

ROME CALL FOR AI ETHICS

INTRODUZIONE

L'"Intelligenza Artificiale" (IA) sta portando profondi cambiamenti nella vita degli esseri umani, e continuerà a farlo. L'IA offre un enorme potenziale quando si tratta di migliorare la convivenza sociale e il benessere personale, aumentando le capacità umane e facilitando o consentendo molte attività che possono essere svolte in modo più efficiente ed efficace. Tuttavia, questi risultati non sono affatto garantiti. Le trasformazioni attualmente in corso non sono solo quantitative. Fra le altre cose, sono qualitative, perché influenzano il modo in cui queste attività vengono svolte e il modo in cui percepiamo la realtà e la natura umana stessa, tanto da poter influenzare le nostre abitudini mentali e interpersonali. La nuova tecnologia deve essere studiata e prodotta secondo criteri che ne garantiscano il vero servizio all'intera "famiglia umana" (Preambolo, Dichiarazione Universale dei Diritti Umani), rispettando la dignità intrinseca di ciascun suo membro e di tutti gli ambienti naturali, e tenendo conto dei bisogni di coloro che sono più vulnerabili. L'obiettivo non è solo garantire che nessuno sia escluso, ma anche espandere quelle aree di libertà che potrebbero essere minacciate dal condizionamento algoritmico.

Dato la natura innovativa e complessa delle questioni poste dalla trasformazione digitale, è essenziale che tutti gli attori coinvolti collaborino e che tutti i bisogni influenzati dall'IA siano rappresentati. Questo Appello è un passo avanti con l'obiettivo di crescere con una comprensione comune e di cercare un linguaggio e soluzioni che possiamo condividere. Sulla base di questo, possiamo riconoscere e accettare responsabilità che tengono conto dell'intero processo di innovazione tecnologica, dalla progettazione alla distribuzione e all'uso, incoraggiando un impegno reale in una serie di scenari pratici. Nel lungo termine, i valori e i principi che siamo in grado di instillare nell'IA contribuiranno a stabilire un quadro che regoli e agisca come punto di riferimento per l'etica digitale, guidando le nostre azioni e promuovendo un uso della tecnologia che vada a beneficio dell'umanità e dell'ambiente.

Ora più che mai, dobbiamo garantire una prospettiva in cui l'IA sia sviluppata con un focus non sulla tecnologia, ma piuttosto per il bene dell'umanità e dell'ambiente, della nostra casa comune e condivisa e dei suoi abitanti, che sono indissolubilmente legati. In altre parole, una visione in cui gli esseri umani e la natura sono al centro di come viene sviluppata l'innovazione digitale, supportata, piuttosto che gradualmente sostituita, da tecnologie che si comportano come attori razionali ma non sono affatto umane. È tempo di iniziare a prepararsi per un futuro più tecnologico in cui le macchine avranno un ruolo più importante nelle vite degli esseri umani, ma anche un futuro in cui è chiaro che

il progresso tecnologico conferma la genialità della razza umana e sia basato sulla sua integrità morale.

ETICA

Tutti gli esseri umani nascono liberi ed eguali in dignità e diritti. Sono dotati di ragione e coscienza e dovrebbero comportarsi gli uni verso gli altri in uno spirito di fratellanza (cfr. Art. 1, Dichiarazione Universale dei Diritti Umani). Questa condizione fondamentale di libertà e dignità deve essere protetta e garantita anche nella produzione e nell'uso dei sistemi di IA. Questo deve essere realizzato salvaguardando i diritti e la libertà degli individui in modo che non siano discriminati dagli algoritmi a causa della loro "razza, colore, sesso, lingua, religione, opinione politica o di altro genere, origine nazionale o sociale, patrimonio, nascita o altro status" (Art. 2, Dichiarazione Universale dei Diritti Umani).

I sistemi di IA devono essere concepiti, progettati e implementati per servire e proteggere gli esseri umani e l'ambiente in cui vivono. Questo punto di vista fondamentale deve tradursi in un impegno a creare condizioni di vita (sia sociali che personali) che consentano sia ai gruppi che ai singoli membri di cercare di esprimersi pienamente quando possibile.

Perché l'avanzamento tecnologico sia in linea con il vero progresso per la razza umana e il rispetto per il pianeta, deve soddisfare tre requisiti. Deve includere ogni essere umano, senza discriminazione; deve porre al centro il bene dell'umanità e di ogni essere umano; infine, deve tenere conto della complessa realtà del nostro ecosistema e deve essere caratterizzato dal modo in cui si prende cura e protegge il pianeta (la nostra "casa comune e condivisa") con un approccio altamente sostenibile, che includa anche l'uso dell'intelligenza artificiale nel garantire sistemi alimentari sostenibili in futuro. Inoltre, ogni persona deve essere consapevole quando interagisce con una macchina.

La tecnologia basata sull'IA non deve mai essere utilizzata per sfruttare le persone in alcun modo, specialmente quelle più vulnerabili. Invece, deve essere utilizzata per aiutare le persone a sviluppare le loro abilità (rafforzamento/consapevolezza) e per sostenere il pianeta.

ISTRUZIONE

Trasformare il mondo attraverso l'innovazione dell'IA significa impegnarsi a costruire un futuro per e con le nuove generazioni. Questo progetto deve riflettersi in un impegno per l'istruzione, sviluppando curricula specifici che spaziano tra diverse discipline nelle scienze umane, nella scienza

e tecnologia, e assumendosi la responsabilità di educare le nuove generazioni. Questo impegno significa lavorare per migliorare la qualità dell'istruzione che i giovani ricevono; questo deve essere fornito attraverso metodi accessibili a tutti, che non discriminino e che possano offrire parità di opportunità e di trattamento. L'accesso universale all'istruzione deve essere raggiunto attraverso principi di solidarietà e equità.

