

I vincoli di Bruxelles sull'IA

Il regolamento Ue sull'intelligenza artificiale ne consente l'uso per la pubblicità commerciale, il merito creditizio, l'assistenza personale ecc.. A certe condizioni

Sdoganata l'intelligenza artificiale usata per la pubblicità commerciale mirata, che assegna un punteggio ai clienti. Via libera anche all'IA che valuta il merito creditizio, a quella con sembianze umane che tiene compagnia e a quella che influenza il comporta-

mento, ma a fin di bene, come far smettere di fumare o insegnare l'inglese. Sono alcune delle eccezioni al lungo elenco degli usi vietati dall'articolo 5 del regolamento Ue sull'IA.

Antonio Ciccia Messina alle pagine 4 e 5

Le indicazioni della Ue sulle pratiche vietate (e non) nelle Linee guida diffuse il 4 febbraio

Publicità mirata, assistenti e amici virtuali: sì all'uso dell'IA

ANTONIO CICCIA MESSINA

Sdoganata l'intelligenza artificiale (IA) usata per la pubblicità commerciale mirata, che assegna un punteggio ai clienti. Via libera anche all'IA che valuta il merito creditizio, a quella con sembianze umane che tiene compagnia e a quella che influenza il comportamento, ma a fin di bene, come far smettere di fumare o insegnare l'inglese.

Sono queste alcune delle eccezioni al lungo elenco dei sistemi vietati dall'articolo 5 del regolamento Ue sull'IA n. 2024/1689 (AI Act). A stilare l'elenco dei casi, in cui l'IA anche se rischiosa non è bandita, è la Commissione Ue con le Linee Guida approvate il 4 febbraio 2025, le quali, pur essendo un'interpretazione autorevole, non sono giuridicamente vincolanti.

A decorrere dal 2 febbraio 2025 sono operativi i divieti delle IA più invasive e potenzialmente dannose per l'uomo. Questi divieti si applicano a tutti i sistemi di IA, indipendentemente dal fatto che siano stati immessi sul mercato o messi in servizio prima o dopo il 2 febbraio 2025.

Le Linee Guida analizzano tutti i divieti (dalle pratiche manipolative al punteggio sociale, dal riconoscimento delle emozioni allo sfruttamento di situazioni di debolezza) e per ciascun gruppo di divieti ritagliano spazi di legittimità per le IA.

IA fuori ambito. Innan-

zi tutto, ci sono casi esclusi dall'ambito di applicazione del regolamento sull'IA. Si tratta delle IA destinate a scopi militari, di difesa o di sicurezza nazionale e a quelle di cui ci si avvale per le attività di cooperazione delle autorità di contrasto e giudiziarie. Si collocano fuori ambito di applicazione dell'AI Act anche le IA usate per scopi di ricerca scientifica e di ricerca, prova o sviluppo relativi a sistemi o modelli di IA. Non sono coperti dalle disposizioni del regolamento 2024/1689, infine, i sistemi di IA usati da persone fisiche nel corso di un'attività non professionale puramente personale e i sistemi di IA rilasciati con licenza libera e open source.

Tecniche pubblicitarie. Le tecniche pubblicitarie, che utilizzano l'intelligenza artificiale per personalizzare i contenuti in base alle preferenze dell'utente, non sono intrinsecamente manipolative, se non utilizzano tecniche subliminali, intenzionalmente manipolative o ingannevoli, tali da ribaltare l'autonomia individuale o sfruttare le vulnerabilità e debolezze delle persone.

Servizi bancari. I sistemi di IA utilizzati per la prestazione di servizi bancari, quali mutui ipotecari e prestiti, che elaborano dati (input) quali l'età o la situazione socioeconomica del cliente, non sono considerati sfruttamento delle vulnerabilità ai sensi del regolamento sull'IA, quando sono progettati per proteggere e so-

stenere le persone vulnerabili a causa della loro età, la disabilità o specifiche circostanze socioeconomiche.

Guida sicura. I sistemi di IA, che rilevano la sonnolenza e l'affaticamento dei conducenti e li avvisano di riposare nel rispetto delle leggi sulla sicurezza, sono utili e non sono considerate una forma di sfruttamento delle vulnerabilità delle persone. Inoltre, le informazioni raccolte dall'IA a proposito delle condotte di guida (eccessiva velocità, inosservanza di regole di prudenza) possono essere utilizzate per aumentare il premio dell'assicurazione contro i sinistri.

Amica del cuore. Una IA da compagnia, che appare con le fattezze di una persona umana e con la capacità di interagire su temi legati alla sfera affettiva, è legittima se tutto ciò è usato per rendere il sistema più attraente e coinvolgere più efficacemente, a condizione che non attui pratiche manipolative o ingannevoli, causa di probabili danni psicologici o fisici, di attaccamento patologico o dipendenza.

Vita sana. È legittima una chatbot terapeutica, che utilizza tecniche subli-



minali per indirizzare gli utenti verso uno stile di vita più sano e per abbandonare le cattive abitudini, come il fumo.

Certo, gli utenti, che seguono i consigli subliminali della chatbot, possono provare disagio fisico e stress psicologico a causa dello sforzo fatto per smettere di fumare. Peraltro, se non ci sono tentativi nascosti di influenzare il processo decisionale, salvo la promozione di abitudini sane, la chatbot non può essere considerata in grado di causare danni significativi.

