

Lettera aperta: l' IA ha bisogno di regole, non di pause

Perché noi, come organizzazioni e singoli cittadini, riteniamo che la moratoria proposta e gran parte dell'attuale dibattito sui sistemi di IA generativa non sia sufficiente a proteggere la nostra società democratica e non colga il punto su come garantire tutele e diritti fondamentali.

Milano, 03/04/2023

Alla gentile attenzione dei membri del Parlamento europeo e italiano,

Noi, le organizzazioni sottoscritte e i singoli cittadini impegnati nella difesa dei diritti civili e nella promozione di tecnologie etiche ed eque, desideriamo riportare la vostra attenzione sulle nostre preoccupazioni riguardo allo sviluppo e all'implementazione di sistemi di intelligenza artificiale (IA), tra cui i Large Language Models (LLM) e i modelli generativi. **Siamo profondamente preoccupati** per le implicazioni di queste tecnologie per la nostra società e per il futuro non solo dell'Europa, e riteniamo che sia cruciale affrontare queste sfide in modo proattivo, collaborativo e responsabile.

Rivolgiamo questa lettera al Parlamento europeo, che il 26 aprile dovrà votare l'accordo politico sul Regolamento sull'Intelligenza Artificiale (AI Act), al Consiglio dell'UE e alla Commissione, che parteciperanno da maggio ai triloghi, e a tutta la società civile interessata a bilanciare i benefici dell'IA indirizzando i problemi legati al suo impatto sulla società, le persone e l'ambiente.



Privacy Network

www.privacy-network.it

Nelle ultime settimane, il focus sugli sviluppi dell'intelligenza artificiale generativa ha portato a un intenso e complesso dibattito sulla regolamentazione e il controllo di queste tecnologie. Modelli come GPT-4, DALL-E, Stable Diffusion e AlphaCode stanno diventando la base di molti sistemi informatici in contesti diversi, come vendite, assistenza clienti, sviluppo software, design, gaming, educazione e molti altri. Tuttavia, l'utilizzo di questi modelli generativi porta con sé rischi e implicazioni potenzialmente problematiche, quali **la diffusione di disinformazione, la generazione di contenuti pregiudizievole e stereotipati, l'emulazione di contenuti estremisti e la creazione di deepfake sempre più verosimili.**

Riconosciamo l'importanza di considerare tematiche geopolitiche e di collaborazione internazionale nell'affrontare questi problemi, e ci impegniamo a lavorare con voi per promuovere un approccio equilibrato e sostenibile all'utilizzo dell'IA non solo nelle applicazioni più recenti e che continueranno a svilupparsi, creando una normativa europea future-proof che si concentri sulla valutazione preventiva del rispetto dei diritti fondamentali e della sicurezza.

Sulla moratoria del *Future of Life Institute*

Vorremmo richiamare la vostra attenzione sulla recente proposta di moratoria avanzata dal *Future of Life Institute* del 28 Marzo 2023. Tale moratoria è una lettera aperta che ha finora raccolto oltre 1.800 firme di esperti e cittadini comuni. La proposta sollecita una sospensione di sei mesi nello sviluppo dei modelli generativi come i LLMs attuali, definiti in modo fin troppo ambiguo come “powerful AI systems with human-competitive intelligence”. L'obiettivo di questa moratoria è chiedere che lo sviluppo dell'IA generativa sia focalizzato su rendere “powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal”.

Sebbene alcune delle preoccupazioni esposte nella lettera siano condivisibili, riteniamo che essa presenti alcune **limitazioni e problemi che stanno distogliendo l'attenzione da ciò che davvero necessita l'impegno.** L'opinione pubblica si sta concentrando giustamente su questi modelli, ma pensiamo che l'hype gonfiato ed esclusivo su queste applicazioni ci stia distraendo dagli impatti concreti e dalle modalità da impiegare per mitigarli. Nei punti successivi di questa

lettera, esporremo le preoccupazioni riguardo alla narrativa che si sta sviluppando e agli aspetti della moratoria proposta che riteniamo insufficienti o problematici:

