Traduzione a cura del Garante per la protezione dei dati personali

Strasburgo, 25 gennaio 2019 T-PD(2019)01

COMITATO CONSULTIVO (CD. T-PD) DELLA CONVENZIONE SULLA PROTEZIONE DELLE PERSONE RISPETTO AL TRATTAMENTO AUTOMATIZZATO DI DATI A CARATTERE PERSONALE
(Convenzione 108)

LINEE-GUIDA IN MATERIA DI INTELLIGENZA ARTIFICIALE E PROTEZIONE DEI DATI

Sistemi, software e dispositivi basati sull'intelligenza artificiale[1] ("IA") (di seguito "applicazioni IA") forniscono nuove e preziose soluzioni per affrontare i bisogni e le sfide in molti e differenti ambiti, quali la domotica, le *smart cities*, l'industria, la sanità e la prevenzione del crimine. Le applicazioni IA possono rappresentare uno strumento utile nei processi decisionali, specie nel supportare politiche inclusive e fondate su evidenze concrete. Come per altre innovazioni tecnologiche, queste applicazioni possono avere ripercussioni negative sugli individui e la società. Per evitare questo rischio, le Parti della Convenzione 108 garantiranno e assicureranno che lo sviluppo e l'utilizzo dell'IA rispettino e garantiscano i diritti alla tutela della vita privata e alla protezione dei dati personali (articolo 8 della Convenzione europea sui diritti dell'uomo), rafforzando così i diritti e le libertà fondamentali.
Le presenti linee guida forniscono una serie di misure di base che i governi, gli sviluppatori, i produttori e i fornitori di servizi di IA dovrebbero adottare per assicurare che le applicazioni IA non compromettano la dignità umana, i diritti e le libertà fondamentali di ogni individuo, in particolare con riferimento al diritto alla protezione dei dati.[2]
Le presenti linee-guida non intendono in alcun modo precludere o limitare le disposizioni della Convenzione europea dei diritti dell'uomo e della Convenzione 108. Le linee-guida tengono conto anche delle nuove tutele previste dalla Convenzione 108 modernizzata (più nota come "Convenzione 108+")[3] .

---

[1] La seguente definizione di IA è attualmente disponibile sul sito web del Consiglio d'Europa https://www.coe.int/en/web/human-rights-rule-of-law/artificial-intelligence/glossary: "Un insieme di scienze, teorie e tecniche il cui scopo è quello di riprodurre, attraverso la macchina, le capacità cognitive di un essere umano. Gli sviluppi attuali mirano, ad esempio, ad affidare a una macchina compiti complessi precedentemente delegati a un essere umano."

[2] Le presenti linee guida si basano sul Rapporto in materia di intelligenza artificiale ("Intelligenza artificiale e protezione dei dati: sfide e possibili rimedi") reperibile al link:
https://rm.coe.int/artificial-intelligence-and-data-protection-challenges-and-possible-re/168091f8a6

[3] Protocollo emendativo CETS n°223 della Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale.

I. Principi generali

1. La protezione della dignità umana e la tutela dei diritti umani e delle libertà fondamentali, in particolare il diritto alla protezione dei dati personali, sono essenziali nello sviluppo e nell'adozione di applicazioni IA che possono avere conseguenze sugli individui e la società. Ciò è particolarmente importante quando le applicazioni IA vengono utilizzate nei processi decisionali.
2. Lo sviluppo dell'IA fondato sul trattamento di dati personali dovrebbe basarsi sui principi della Convenzione 108+. Gli elementi chiave di questo approccio sono: liceità, correttezza, specificazione della finalità, proporzionalità del trattamento, protezione dei dati fin dalla progettazione (privacy by design) e protezione per impostazione predefinita (privacy by default), responsabilità e dimostrazione della conformità (accountability), trasparenza, sicurezza dei dati e gestione dei rischi.
3. Un'innovazione responsabile nel settore dell'IA necessita di un approccio incentrato sulla prevenzione e attenuazione dei potenziali rischi del trattamento dei dati personali.
4. In linea con le indicazioni in materia di valutazione del rischio fornite nelle linee-guida sui Big Data adottate dal Comitato della Convenzione 108 nel 2017[4], dovrebbe essere adottata una prospettiva più ampia quanto alle possibili conseguenze derivanti dal trattamento dei dati. Tale prospettiva dovrebbe considerare non solo i diritti umani e le libertà fondamentali, ma anche il funzionamento delle democrazie e i valori sociali ed etici.
5. Le applicazioni IA devono sempre rispettare pienamente i diritti degli interessati, in particolare alla luce dell'articolo 9 della Convenzione 108+.
6. Le applicazioni IA dovrebbero consentire un controllo significativo da parte degli interessati sul trattamento dei dati e sulle conseguenze correlate per gli individui e la società.