L'accesso alla formazione continua deve essere garantito anche agli anziani, che devono avere l'opportunità di accedere a servizi offline durante la transizione digitale e tecnologica. Inoltre, queste tecnologie possono rivelarsi estremamente utili nell'aiutare le persone con disabilità a imparare e a diventare più indipendenti: l'istruzione inclusiva significa quindi anche utilizzare l'IA per supportare e integrare ogni singola persona, offrendo aiuto e opportunità per la partecipazione sociale (ad esempio, lavoro da remoto per coloro che hanno limitata mobilità, supporto tecnologico per coloro che sono affetti da disabilità cognitive, ecc.).

L'impatto delle trasformazioni apportate dall'IA nella società, nel lavoro e nell'istruzione ha reso essenziale rivisitare i curricula scolastici al fine di rendere il motto educativo "nessuno deve essere lasciato indietro" una realtà. Nel settore dell'istruzione, sono necessarie riforme volte a stabilire standard alti ed obiettivi, che possono migliorare i risultati individuali. Questi standard non dovrebbero limitarsi allo sviluppo delle competenze digitali, ma dovrebbero invece concentrarsi sul garantire che ogni persona possa esprimere appieno le proprie capacità e sul lavorare per il bene della comunità, anche quando non vi è alcun beneficio personale da trarne.

Mentre progettiamo e pianifichiamo la società di domani, l'uso dell'IA deve seguire forme di azione che siano socialmente orientate, creative, in grado di instaurare relazioni, produttive, responsabili e in grado di avere un impatto positivo sulla vita personale e sociale delle nuove generazioni. L'impatto sociale ed etico dell'IA deve essere anche al centro delle attività educative della stessa.

L'obiettivo principale di questa educazione deve essere quello di sensibilizzare sulle opportunità e anche sulle possibili criticità poste dall'IA dal punto di vista dell'inclusione sociale e del rispetto dell'individuo.

DIRITTI

Lo sviluppo dell'IA al servizio dell'umanità e del pianeta deve riflettersi in normative e principi che proteggano le persone - in particolare i deboli e i meno privilegiati - e gli ambienti naturali. L'impegno

etico di tutti gli attori coinvolti è un punto di partenza cruciale; per rendere questo futuro una realtà, valori, principi e, in alcuni casi, norme di legge, sono assolutamente indispensabili per sostenere, strutturare e guidare questo processo.

Per implementare sistemi di IA che apportino beneficio all'umanità e al pianeta, agendo come strumenti per costruire e mantenere la pace internazionale, lo sviluppo degli stessi deve andare di pari passo con robuste misure di sicurezza digitale.

Affinché l'IA agisca come strumento per il bene dell'umanità e del pianeta, dobbiamo porre al centro del dibattito pubblico il tema della protezione dei diritti umani nell'era digitale. È giunto il momento di interrogarsi se le nuove forme di automazione e attività algoritmiche richiedano lo sviluppo di maggiori responsabilità. In particolare, sarà essenziale considerare una forma di "dovere di spiegazione": dobbiamo pensare a rendere comprensibili non solo i criteri decisionali degli agenti algoritmici basati sull'IA, ma anche il loro scopo e i loro obiettivi. Questi dispositivi devono essere in grado di offrire agli individui informazioni sulla logica alla base degli algoritmi utilizzati per prendere decisioni. Ciò aumenterà la trasparenza, la tracciabilità e la responsabilità, rendendo il processo decisionale assistito dal computer più valido.

Deve essere incoraggiata l'introduzione di nuove forme di regolamentazione per promuovere la trasparenza e il rispetto dei principi etici, specialmente per le tecnologie avanzate che hanno un rischio più elevato di impattare sui diritti umani, come il riconoscimento facciale.

Per raggiungere questi obiettivi, dobbiamo partire fin dall'inizio dello sviluppo di ciascun algoritmo con una visione "algor-etica", ovvero un approccio etico durante la progettazione. Progettare e pianificare sistemi di IA di cui possiamo fidarci comporta la ricerca di un consenso tra i decisori politici, le agenzie del sistema ONU e altre organizzazioni intergovernative, i ricercatori, il mondo accademico e i rappresentanti delle organizzazioni non governative riguardo ai principi etici che dovrebbero essere integrati in queste tecnologie. Per questo motivo, i promotori dell'Appello esprimono il desiderio di lavorare insieme, in questo contesto e a livello nazionale e internazionale, per promuovere l'"algor-etica", nel senso dell'uso etico dell'IA come definito dai seguenti principi:

- **Trasparenza**: in linea di principio, i sistemi di IA devono essere comprensibili quanto al loro funzionamento;

- **Inclusione**: devono essere presi in considerazione i bisogni di tutti gli esseri umani in modo che tutti possano beneficiare dei sistemi di IA e che a tutti gli individui possano essere offerte le migliori condizioni possibili per esprimersi e svilupparsi;
- **Responsabilità**: coloro che progettano e implementano l'uso dell'IA devono procedere con responsabilità e trasparenza;
- **Imparzialità**: non creare o agire sulla base di pregiudizi, salvaguardando così equità e dignità umana;
- **Affidabilità**: i sistemi di IA devono essere in grado di funzionare in modo affidabile;
- **Sicurezza e privacy**: i sistemi di IA devono funzionare in modo sicuro e rispettare la privacy degli utenti.

Questi principi sono elementi fondamentali di una buona innovazione.

Roma, 28 febbraio 2020