- La sostanziale invalidità di qualsiasi richiamo e/o lasso temporale proposto a sospendere lo sviluppo di tali sistemi, data l'inattualità delle soluzioni ivi contenute a livello **geopolitico e di mercato**, nonché la **pericolosa abilitazione di una narrazione sull'IA** che la ritrae come qualcosa che ci supererà nelle abilità umane cognitive e ci renderà obsoleti;
- La necessità di affrontare problemi più concreti e pressanti relativi all'ecosistema delle tecnologie di intelligenza artificiale generativa, come **la crescente concentrazione di potere nelle mani delle aziende tecnologiche**, le sfide per una regolamentazione adeguata e le questioni legate al **colonialismo digitale**;
- L'importanza di un **approccio proattivo al risk management**, e alla valutazione preventiva di queste tecnologie che coinvolga una vasta gamma di stakeholder nel processo decisionale;
- La promozione della **trasparenza in tutte le fasi del processo di sviluppo e utilizzo** dell'IA, al fine di garantire un approccio efficace al risk management e al controllo delle tecnologie;
- L'importanza di considerare i potenziali **danni psicologici** causati dall'IA generativa e di sviluppare strategie per prevenire e mitigare tali effetti negativi sulla salute mentale degli individui;
- La necessità di **cooperazione internazionale** e di **sviluppo di normative e politiche condivise** a livello globale per affrontare i rischi e le sfide dell'IA generativa in modo armonizzato.

Riteniamo che questi temi siano fondamentali per garantire un approccio equilibrato e sostenibile all'impiego dell'IA generativa e per promuovere un futuro più sicuro, etico ed equo per tutti. Vi invitiamo a prendere in considerazione queste preoccupazioni nella vostra valutazione della moratoria proposta e nel vostro impegno per la regolamentazione e la supervisione di tali tecnologie.

Prospettive, problemi e soluzioni possibili

Partendo dalla moratoria, la narrazione con cui si rappresenta lo sviluppo di tali sistemi IA sembra voler legittimare una narrazione di un'IA che sarà presto cosciente, dunque abile nel esercitare un potere ed una minaccia all'umanità tutta.

Queste narrazioni sci-fi **spostano l'attenzione dai rischi e gli impatti già problematici, come le discriminazioni o l'esclusione di alcuni gruppi sociali, e impediscono di individuare le responsabilità.** È fondamentale considerare come gli stati, le loro politiche e le dinamiche di potere tra di loro influenzino significativamente il modo in cui vengono sviluppati e utilizzati i modelli di intelligenza artificiale. La competitività globale nel campo dell'IA e l'interesse strategico delle principali potenze mondiali in tale tecnologia rendono imprescindibile affrontare le questioni geopolitiche per garantire uno sviluppo equo ed etico dell'IA. Tali questioni includono:

- 1. Le divergenze normative e di valori** possono ostacolare la creazione di un quadro regolamentare condiviso a livello globale, ma allo stesso vanno riconosciute come fondamentali nell'autodeterminazione delle culture.
- 2. Il risk management, l'assessment e la trasparenza nell'IA** hanno un ruolo fondamentale nel garantire la responsabilità e la fiducia nel processo di sviluppo e utilizzo dell'IA. È pertanto essenziale promuovere l'adozione di strumenti e prassi che consentano di valutare e gestire efficacemente i rischi associati a queste tecnologie. Per questo vi esortiamo a:
 - a. Sviluppare strumenti e metodologie che permettano di identificare e quantificare anticipatamente i rischi legati all'IA in termini di potenziali impatti negativi sui **diritti fondamentali, la società, la sicurezza e l'ambiente** (indirizzando direttamente le emissioni legate allo sviluppo dell'IA), e che richiedano revisioni periodiche
 - b. L' auditing di tali sistemi, effettuato da entità indipendenti e qualificate, contribuisce a garantire il rispetto delle normative vigenti, una chiara documentazione delle metodologie di sviluppo, dei modelli di apprendimento e dei criteri di valutazione utilizzati, nonché la possibilità di esaminare e verificare i processi decisionali sottostanti.
 - c. Sostenere la *ricerca e lo sviluppo di strumenti e metodologie per il risk management e l'auditing* nell'IA, al fine di garantire la sicurezza, l'affidabilità e la conformità **anche qualora non siano considerati ad alto rischio.** It would be functional to **expand transparency and publicity obligations to all AI systems in use by "public" entities** regardless of the

level of risk. The enrolment of all AI systems in use by the public, regardless of the degree of risk.

