di IA caratterizzati da capacità operative sempre crescenti avranno un impatto significativo sui mercati del lavoro per cui è essenziale che vi sia un approccio congiunto di governi, industria, educatori e società civile per garantire un'equa distribuzione dei benefici. I governi devono prendere in considerazione incentivi per sistemi di IA socialmente utili che rispondano a bisogni disattesi piuttosto che limitarsi a sostituire gli esseri umani nei ruoli esistenti. Le politiche devono dare priorità allo sviluppo di sistemi di IA che affrontino le sfide più importanti per il pianeta, per la società e per l'economia.
10. La cooperazione tra il settore pubblico e quello privato assicura uno sviluppo responsabile dell'IA. Gli enti privati innovano rapidamente, mentre le istituzioni pubbliche offrono quadri normativi ed etici e promuovono ricerche indipendenti pluriennali a beneficio dell'umanità. Lo scambio di talenti consente di unire le migliori pratiche con gli obiettivi. Questa sinergia unisce l'innovazione con la responsabilità, ottimizzando i benefici sociali dell'IA, ed è sfruttata al meglio quando la società, attraverso la normativa e i processi democratici, traccia la linea rossa tra ciò che è accettabile e ciò che non lo è e gli sviluppatori si assumono l'onere di dimostrare che i loro sistemi non oltrepasseranno mai tale linea.
11. Lo sviluppo e la diffusione responsabili e sicuri dell'IA non sono esclusivamente una questione scientifica e tecnologica: richiedono un approccio collaborativo che coinvolge diverse discipline accademiche, comprese le scienze umane e sociali, per comprendere e modellare il complesso sistema socio-tecnologico all'interno del quale l'IA opera e per definire meccanismi di *governance* adeguati. Unendo i punti di vista tecnici con quelli non tecnici, possiamo affrontare le sfide che l'IA ha davanti come ad esempio pregiudizi ed equità, garantendo un suo sviluppo inclusivo, affidabile, trasparente ed equo.
12. L'istruzione deve offrire un quadro chiaro delle capacità dell'IA, sfatare i miti e consentire di valutarne i pro e i contro. I cittadini devono essere tecnicamente alfabetizzati e resi consapevoli delle implicazioni sociali ed etiche dell'IA. Il dibattito pubblico, facilitato dalle varie parti interessate, consente ai singoli individui di modellare attivamente le politiche sull'IA e allineare l'IA con l'interesse pubblico. Infine, si devono bilanciare attentamente i rischi e le opportunità derivanti dall'integrazione di potenti sistemi di IA nel sistema educativo.