Scudo contro hacker. Non sono bandite le tecniche manipolative e ingannevoli usate dall'intelligenza artificiale nella formazione sulla sicurezza e in simulazioni di apprendimento, le quali, ad esempio, imitano i tentativi di phishing per educare gli utenti sulle minacce alla sicurezza informatica.

Questi sistemi possono avvalersi di tecniche di manipolazione intenzionali (ad esempio, sfruttando i pregiudizi cognitivi), ma ciò viene fatto temporaneamente con l'obiettivo di addestrare e sensibilizzare e senza causare danni significativi.

Analisi delle emozioni. Costituisce persuasione legittima l'attività svolta da un sistema di intelligenza artificiale, che opera in modo trasparente e analizza le emozioni dei clienti per migliorare le interazioni e fornire assistenza specifica. Al contrario, commette una manipolazione un sistema occulto di riconoscimento delle emozioni dei consumatori, usato per offrire prodotti più costosi quando l'utente è più propenso ad acquistarli.

Studiare una lingua. Un sistema di IA, che, nell'assistere gli utenti ad apprendere più velocemente una lingua straniera, usa tecniche subliminali, non è manipolativo, se opera in modo trasparente e rispetta

scelta libera e informata dell'utente di acconsentire o meno all'uso del sistema.

Merito creditizio. Non incorrono nei divieti i sistemi di valutazione del merito creditizio, utilizzati da operatori finanziari e agenzie di informazioni creditizie per valutare l'affidabilità o i debiti insoluti di un cliente, fornire un punteggio di credito o determinare la valutazione del merito creditizio. Tali sistemi, che si basano sulle entrate e sulle spese del cliente e su altre circostanze finanziarie ed economiche, superano il vaglio, se sono conformi alle norme di protezione dei consumatori e degli utenti.

Anti frode. Le imprese hanno un interesse legittimo a valutare i clienti in funzione di contrasto alle frodi finanziarie. È legittimo, dunque, usare l'IA, purché la valutazione si basi su dati pertinenti.

Shopping online. Il punteggio assegnato con l'uso dell'IA da parte di una piattaforma di shopping online, al fine di offrire vantaggi ai clienti con un notevole frequenza di acquisti e un basso tasso di restituzione dei prodotti, non ricade nei divieti previsti dal Regolamento sull'IA: i vantaggi sono giustificati e proporzionati rispetto allo scopo di premiare il comportamento positivo e gli altri utenti non subiscono pregiudizio.

Emozioni commerciali. I sistemi di riconoscimento delle emozioni, utilizzati in un contesto commerciale per interagire con i clienti, non rientrano nelle IA vietate. I sistemi di IA, che consentono il riconoscimento delle emozioni basato sulla pressione di un tasto o sulla base di messaggi vocali dei clienti (ad esempio, messaggi di chat, uso di assistenti vocali virtuali), utilizzati nel marketing online per applicazioni per la visualizzazione di messaggi personalizzati e per scopi pubblicitari, anche in ambienti intelli-

genti (cartelloni pubblicitari intelligenti) non sono coperti dal divieto.

Sanzioni rinviate. Per i sistemi di IA vietati, l'articolo 99 dell'AI Act prevede sanzioni fino a 35 milioni di euro o, se superiore, fino al 7% del fatturato mondiale. Peraltro, ai sensi dell'articolo 113, l'applicazione delle sanzioni non scatta in contemporanea con l'applicazione dei divieti: se questi ultimi sono diventati operativi il 2 febbraio 2025, per le sanzioni bisognerà attendere il 2 agosto 2025 (si veda *ItaliaOggi* del 23 agosto 2024).

Mezzi di tutela. Nonostante questa partenza in due tempi, gli interessati non sono completamente disarmati. Il fatto che un sistema di IA sia estraneo ai divieti dell'articolo 5 non significa che sia liberalizzato. In effetti, continuano ad applicarsi i divieti derivanti da altre normative e, rimanendo all'interno del regolamento sull'IA, la maggior parte dei sistemi di IA, che rientrano in un'eccezione a un divieto del citato articolo 5, si qualificherà come ad alto rischio.

Questo significa che fornitori e utilizzatori dovranno osservare gli adempimenti previsti per i sistemi di alto rischio (articoli 6 e seguenti).

Inoltre, anche se le disposizioni in materia di sanzioni amministrative, si applicano solo in un secondo momento, i divieti sono già operativi e i soggetti interessati possono agire in giudizio per ottenere provvedimenti inibitori delle pratiche vietate. Inoltre, quando mediante i sistemi di IA vietati si trattano dati personali, gli interessati godono dei mezzi di tutela previsti dal regolamento Ue sulla privacy n. 2016/679. Infine, se le pratiche vietate provocano danni patrimoniali o non patrimoniali, il danneggiato può sempre agire in giudizio per il risarcimento.

— © Riproduzione riservata — ■

Gli strumenti di tutela contro le IA vietate

DS6901

DS6901

Divieti	Sono elencati all'articolo 5 dell'AI Act ed operativi dal 2 febbraio 2025
Sanzioni amministrative previste da AI Act	Possono arrivare fino a 35 milioni di euro o se superiore, fino al 7% del fatturato mondiale (operative dal 2 agosto 2025)
Tutele civili contro pratiche proibite	Azioni per ottenere provvedimenti inibitori
	Azioni per il risarcimento dei danni patrimoniali e non patrimoniali