- d. Promuovere la formazione di personale specializzato nell'analisi e nella gestione dei rischi associati all'IA, per assicurare una valutazione competente e indipendente dei sistemi di IA. In questa ottica, il neo-istituto *European Centre for Algorithmic Transparency* (ECAT) rappresenta una direzione quanto più favorevole a tali punti.

3. Ecosistema di norme:

- a. La Digital Services Act (DSA) dell'UE rappresenta un passo importante verso la regolamentazione dell'IA generativa, incoraggiando il coinvolgimento delle comunità delle piattaforme online nel processo di monitoraggio e valutazione delle tecnologie IA.
- b. Il GDPR dell'UE, come testimoniato dal blocco di ChatGPT a livello nazionale da parte dell'Autorità Italiana per la Protezione dei Dati (Garante), è una base legale complementare per il trattamento dei dati che dovrebbe costituire, insieme alla prossima Data Act e Data Governance Act, il punto di riferimento normativo in cui i fornitori di IA generativa dovrebbero informare proattivamente le loro politiche di trattamento dei dati.

4. Asimmetrie nella distribuzione delle risorse e delle competenze: i paesi con una maggiore capacità tecnologica e finanziaria possono influenzare lo sviluppo e l'utilizzo dell'IA a scapito di quelli meno avanzati, aumentando le disuguaglianze e il rischio di colonialismo digitale. Per affrontare queste sfide, esortiamo l'Unione Europea a:

- a. Disincentivare meccanismi di competitività e sopraffazione in caso di momentanee sospensioni, incoraggiando la creazione di reti di collaborazione tra istituzioni di ricerca, aziende tecnologiche e governi, al fine di mitigare gli effetti delle rivalità geopolitiche sulla cooperazione nel settore dell'IA.

5. Colonialismo digitale: è quel fenomeno che va dall'estrazione all'appropriazione di dati e risorse digitali da parte di paesi e aziende più avanzati a scapito dei paesi in via di sviluppo. Le principali conseguenze e questioni etiche del

colonialismo digitale: sono lo **sfruttamento umano** - come il caso dei lavoratori e delle lavoratrici kenyote sfruttate da OpenAI a due dollari l'ora per allenare la loro IA; lo **sfruttamento economico** - l'accaparramento delle risorse digitali da un lato e l'impossibilità di sviluppare una propria industria locale; la **creazione di dipendenze tecnologiche** - il dominio di un solo modello di infrastrutture digitali crea delle dipendenze e impedisce lo sviluppo di architetture locali. Per contrastare il fenomeno, l'Unione Europea dovrebbe:

- a. Promuovere una regolamentazione equa e trasparente per la raccolta e l'utilizzo dei dati, garantendo il rispetto della privacy e dei diritti umani delle popolazioni coinvolte;
- b. Sostenere iniziative di cooperazione internazionale che favoriscano il trasferimento di tecnologie e competenze ai paesi in via di sviluppo, permettendo un apporto autonomo alle offerte dall'IA generativa;
- c. Incoraggiare la creazione di infrastrutture digitali e servizi di intelligenza artificiale accessibili e democratici, che possano garantire un maggiore controllo e autonomia ai paesi meno avanzati.
- d. La concentrazione del controllo sull'IA portano ad un controllo quasi monopolistico sul loro sviluppo, diffusione e l'implementazione. Le decisioni riguardanti gli standard etici, le priorità di ricerca e le applicazioni pratiche dell'IA tendono ad essere guidate dagli interessi di queste entità anziché dal consenso globale o dall'interesse comune.