II. Linee-guida per sviluppatori, produttori e fornitori di servizi

1. Gli sviluppatori, i produttori e i fornitori di servizi di IA dovrebbero adottare un approccio orientato ai valori nella progettazione dei loro prodotti e servizi, in linea con la Convenzione 108+, in particolare con l'art. 10, par. 2 di quest'ultima, e con gli altri strumenti pertinenti del Consiglio d'Europa.
2. Gli sviluppatori, i produttori e i fornitori di servizi di IA dovrebbero valutare le possibili conseguenze negative delle applicazioni IA sui diritti umani e le libertà fondamentali e, alla luce di tali conseguenze, adottare un approccio precauzionale basato su appropriate misure di prevenzione e attenuazione dei rischi.
3. In tutte le fasi del trattamento, compresa la raccolta dei dati, gli sviluppatori, i produttori e i fornitori di servizi di IA dovrebbero adottare un approccio volto a tutelare i diritti umani fin dalla progettazione dei tali servizi ("human rights by design") ed evitare qualsiasi potenziale pregiudizio (*bias*), anche involontario o occulto, il rischio di discriminazione o altri effetti negativi sui diritti umani e le libertà fondamentali degli interessati.
4. Gli sviluppatori di IA dovrebbero vagliare accuratamente la qualità, la natura, l'origine e la quantità di dati personali utilizzati, riducendo i dati inutili, ridondanti o marginali durante lo sviluppo e le fasi di addestramento e poi monitorando l'accuratezza del modello man mano che viene alimentato con nuovi dati. L'uso di dati sintetici[5] può rappresentare una soluzione atta a

[4] https://rm.coe.int/t-pd-2017-1-bigdataguidelines-en/16806f06d04
[5] I dati sintetici sono generati da un modello di dati costruito su dati reali. Dovrebbero essere rappresentativi dei dati reali originali. Vedi la definizione di dati sintetici in OCSE. "Glossario dei termini statistici", 2007. http://ec.europa.eu/eurostat/ramon/coded_files/OECD_glossary_stat_terms.pdf ("Un approccio alla riservatezza

minimizzare la quantità di dati personali trattati dalle applicazioni IA.
5. Nello sviluppo e nell'utilizzo di applicazioni IA si dovrebbe tenere in adeguata considerazione il rischio di impatti negativi sugli individui e la società dovuto all'impiego di dati[6] e modelli algoritmici decontestualizzati[7].
6. Gli sviluppatori, i produttori e i fornitori di servizi di IA sono invitati a istituire e consultare comitati indipendenti di esperti provenienti da più ambiti, nonché a collaborare con istituzioni accademiche indipendenti, che possono contribuire alla progettazione di applicazioni IA fondate sui diritti umani e ispirate a considerazioni di natura etica e sociale, nonché all'individuazione di potenziali pregiudizi (*biases*). Tali comitati potrebbero svolgere un ruolo particolarmente importante in quei settori ove assicurare la trasparenza e il coinvolgimento degli interessati può risultare più difficile a causa di interessi e diritti confliggenti, come negli ambiti della giustizia predittiva, della prevenzione e repressione dei reati.
7. Dovrebbero essere incoraggiate forme partecipative di valutazione del rischio, basate sul coinvolgimento attivo degli individui e dei gruppi potenzialmente interessati dalle applicazioni IA.
8. Tutti i prodotti e servizi dovrebbero essere progettati in modo tale da garantire il diritto delle persone di non essere sottoposte a una decisione basata unicamente su trattamenti automatizzati che abbia effetti significativi su di esse, senza che le loro opinioni vengano prese in considerazione.
9. Al fine di rafforzare la fiducia degli utenti, gli sviluppatori, i produttori e i fornitori di servizi IA sono invitati a progettare i loro prodotti e servizi in modo da salvaguardare la libertà di scelta degli utenti rispetto all'utilizzo dell'IA, fornendo alternative praticabili alle applicazioni IA.
10. Gli sviluppatori, i produttori e i fornitori di servizi di IA dovrebbero adottare forme di vigilanza sugli algoritmi che promuovano la responsabilizzazione di tutte le parti interessate durante l'intero ciclo di vita di tali applicazioni, al fine di garantire l'osservanza dei principi e delle norme in materia di protezione dei dati e diritti umani.
11. Gli interessati dovrebbero essere informati se interagiscono con un'applicazione IA e hanno il diritto di ottenere informazioni sulla logica alla base dei trattamenti di dati che li coinvolgono. Le informazioni da fornire dovrebbero comprendere le conseguenze derivanti dall'applicazione di tale logica.
12. Dovrebbe essere garantito il diritto di opporsi al trattamento basato su tecnologie che influenzano le opinioni e lo sviluppo personale degli individui.