6. Ridurre gli impatti complessivi dell'IA sul clima e definire requisiti di impatto per **le emissioni associate allo sviluppo e all'uso dell'IA**. Per agire in modo responsabile nel contesto dei cambiamenti climatici, è importante evitare un approccio "tecnosoluzionista" e non sovrastimare i costi ecologici dell'IA. Pertanto, è essenziale allineare l'intero ecosistema dell'IA con strategie di sostenibilità ambientale. I governi dovrebbero adottare regolamenti, strategie, finanziamenti e programmi di approvvigionamento di IA che **tengano conto dell'impatto climatico**, ad esempio evitando finanziamenti pubblici per applicazioni contrarie agli obiettivi climatici e acquistando servizi di IA solo da aziende impegnate in un obiettivo di emissioni zero;

- a. Assicurare un'adeguata inclusione del cloud computing nelle politiche di rendicontazione e prezzi del carbonio;
- b. Stabilire requisiti di rendicontazione e disponibilità dei dati per le emissioni del ciclo di vita dell'IA;

- c. Definire standard metodologici per le valutazioni di impatto ambientale a livello nazionale e internazionale.

7. Cooperazione internazionale per la governance dell'IA tra gli attori del settore, i regolatori e la società civile.

La natura pervasiva e transfrontaliera dell'IA richiede uno sforzo coordinato affinché le normative e le politiche adottate a livello globale siano quanto più allineate e complementari. Ciò può contribuire alla condivisione delle migliori pratiche tra i paesi, facilitando l'adozione di standard comuni e la creazione di un quadro normativo condiviso che tenga conto delle specificità culturali, economiche e sociali delle diverse regioni del mondo.

- a. Attraverso la partecipazione a iniziative e forum internazionali e la collaborazione con altre organizzazioni internazionali per promuovere l'adozione di principi comuni e linee guida per l'IA a livello globale.
- b. Pianificando task force dedicate che includano cittadini, società civile e stakeholder rilevanti per discutere e identificare possibili insidie sociali dell'IA e elaborare soluzioni democratiche e condivise specifiche.
- c. Inoltre, esistono anche altri meccanismi di notifica che consentono agli utenti di segnalare risultati falsi, pregiudiziali o dannosi creati dalle tecnologie IA, come ad esempio il meccanismo di segnalazione di Facebook o la piattaforma di segnalazione di Google per le informazioni errate. Tuttavia, questi meccanismi devono essere ulteriormente sviluppati e potenziati, come previsto nella prossima Digital Services Act dell'UE, per garantire una maggiore efficacia nella gestione dei rischi e delle sfide associate all'IA generativa.

Conclusione

L'Europa ha la possibilità di assumere un ruolo di leadership nella promozione di un approccio democratico e inclusivo allo sviluppo dell'IA, che garantisca equità e giustizia nella distribuzione dei benefici offerti da queste tecnologie. La promozione della trasparenza e dell'adeguata gestione dei rischi nell'IA generativa è fondamentale per assicurare la sicurezza, la stabilità e la fiducia dei cittadini nelle tecnologie di intelligenza artificiale.

In particolare, riteniamo che sia necessario affrontare questioni come il **colonialismo digitale e la concentrazione del controllo** sull'IA nelle mani di poche aziende

tecnologiche e nazioni, al fine di garantire l'equità e la giustizia nella distribuzione dei benefici dell'IA generativa. Inoltre, la **promozione di un approccio proattivo al risk management**, che coinvolga una vasta gamma di stakeholder nel processo decisionale, è fondamentale per prevenire possibili effetti negativi derivanti dall'impiego dell'IA generativa.

Per conseguire questi obiettivi, l'Unione Europea ha il potenziale di guidare lo sviluppo di un quadro normativo e di pratiche innovative che pongano al centro la responsabilità e l'etica nell'era dell'intelligenza artificiale. La cooperazione internazionale nella governance dell'IA è essenziale per garantire un approccio equilibrato, etico e sostenibile allo sviluppo e all'utilizzo delle tecnologie di intelligenza artificiale.

Pertanto, ci appelliamo ai Membri del Parlamento Europeo affinché promuovano un approccio collaborativo e inclusivo nella governance dell'IA generativa e nel garantire che le tecnologie di intelligenza artificiale siano utilizzate per il bene comune.