II. Linee-guida per legislatori e policy makers

1. Il rispetto del principio di responsabilizzazione (accountability), l'adozione di procedure di valutazione del rischio e l'applicazione di altre misure adeguate, come i codici di condotta e i meccanismi di certificazione, possono rafforzare la fiducia nei confronti di prodotti e servizi di IA.
2. Fatta salva la riservatezza tutelata dalla legge, le procedure di appalto pubblico dovrebbero imporre a sviluppatori, produttori e fornitori di servizi di IA specifici obblighi di trasparenza, una valutazione preliminare dell'impatto del trattamento dei dati sui diritti umani e sulle libertà fondamentali e l'obbligo di vigilanza sugli effetti negativi e le conseguenze potenzialmente derivanti dalle applicazioni IA (di seguito, "vigilanza sugli algoritmi"[8]).

---

secondo il quale invece di diffondere dati reali, vengono diffusi dati sintetici generati da uno o più modelli di popolazione") .

[6] Si tratta del rischio di ignorare le informazioni contestuali che caratterizzano le situazioni specifiche in cui si prevede di utilizzare le soluzioni basate sull'intelligenza artificiale.

[7] Tale rischio si materializza quando i modelli IA, originariamente progettati per un'applicazione specifica, vengono utilizzati in un contesto diverso o per finalità diverse.

[8] Sulla nozione di vigilanza sugli algoritmi, in quanto adozione di pratiche ispirate a responsabilizzazione, consapevolezza e gestione del rischio in rapporto a possibili conseguenze ed effetti negativi lungo l'intero ciclo di vita di queste applicazioni, si veda anche la 40ª Conferenza internazionale delle autorità di protezione dati, Dichiarazione

3. Le autorità di controllo dovrebbero essere dotate di risorse sufficienti per sostenere e monitorare i programmi di vigilanza sugli algoritmi posti in essere dagli sviluppatori, produttori e fornitori di servizi di IA.
4. L'eccessivo affidamento sulle soluzioni fornite dalle applicazioni IA e i timori di contestare le decisioni suggerite dalle applicazioni IA rischiano di alterare l'autonomia dell'intervento umano nei processi decisionali. Dovrebbero pertanto essere preservati il ruolo dell'intervento umano nei processi decisionali e la libertà dei decisori umani di non fare affidamento sulle raccomandazioni fornite attraverso l'impiego dell'IA.
5. Gli sviluppatori, i produttori e i fornitori di servizi di IA dovrebbero consultare le autorità di controllo quando le applicazioni di IA possono incidere in modo significativo sui diritti umani e sulle libertà fondamentali degli interessati.
6. Si dovrebbe promuovere la cooperazione tra le autorità di protezione dei dati e altri organismi che hanno una competenza in materia di IA, quali le autorità di tutela dei consumatori o della concorrenza, quelle incaricate del contrasto alle discriminazioni, le autorità di regolamentazione settoriali e le autorità di regolamentazione dei media.
7. Dovrebbero essere istituiti meccanismi adeguati al fine di garantire l'indipendenza dei comitati di esperti di cui alla sezione II.6.
8. Si dovrebbero garantire l'informazione e il coinvolgimento attivo di individui, gruppi e altre parti interessate nel dibattito sul ruolo da attribuire all'IA nel plasmare le dinamiche sociali e nei processi decisionali che li riguardano.
9. I policy maker dovrebbero investire risorse nell'alfabetizzazione digitale per aumentare la conoscenza e la comprensione degli interessati riguardo alle applicazioni IA e ai loro effetti. Dovrebbero promuovere, inoltre, la formazione professionale rivolta agli sviluppatori di IA al fine di aumentare la conoscenza e la comprensione dei potenziali effetti dell'IA sugli individui e la società. Dovrebbero sostenere la ricerca incentrata su un'IA orientata ai diritti umani.

---

sull'etica e la protezione dei dati nell'intelligenza artificiale, principio guida n. 2. Si veda anche il Rapporto sull'intelligenza artificiale (nota 2), sezione II.4